Second International Conference on Research in Air Transportation

June 24-28, 2006
Belgrade, Serbia and Montenegro

Contents

Doctoral Symposium

Message from Program and Conference Chairs

Welcome to the Second International Conference on Research in Air Transportation!

On the behalf of the ICRAT 2006 Organisation Committee, we would like to express here our deep gratitude to the senior and young researchers in Air Transportation for having contributed to this young but challenging and exciting conference.

For this second edition of ICRAT, there were 82 qualified submissions by authors from 22 countries. The referee process resulted in 53 acceptances, for an acceptance rate of 64%, among which 33 submissions were selected as standard papers, and 20 as short papers, representing respectively 40% and 24%. All selected papers, long and short, are of good quality, and we are very proud of the professionalism of all authors, reviewers, and Program Committee members. Thank you so much for your contributions and collaborations.

This is also the first year that Tutorials, Workshops, and most importantly a Doctoral Symposium are included in the conference program. Four tutorials on the practice Air Transport are expected to bring up the understanding of how things work for the young scientists. The two workshops on the vision for the future are expected to create the forum for constructive discussions about research directions for our small research community. The Doctoral Symposium is expected to create a forum for young researchers to discuss their research approaches with senior researchers to obtain guidelines and supports. The program will be even more exciting with the four invited keynote speakers, all senior research scientists or strategists in Air Transportation. We are very grateful for their presence, contributions, and support.

The proceedings you are handling are the result of much hard work from many people. We would like to thank:

- The authors and co-authors of the paper submissions. They are, of course, what makes the conference program great.
- The invisible tertiary reviewers, who often supply the most expert and informed comments on their review, and the ICRAT'06 Scientific Program Committee. There were 46 members who had spent most of their free time during the referee process to review the submitted papers, and to return with careful comments. They are the guardians for the quality of the conference.
- The logistic team, also known as the conference secretariat team who worked hard to ensure the on-line processes with the authors, to collect, compile, and edit the final camera-ready proceedings.
- The Local Organising committee members and volunteers, for the local arrangements, the printing of the proceedings, and all the logistics at the conference place.
- The various institutions that provided the support for the paper process. The list includes the employers of all authors and co-authors; the employers of all reviewers and committee members; and the sponsors: EUROCONTROL, The Ministry of Science and Environmental Protection of Serbia Montenegro, The University of Belgrade - Faculty of Transport and Traffic Engineering, European Commission DG-TREN, and NASA.

Thank you all again, authors and reviewers, for your contribution to ICRAT'06 that surely be exciting. Thanks once more to the conference secretaries: Magnus Axholt, Ella Pinska, and Stephen Peterson to be the bridge between the authors, the Program Committee, and the Local Organisers. The success of this conference will be yours!

Vu Duong, Program Chair
Vojin Tosic, Conference Chair

Chairs and Committees

General Chair
Jean-Marc Garot — Ministry of Transport, FR

Co-chair
Vojin Tosic — Univ. Belgrade, YU

Past General Chairs
George Donohue — George Mason Univ., US
Antonin Kazda — Univ. Zilina, SK

Program Chair
Vu Duong — EUROCONTROL, FR

Local Organization Chair
Vojin Tosic — Univ. Belgrade, YU

Grants & Awards Chair
Christian Pusch — EUROCONTROL, FR

Tutorial & Summer School Chair
Marc Brochard — EUROCONTROL, FR

Publication Chair
Marc Bui — EPHE, FR

Internet Chair
Patrick Bellot — ENST, FR

Program Committee

Giovanni Andreatta	Univ. Padova, IT
Alexandre d'Aspremont	Princeton Univ., US
Obrad Babic	Univ. Belgrade, YU
Philippe Baptiste	Ecole Polytechnique, FR
Patrick Bellot	ENST, FR
Marc Bourgois	EUROCONTROL, FR
Marc Bui	EPHE, FR
John-Paul Clarke	MIT, US
Dominique Collin de Verdière	DGAC, FR
Kevin Corker	San José State Univ., US
Joseph Dodoo	UMES, US
George Donohue	George Mason Univ., US
Nicolas Durand	CENA/ENAC, FR
Tom Edwards	NASA, US
Laurent El Ghaoui	UC Berkeley, US
Heinz Erzberger	NASA Ames, US
Hartmut Fricke	TU Dresden, DE
Mark Hansen	UC Berkeley, US
John Hansman	MIT, US
Karel Havel	Univ. Zilina, SK
Ian Humphreys	Loughborough Univ., UK
Antonin Kazda	Univ. Zilina, SK
Peter Lindsay	Univ. Queensland, AU
Sandy Lozito	NASA Ames, US
Gört Ludtke	DLR, BE
John Lygeros	Univ. Patras, GR
Jean-Luc Marchand	Eur. Commission, BE
Bernard Miaillier	EUROCONTROL, BE
José Miguel de Pablo Guerrero	AENA, ES
Andy Price	NATS, UK
Amy Pritchett	Georgia Tech, US
Christian Pusch	EUROCONTROL, FR
Jozsef Rohacs	BUTE, HU
Lance Sherry	George Mason Univ., US
Ladislav Smrcek	Univ. Glasgow, UK
Monica Tavanti	DeepBlue, IT
Callum Thomas	MMU, UK
Uwe Völckers	DLR, DE

Session 1

Environment

Effect of Low Noise Arrival Procedures on External Cost around a Generic Airport

Torsten Bähr, Stefan Schwanke

Abstract— **Air transport is expected to grow the next decades constantly. Along with this growth also the environmental problems of this kind of transportation will increase such as exhaust gas and noise emission. Especially the airport surrounding neighbourhoods are affected by these issues as noise and gas emissions are very high close to the airport during departure and arrival periods.**

Concerning noise there are some principles to improve the existing or upcoming situation. One is the implementation of noise abatement approach and departure procedures. This study concerns the economic advantages of such approach procedures according to their noise emissions. Taking into account the social costs of aircraft noise around a generic airport, the impact of an A340 flying four different low noise approaches in a complete airport scenario on the resulting social costs is elaborated. Therefore the total annual, aircraft specific and marginal costs of the emitted noise using different evaluation modes such as the Hedonic House Price Method or an approach according to the European Directive 2002/49/EC on noise emissions are estimated.

The noise simulations of this study are accomplished for a generic airport based on the layout and the population distribution of Frankfurt airport. The flight mix is linked to the real Frankfurt flight route usage and the Airbus Global Market Forecast. Current situations are evaluated as well as realistic future scenarios for the years 2012 and 2022. The results have been validated as far as possible with data provided by German airport authorities and recent noise valuation studies from other parties.

Index Terms—**aircraft noise, external cost, noise abatement procedures, noise control**

I. Introduction

AIR transportation has been an important component of the mobility behaviour in the industrial nations for more than the last decade. Especially in the last years an increased demand for air passages goes along with an increasing offer, both at large international airports and at regional ones. However with this increasing traffic volume the problems and negative effects increase: at the airports themselves with capacity bottlenecks and resulting delays as for the adjacent municipalities with exhaust gas and noise emissions. A rising sensitisation within the population for these problems forces all participants of the aviation system, authorities and research institutes to focus on reducing these negative effects.

From the point of view of economic efficiency it is of paramount importance to quantify the costs and benefits of these improvements. Thus, aircraft operators are enable to evaluate the effects of low noise operational procedures and airports and authorities may balance noise restrictions. Furthermore the number of European airports that apply noise dependent landing charges provokes additional incentives.

Regarding the efforts of the authorities of the European Union and some national governments to take into account the external effects of transportation an estimation of the social costs of the aircraft noise seems to be reasonable. Principal problems in evaluating silence economically are the difficulties to convert an immaterial issue such as noise into monetary units. Noise causes a lot of negative effects on people, but neither they are paid for being exposed to noise nor are charged directly for living in a quiet neighbourhood. In the last decades there were different methodical approaches to evaluate the effects of aircraft noise on people. In comparison to former studies on aircraft noise which basically followed one approach only, the methodological framework of this study adapts and summarises the most reliable and important ones to calculate total annual, aircraft specific and marginal costs of aircraft noise. The final target is to internalise these social costs to reflect true and realistic market prices, e.g. of an air trip or air freight shipped.

The investigations of this study are conducted applying the noise prediction tool INM and focuses on the implementation of low noise arrival procedures, especially various continuous descent approach (CDA) procedures. The consequences of noise spreading and the resulting social costs are determined at a fictive airport which can represent every characteristic condition (layout, structure of the air routes, flight mix, social and environmental factors) of every kind of airport. It is not subjected to administrative and technical restrictions for design and development. As it is a very demanding topic, in this first step a quasi-generic airport based on the layout of Frankfurt airport with the fleet mix as a variable parameter for representing different and functions airport sizes is designed. Other factors are kept constant.

The goal of this approach is to find the impacts of the implemented low noise operational procedures on area and population in a complete realistic current and future airport scenario. The computations are processed with the planed noise metric LDEN of the European Union.

Manuscript received February 26, 2006.

T. Bähr is with the Technical University of Munich, Institute of Aeronautical Engineering, Boltzmannstraße 15, 85747 Garching, Germany (e-mail: torsten.baehr@web.de).

S. Schwanke is with the Technical University of Munich, Institute of Aeronautical Engineering, Boltzmannstraße 15, 85747 Garching, Germany (phone: +49-89-289-15988, e-mail: schwanke@tum.de).

II. Methodological Framework

The methodological framework which builds the underlying structure of the noise cost evaluation is shown in Figure A1. The framework consists of four items:

1) Noise abatement measure to choose
2) The INM model of the generic airport
3) TESCOAN for calculating social cost
4) Results and interpretation of cost calculation

This procedure is applied to economically evaluate various scenarios and the effects of noise abatement measures. In the first step the definition of the noise control measure to be analysed is conducted. Up to now fully developed INM models for two German airports – Frankfurt and Munich – are available and implemented in INM to calculate noise carpets. In case of availability of sufficient data the examination can be extended to other airports as well. Using the INM input, such as the flight mix, and output data, such as the noise contour area and population data, the self developed calculation tool TESCOAN generates total annual, aircraft specific and marginal social cost of aircraft noise.

A. Airport Model in INM

The generic airport designed for this study bases on Frankfurt airport as described in chapter 1. This model comprises the triple runway layout, SIDs and STARs according to current DFS standard and precise elevation data from DLR in an area 70 x 40 km around the airport reference point. The required population data includes about 2,7 million residents clustered with about 23.000 population points. Figure 1 provides an overview of the airport layout in INM.

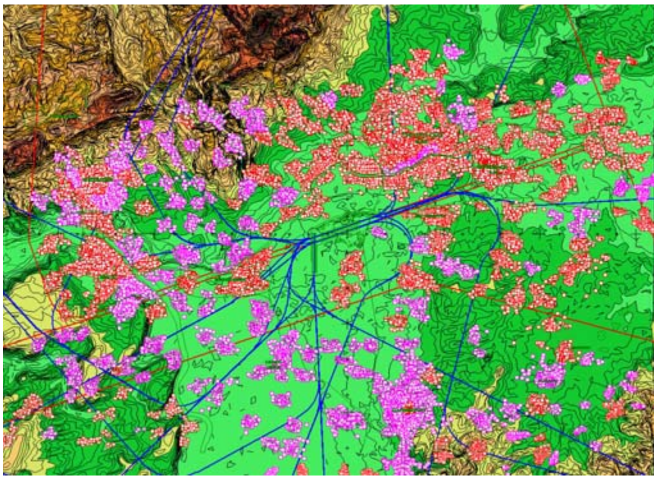


Figure 1: INM model of the used generic airport, source: [2]

The flight mix data is the variable parameter in the simulation. Baseline for the definition of the flight mix is the AIRBUS Global Market Forecast 2003 [1]. The aircraft mix and utilization drawn from there is related by percentage to the number of yearly movements of FRA. Considered are GMF[1] aircraft only for the year 2002 with a real day – evening – night distribution and runway use. The flight mix for the forecast scenario 2022 is defined according to the growth and utilisation prediction of the Airbus GMF for each seat segment. The values for the scenario 2012 are interpolated.

Table 1 indicates the distribution of movements for all three base scenarios considered, distinguished into day, evening, night time and total daily movements.

TABLE I
Day-evening-night movements per day for each scenario (only GMF A/C)

Year	Day time movements	Evening movements	Night time movements	Total movements
2002	778,15	138,22	164,35	1080,72
2012	1150,74	231,43	317,93	1700,10
2022	1514,82	305,42	471,39	2291,63

B. Implementation of Low Noise Operational Procedures

One major problem for airports today is the noise distribution of conventional aircraft approach procedures. For example, aircraft often descend to intermediate level segments in the range of 2000 to 4000 ft AGL, before capturing the glide slope and transitioning on the final descent to the runway. Such procedures result in spreading noise emissions onto nearby communities, sometimes as far away as 20 miles from the runway threshold.

Continuous descent approach (CDA) procedures have been proposed and introduced at some European airports especially at night to reduce noise emissions by delaying descent below 7000 ft as late as possible and descending incessantly at idle or near idle thrust from about 250 kts until final approach speed. Furthermore the landing gear is set as late as possible and the final approach configuration (flaps fully deployed) at approximately 5 nm to threshold.

In this study four different CDA procedures of an A340 with a reference procedure on their effects on the noise contour in the airport surroundings are compared. In table 2 the procedures are labeled and summarised corresponding to their main characteristics.

TABLE 2
Approach procedures in this study

Procedure name	category	characterisations	A/P
APP1	reference	2,5°/3,0° ILS 3000 ft	off
APP2	CDA	TSA 2,0°/4,0° 2900 ft	on
APP3	CDA	3° CDA	on
APP4	CDA	TSA 2,0°/4,0° 2900 ft	off
APP5	CDA	3° CDA	off

The reference procedure (APP1) is an approach with a first descent segment with a descent angle of 2,5° followed by a second descent segment with a descent angle of 3,0°. Between these two descent segments there is a horizontal flight segment at an altitude of 3000 ft. At this procedure the AP is deactivated.

[1] GMF aircraft include only commercial aircraft > 100 seats

The procedures APP2 and APP4 are "Two Segment Approaches" which have two descent segments: the first segment with a descent angle of 2,0° and the second segment with a descent angle of 4,0°. At an altitude of 2900 ft there is a direct transition from the first to the second segment without a horizontal flight segment. APP2 is flown with AP activated; APP4 is flown with AP deactivated.

APP3 and APP5 are CDA procedures with a sole descent angle of 3,0°. APP3 is flown with AP activated; APP5 is flown with AP deactivated.

Figure 2 shows the parameters altitude in [ft], true air speed (TAS) in [kts] and net thrust per engine in [lbs] vs. the distance to runway in [ft], providing the data needed for the implementation in the INM simulation.

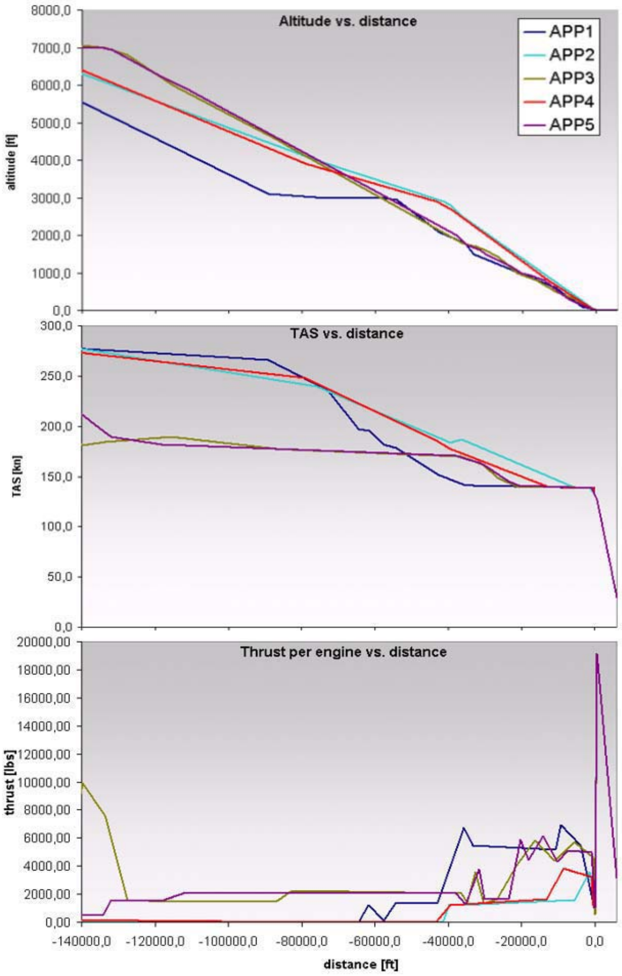


Figure 2: Parameters of the 5 approach procedures to be researched, source: [3]

These different approach procedures are implemented in INM with the fixed point profile method. It means that data triples (altitude, TAS, net thrust) are defined at certain distances from the runway threshold [3].

To evaluate the effects of these approach procedures on the noise zones each of the scenarios for the three years 2002, 2012 and 2022 is split into four groups of market penetration: 0 %, 40 %, 60 % and 100 % of the A340 flying the low noise approach procedure. 0 percent means that all A340 aircraft fly the reference procedure APP1; by contrast 100% means all A340 are flying a certain low noise procedure (APP2-APP5) independently of time of day. All other movements (see table 1) are left unchanged.

The reference procedure used is a little different to the INM standard approach in order to obtain a comparable data base in terms of aircraft performance and meteorological conditions.

C. Determination of External Cost of Aircraft Noise

The estimation of social cost of the emitted aircraft noise builds the major part of this analysis. The underlying software covers a wide spread range of approaches to calculate total annual, aircraft specific and marginal cost.

Most of the precedent studies on aircraft noise focused on one single method to convert noise into a monetary unit. There are two methods, which were frequently used in the past: the Contingent Valuation and the Hedonic House Price Method for the calculation of total annual social costs. The aim of the preceding document [4] was to research and implement other methodical approaches into the framework to calculate rapidly total annual, aircraft specific and marginal social cost of aircraft noise. As the theoretic background of the economic evaluation methods is elaborated in [4] this document concentrates on the results that derive from the methods.

The determination of the different cost categories is therefore assumed to be economically justified. To assess the load for the surrounding municipalities or a change of their situation through the employment of noise abatement measure total annual costs are evaluated. Aircraft specific or marginal costs can be considered to obtain the costs of each aircraft type. It has to be mentioned that the last two cost categories cannot be compared with one another because they do not contain the same information [4]. However, both categories are worth being considered.

The following section briefly describes the implemented cost calculating approaches.

1) Total annual costs

Total annual costs for a scenario are the baseline for further calculations of aircraft specific or marginal costs. Usually the reference period to evaluate external cost is one year. The methods in the framework are:

- Hedonic House Price Method HPM with a fix linear NDI [5, 6] – This is the most frequently used approach for calculating social costs of aircraft noise. It is a revealed preference method which uses surveys of the housing markets around the airport or meta-studies to calculate the monetary value NDI for the evaluation of the good "quiet" [4].
- Hedonic House Price Method HPM with a noise dependent, variable nonlinear NDI [4,7] – As the HPM with a fix linear NDI has the disadvantage of missing correlation between felt annoyance of noise and housing prices, a method was developed to consider this problem.

Result is a calculation approach with a variable nonlinear NDI linked to published community noise annoyance behaviour.

- Contingent Valuation CVM [8] – This method for estimation of social costs also applies the stated preference method.
- Approach according to the appendix to the EU directive 2002/49/EC (Experts [9] give the advice to pay 38,75€ (2000)/dB LDEN per household for a threshold at 50 or 55 dB(A) LDEN) [10]).
- External cost due to land use restrictions (loss of land value) – In some countries there are restrictions that land in noise contours above a certain limit is not applicable for residential construction.
- Costs due to the disturbance of night sleep – Monetary evaluation of the loss of sleep quality according to [11]. This issue is a very important one under research. There are a lot of quantitative research results on the effects of aircraft noise on the human body during night sleep.
- Calculation of health costs due to aircraft noise – Cost divided into three parts – resource costs, opportunity costs, costs due to a loss of utility –according to [10, 11]. Because of high noise thresholds (>70 LDEN) the resulting costs are very low or against zero.

2) Aircraft specific costs

Aircraft specific costs are the share of one aircraft of the calculated absolute external costs. The noise characteristic of the aircraft (noise certification level for T/O, S/L and APP) are used to derive aircraft specific costs for a movement or a LTO:

- European Landing Charges methodology (as in [12]) for departure, approaches and LTO – This methodology considers the noise thresholds at departures and at arrivals corresponding to categories of relatively quiet aircraft for the considered airport as well as the relative importance of noise emissions at arrivals and departures for the impacted population.
- Specific costs – cost doubling every 3 EPNdB (attempt using the noise energy – cost doubling when doubling of the noise energy) or 10 dB (attempt using the perceived sound intensity – cost doubling when doubling of the sound intensity), calculated for a LTO [13], for an approach only and for a departure only [4].

3) Marginal costs of one additional aircraft

Marginal costs mean the costs of one additional aircraft to an existing mix. Two approaches for calculating marginal costs are applied within this study [4]:

- Calculation of a reference scenario (noise and costs), calculation of noise contours of one annual movement of one aircraft (equivalent to 0,0027 movements per day), logarithmical summation of reference noise value and one additional aircraft noise value, computation of costs for the new scenario – marginal costs are the difference between new scenario and reference scenario.
- Calculation of a reference scenario (noise and costs), calculation of a scenario (noise and costs) with 365 additional aircraft (1 daily movement) – the marginal costs are the $1/365^{th}$ share of the difference between new scenario and reference scenario.

TESCOAN [4] is a self developed software for user friendly and fast calculation of total annual, aircraft specific and marginal social cost of aircraft noise. It requires input, such as population and operational data and output information, such as noise contours of INM to compute economic noise values using one or more of the calculation approaches explained before.

In this study total annual costs are estimated with the approaches HPM with a constant NDI (HP), the method according to the appendix of the EU directive (EU) and cost of illness (COI). Costs due to land use restrictions (LU) also are calculated but not validated as there is few information on current national and international noise thresholds. Nevertheless to compare the cost of annoyance of people with the effects on land use it is reasonable to compute them as well.

TABLE 3
PARAMETERS FOR CALCULATION OF SOCIAL COST OF AIRCRAFT NOISE

hedonic house price method		land use restrictions	
NDI [%/dB]	0,61	share of vacant land [%]	30
noise threshold [dB]	50/55	noise threshold [dB]	50/55
house price [€]	250.000	interest rate [%/a]	3,15
house life time [a]	25	development factor	0,2
interest rate [%/a]	3	value of farmland [€/m²]	8,5
size of household [persons/household]	2,2	value of housing land [€/m²]	60,0
		size of airport [km²]	14

Table 3 shows the parameters for the cost calculation with the HPM and the land use restriction method. The monetary value for the method according to the EU directive is 38,75 €/dB LDEN per household and noise thresholds are 50 and 55 dB LDEN. Parameters for the COI can be checked in [11].

Aircraft specific costs are estimated with the EU landing charges methodology and marginal cost will be derived with both methods explained (see paragraph 3).

III. RESULTS

A. Results of the noise calculation

The results of the noise calculation in INM provide important information. Figure A2 illustrates the noise zones for the three reference scenarios for the years 2002, 2012 and 2022. Table 4 indicates the size of the responding noise contour area while table 5 indicates the number of people affected in the noise contours 50, 55 and 60 dB LDEN for the 5 procedure scenarios with a market penetration of 100 %.

It is apparent that the size and number of affected persons increases with the years. There are significant differences between the five approaches within one year.

TABLE 4
SIZE OF NOISE CONTOURS 50, 55, 60 DB LDEN FOR MARKET PENETRATION OF 100 %, [KM²]

LDEN	APP1	APP2	APP3	APP4	APP5
			scenario 2002		
60,0	69,641	69,478	69,581	69,493	69,563
55,0	163,290	162,970	163,235	163,003	163,182
50,0	375,887	375,218	375,729	375,270	375,662
			scenario 2012		
60,0	85,903	85,442	85,709	85,479	85,687
55,0	208,494	207,506	208,224	207,609	208,107
50,0	470,742	468,756	470,092	468,865	469,955
			scenario 2022		
60,0	106,039	105,222	105,829	105,301	105,724
55,0	261,070	259,189	260,649	259,359	260,425
50,0	573,980	571,240	573,514	571,269	573,185

TABLE 5
AFFECTED PEOPLE IN NOISE CONTOURS OF 50, 55 AND 60 DB LDEN FOR A MARKET PENETRATION OF 100 %

LDEN	APP1	APP2	APP3	APP4	APP5
			scenario 2002		
60,0	9.161	8.735	8.892	8.735	8.892
55,0	94.607	93.510	94.405	93.747	94.317
50,0	306.777	305.499	306.379	305.630	306.672
			scenario 2012		
60,0	32.508	31.331	32.336	31.331	32.067
55,0	157.140	155.198	156.406	155.452	156.216
50,0	407.756	406.282	408.043	406.367	408.170
			scenario 2022		
60,0	55.031	53.638	54.723	53.726	54.454
55,0	208.300	205.354	207.636	205.712	207.338
50,0	511.131	509.573	512.352	509.794	512.340

Figure 3 indicates the absolute differences of affected people and area of the noise contours between case APP1 and the other four scenarios.

As presented in figure 3, APP2 and APP4 scenarios cover less area and less person are affected by a certain noise level. The investigation of the scenarios APP3 and APP5 did not lead to absolute consistent results: In all three annual scenarios these approach procedures also provoke less area covered by a certain noise level, but for the years 2012 and 2022 there are more people affected by a noise level of 50 dB LDEN or higher. However, in the contour 55 dB LDEN still less people are affected. This leads to the conclusion, that these two approach procedures are noise abating only within a certain distance to the airport. Because of the different flight altitude and noise emission angle noise might be emitted on a smaller, but a higher populated area so more persons are affected by noise emissions.

Thus diversity is to be expected concerning resulting costs between the scenarios.

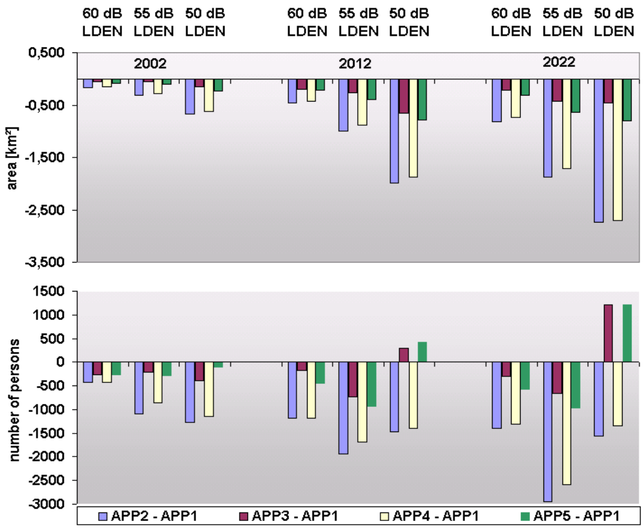


Figure 3: Changes of people affected and land covered at a certain noise level between APP1 and the four other approach procedures

B. Results of the social cost estimation

1) Total annual cost

a) Reference scenario

As the noise threshold for the calculation approaches for the costs of illness is 70 dB LDEN [10], the estimated costs (COI) are equal to zero and therefore disregarded. Table 6 indicates the total annual cost for the reference scenarios APP1 estimated with the other three calculation approaches. As mentioned the costs due to land use restrictions are calculated but are not validated. Therefore, they will not be considered further.

TABLE 6
TOTAL ANNUAL SOCIAL COSTS FOR THE REFERENCE SCENARIOS APP1 FOR THE YEARS 2002, 2012, 2022 [€/A]

APP1	55 dB LDEN	50 dB LDEN
	2002	
HP	9.079.209	46.493.534
EU	4.017.234	20.571.769
LU	14.153.550	35.007.402
	2012	
HP	17.899.555	70.763.219
EU	7.919.929	31.310.259
LU	18.587.655	44.311.824
	2022	
HP	28.342.581	96.171.113
EU	12.540.604	42.552.367
LU	23.744.888	54.438.543

It indicates an increase of the social costs with the years, which is caused by higher traffic volume. Furthermore the appearance and integration of quieter aircraft will not compensate this higher traffic. Also it indicates an increase of the social costs for lower noise thresholds, which means that more persons or more land are affected.

b) Low noise approach procedure scenario

As the results of both HPM and EU depend on the noise level that people are exposed to, there is a direct relation between these two costs. This is the reason for giving the results below only for the EU approach.

There are different possibilities to determine the effects of low noise procedures on the total annual cost:

1. Effects of different market penetrations within one year

Total annual cost of three different market penetrations are compared within one annual scenario. In this case the costs are compared for the noise thresholds 50 and 55 dB LDEN for the scenario 2002 calculated with the EU approach.

This results in a decrease of the total costs with an increase of the market penetration in comparison with APP1 for all four low noise approaches as (table 7). There is also diversity in the relative differences between the effects on the costs for the two noise thresholds. I.e. for 50 dB LDEN the relative difference is higher than for the 55 dB LDEN threshold.

TABLE 7
EFFECT OF DIFFERENT MARKET PENETRATIONS ON TOTAL ANNUAL COSTS, HERE FOR EU APPROACH, SCENARIO 2002 [€/A]

m. p.	APP2	APP3	APP4	APP5
55 dB LDEN				
40 %	3.995.909	4.013.087	3.997.621	4.008.353
60 %	3.986.693	4.010.438	3.988.973	4.004.201
100 %	3.963.702	4.004.390	3.968.007	3.998.690
50 dB LDEN				
40 %	20.524.499	20.561.233	20.529.107	20.554.217
60 %	20.498.769	20.555.989	20.505.385	20.545.176
100 %	20.453.006	20.545.317	20.462.843	20.529.419

m. p. – market penetration

2. Effects of one market penetration over the respective years

The total annual costs of three annual scenarios within one market penetration are compared. In this case the costs are compared for the noise thresholds 50 and 55 dB LDEN for the scenario of 100 % market penetration calculated with the EU approach.

This results in a decrease of the total costs for all years compared with APP1 (table 8). There is also an increase of the relative differences with the years. It is apparent, that for the noise threshold 50 dB LDEN the relative difference is higher than for the 55 dB LDEN.

2) Aircraft specific cost

a) Reference scenario

As the flight mix for the complete case covers up to 61 aircraft the exposition of the result is limited to 10 characteristic aircraft as presented in table A1 where an overview of the aircraft specific cost for a LTO applying the EU landing charge proposal [12] is given.

TABLE 8
EFFECT OF DIFFERENT ANNUAL SCENARIOS ON TOTAL ANNUAL COSTS, HERE FOR EU APPROACH, SCENARIO 100 % [€/A]

n. t.	APP2	APP3	APP4	APP5
2002				
55	3.963.702	4.004.390	3.968.007	3.998.690
50	20.453.006	20.545.317	20.462.843	20.529.419
2012				
55	7.743.705	7.883.614	7.758.632	7.855.446
50	31.004.509	31.259.415	31.035.601	31.219.043
2022				
55	12.207.381	12.467.360	12.233.953	12.420.240
50	42.036.245	42.473.678	42.077.785	42.407.195

n. t. – noise threshold in dB LDEN

b) Low noise approach procedure scenario

Taking a look on the changes of the aircraft specific cost by applying the low noise approaches, two ways to evaluate the effects of these procedures are considerable:

1. split up the cost benefit on all aircraft represented in the mix:

As indicated in table A1 the cost benefits can be split up on all aircraft. The changes of these aircraft specific costs are rather small. It might be reasonable to ask, whether all aircraft – also the ones which do not fly low noise procedures – should be considered by the cost-benefit-analysis.

2. split up the cost benefit only on the aircraft flying a low noise approach procedure:

Using this approach just means giving the financial benefit to the A340 that have flown the low noise approaches because these are the aircraft that decrease the total cost.

TABLE 9
BENEFIT OF ONE A340 APP, HERE FOR EU APPROACH, [€/MOVEMENT]

m. p.	APP2	APP3	APP4	APP5
2002 55 dB LDEN				
40%	4,25	0,83	3,91	1,77
60%	9,13	2,03	8,45	3,90
100%	26,67	6,40	24,53	9,24
50 dB LDEN				
40%	9,42	2,10	8,50	3,50
60%	21,82	4,72	19,85	7,95
100%	59,17	13,18	54,27	21,10
2012 55 dB LDEN				
40%	6,07	1,24	5,47	2,09
60%	13,63	2,93	12,48	4,76
100%	37,27	7,68	34,12	13,64
50 dB LDEN				
40%	10,13	1,86	9,11	3,13
60%	23,09	3,91	20,85	6,92
100%	64,67	10,75	58,09	19,29

m. p. – market penetration

The advantages of the approach procedures are more clearly identifiable applying the second procedure. Its results for costs calculated with the EU approach for the years 2002 (2007 A340 arrivals) and 2012 (4728 arrivals) are indicated in table 9.

There are significant differences between the APP2 and APP4, which obviously are more noise abating approaches than APP3 / 5. The benefit per movement increases with market penetration increased. It is remarkable that in spite of more than twice as much A340 movements in the 2012 scenarios the benefits are higher than in the 2002 scenarios.

3) Marginal cost

Another approach to compare the differences of the approach procedures on the social costs is to estimate marginal costs. In comparison to aircraft specific costs marginal costs contain a much more parameter that might be change: i.e. flight track, flight time and runway.

Marginal costs are estimated with both methods TESCOAN offers. As both approaches provide rather similar results, only costs estimated with the method "one additional aircraft a year" are specified in this study for the noise threshold 50 dB LDEN. These costs are split into day, evening and night for approaches on runway 07L and 25R.

a) Reference approaches

Using the reference approach procedure APP1 the reference marginal costs for an additional A340 landing is determined for the comparison with the other procedures (table 10).

TABLE 10
MARGINAL COSTS FOR THE REFERENCE APPROACH PROCEDURES [€/MOVEMENT]

	07L day	07L evening	07L night	25R day	25R evening	25R night
2002	38,09	190,41	380,89	39,70	198,45	396,96
2012	33,91	169,53	339,10	33,69	168,44	336,89
2022	29,60	147,96	296,01	28,35	141,72	283,47

Two phenomena can be identified in this table: on the one side the marginal cost decrease with the year and higher traffic volume. On the other side the marginal costs in the evening are higher than at day time but lower than at night time according to the LDEN weighting.

b) Low noise approach procedure scenario

Comparing the marginal costs of the four low noise approaches to the reference ones, it is clearly identifiable which approach is a low noise one and which is less noise abating. APP2 seems to be the most quite approach procedure according to the social costs of the noise emitted. According to the costs, APP3 is the procedure with the least improvement on noise emission.

In comparison with the reference approaches there are changes in social marginal costs up to almost 40 %. There is a difference of costs between approaches on runway 07L and 25R. This is reasoned by the different distribution of population around the airport. That is why not the same amount of people is affected by the additional arrival.

TABLE 11
MARGINAL COSTS FOR THE LOW NOISE APPROACH PROCEDURES [€/MOVEMENT]

	07L day	07L evening	07L night	25R day	25R evening	25R night
			APP2			
2002	23,31	116,60	233,13	26,63	133,13	266,29
2012	21,96	109,82	219,62	24,56	122,82	245,65
2022	20,10	100,53	201,02	21,68	108,36	216,75
			APP3			
2002	34,47	172,38	344,64	37,47	187,35	374,65
2012	31,64	158,24	316,40	32,27	161,36	322,72
2022	27,95	139,80	279,54	27,52	137,60	275,19
			APP4			
2002	24,95	124,71	249,53	27,78	138,94	277,82
2012	23,35	116,68	233,45	25,34	126,72	253,40
2022	21,16	105,79	211,62	22,22	111,12	222,22
			APP5			
2002	32,09	160,49	320,93	35,68	178,35	356,78
2012	29,62	148,14	296,22	31,26	156,31	312,60
2022	26,46	132,31	264,60	26,78	133,87	267,76

IV. CONCLUSION

Summarizing this study indicates the advantages of noise abatement approach procedures for the adjacent municipalities taking into account the social costs of the emitted aircraft noise.

The estimated total annual, aircraft specific and marginal costs of the emitted aircraft noise determined in this study indicate the same result: The procedures APP2 and APP4 as two segment approaches are quieter than APP3 and APP5 as 3° CDA approaches. The differences of costs are rather high. APP2 is the most noise efficient approach, APP4 the second, APP5 the third and APP3 is few efficient of these approach procedures. There is no direct influence of using the auto pilot as APP2 (AP on) is less noisy than APP4 (AP off) but APP3 (AP on) is more noisy than APP5 (AP off); nevertheless there is a cost difference of up to 10% between same procedures (AP on and AP off).

The study shows that marginal social costs are suitable for the evaluation of effects of noise abating operational procedures. One conclusion that derives from this study is for example that the positive effect of those procedures is much higher in the evening and at night time. As the costs estimated are directly related to the person's annoyance of aircraft noise, it is much easier to discuss the sense of flying noise abating approach procedures at least at night or during the evening hours using these results. On the other side it is difficult to evaluate the effect of such procedures using aircraft specific costs as they are estimated with the noise certification levels according to the ICAO. Flying a noise abatement procedure does not change these noise levels. Thus getting an idea on the influence of flying low noise approaches on the whole scenario it is reasonable to take these costs into account. To categorize these procedures this kind of cost is not very efficient. Total costs give a good impression on the whole scenario as well, but cannot evaluate well the quality of a low noise approach.

The comparison of the increase of costs of annoyance and costs of land use indicates that there is no proportional connection between persons and land affected. Therefore a calculation of both approaches is reasonable.

This investigation indicates that marginal costs seem to be the most effective approach to evaluate the efficiency of flying noise abatement approach procedures. It also points out differences of the impact of flying such procedures during different periods of the day.

APPENDIX

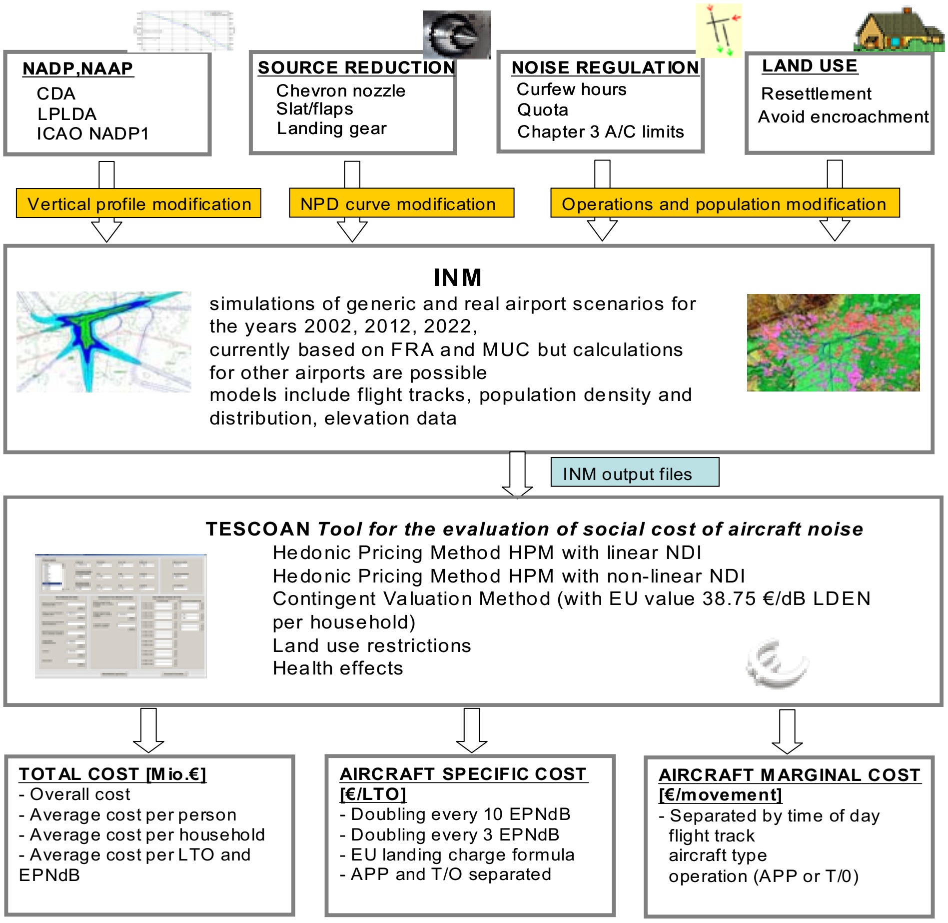


Figure A1: Methodological framework to estimate the economical influence of noise abatement measures

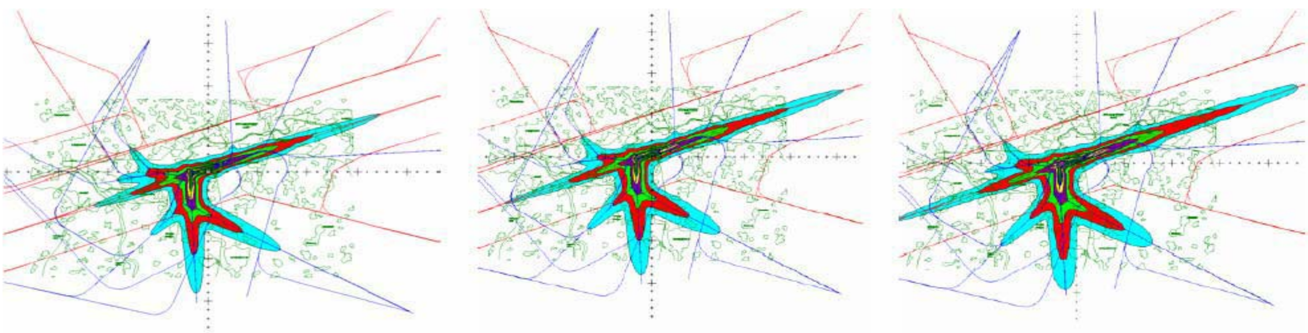


Figure A2: Noise contours of reference scenarios 2002 (left), 2012(middle) and 2022 (right) (LDEN > 50 in 5dB steps)

TABLE A1

AIRCRAFT SPECIFIC COSTS FOR EU APPROACH, REFERENCE AND LOW NOISE SCENARIOS

cost [€/LTO]	Ref	40%				60%				100%			
	APP1	APP2	APP3	APP4	APP5	APP2	APP3	APP4	APP5	APP2	APP3	APP4	APP5
EU 2002 55													
A30062	30,24	30,08	30,21	30,10	30,18	30,01	30,19	30,03	30,15	29,84	30,15	29,87	30,10
A320	13,92	13,85	13,91	13,85	13,89	13,81	13,90	13,82	13,88	13,73	13,88	13,75	13,86
A33034	21,53	21,42	21,51	21,43	21,48	21,37	21,49	21,38	21,46	21,24	21,46	21,27	21,43
A34030	30,84	30,67	30,81	30,69	30,77	30,60	30,79	30,62	30,74	30,43	30,74	30,46	30,70
B737300	16,31	16,22	16,29	16,23	16,28	16,19	16,28	16,20	16,26	16,09	16,26	16,11	16,24
B737700	10,75	10,69	10,74	10,69	10,72	10,66	10,73	10,67	10,71	10,60	10,71	10,61	10,70
B747400	112,36	111,77	112,25	111,81	112,11	111,51	112,17	111,57	112,00	110,86	112,00	110,98	111,84
B757RR	8,33	8,29	8,32	8,29	8,31	8,27	8,32	8,27	8,30	8,22	8,30	8,23	8,29
B767400	25,83	25,69	25,81	25,71	25,77	25,64	25,79	25,65	25,75	25,49	25,75	25,52	25,71
F10062	5,65	5,62	5,64	5,62	5,64	5,61	5,64	5,61	5,63	5,57	5,63	5,58	5,62
EU 2002 50													
A30062	154,87	154,52	154,79	154,55	154,74	154,32	154,76	154,37	154,67	153,98	154,68	154,05	154,56
A320	71,28	71,12	71,25	71,14	71,22	71,03	71,23	71,05	71,19	70,87	71,19	70,91	71,14
A33034	110,26	110,00	110,20	110,03	110,16	109,87	110,17	109,90	110,11	109,62	110,12	109,67	110,03
A34030	157,92	157,56	157,84	157,59	157,78	157,36	157,80	157,41	157,71	157,01	157,71	157,08	157,59
B737300	83,53	83,34	83,49	83,35	83,46	83,23	83,46	83,26	83,42	83,05	83,42	83,09	83,36
B737700	55,03	54,90	55,00	54,92	54,98	54,83	54,99	54,85	54,96	54,71	54,96	54,74	54,92
B747400	575,39	574,07	575,10	574,20	574,90	573,35	574,95	573,53	574,65	572,07	574,65	572,34	574,21
B757RR	42,66	42,56	42,63	42,57	42,62	42,50	42,62	42,52	42,60	42,41	42,60	42,43	42,57
B767400	132,28	131,98	132,21	132,01	132,17	131,81	132,18	131,85	132,11	131,52	132,11	131,58	132,01
F10062	28,92	28,86	28,91	28,86	28,90	28,82	28,90	28,83	28,89	28,76	28,89	28,77	28,86
EU 2012 55													
A30062	51,74	51,27	51,64	51,31	51,57	51,03	51,59	51,09	51,49	50,58	51,50	50,68	51,31
A320	22,84	22,63	22,79	22,65	22,76	22,53	22,77	22,55	22,73	22,33	22,73	22,37	22,65
A33034	33,77	33,47	33,71	33,50	33,67	33,31	33,67	33,35	33,61	33,02	33,62	33,08	33,50
A34030	46,79	46,37	46,71	46,41	46,65	46,16	46,66	46,21	46,57	45,75	46,58	45,84	46,41
B737300	30,46	30,19	30,41	30,22	30,37	30,05	30,38	30,09	30,32	29,79	30,33	29,84	30,22
B737700	18,22	18,05	18,18	18,07	18,16	17,97	18,16	17,99	18,13	17,81	18,13	17,85	18,07
B747400	180,71	179,07	180,37	179,23	180,15	178,26	180,18	178,47	179,85	176,69	179,88	177,03	179,24
B757RR	14,25	14,12	14,22	14,13	14,20	14,06	14,21	14,07	14,18	13,93	14,18	13,96	14,13
B767400	42,09	41,71	42,02	41,75	41,96	41,52	41,97	41,57	41,89	41,16	41,90	41,24	41,75
F10062	9,53	9,45	9,51	9,45	9,50	9,40	9,50	9,41	9,49	9,32	9,49	9,34	9,45
EU 2012 50													
A30062	204,53	203,75	204,39	203,83	204,29	203,34	204,33	203,46	204,17	202,53	204,20	202,74	203,94
A320	90,28	89,93	90,22	89,97	90,17	89,75	90,19	89,80	90,12	89,40	90,13	89,49	90,02
A33034	133,51	133,00	133,42	133,05	133,35	132,74	133,38	132,81	133,28	132,21	133,30	132,34	133,12
A34030	184,99	184,29	184,86	184,36	184,77	183,92	184,81	184,02	184,67	183,19	184,69	183,37	184,45
B737300	120,44	119,98	120,35	120,02	120,30	119,74	120,32	119,81	120,23	119,26	120,24	119,38	120,09
B737700	72,02	71,74	71,97	71,77	71,93	71,60	71,95	71,64	71,89	71,32	71,90	71,39	71,81
B747400	714,40	711,67	713,90	711,95	713,56	710,25	713,70	710,65	713,16	707,43	713,24	708,14	712,32
B757RR	56,33	56,12	56,29	56,14	56,27	56,01	56,28	56,04	56,23	55,78	56,24	55,84	56,17
B767400	166,41	165,78	166,30	165,84	166,22	165,45	166,25	165,54	166,12	164,79	166,14	164,95	165,93
F10062	37,68	37,54	37,66	37,55	37,64	37,46	37,65	37,48	37,62	37,31	37,62	37,35	37,57
EU 2022 55													
A30062	64,67	63,95	64,49	64,02	64,39	63,64	64,41	63,72	64,28	62,95	64,29	63,09	64,05
A320	28,45	28,13	28,37	28,16	28,33	28,00	28,33	28,03	28,28	27,69	28,28	27,75	28,17
A33034	41,91	41,44	41,80	41,49	41,73	41,24	41,74	41,30	41,66	40,80	41,67	40,89	41,51
A34030	57,90	57,25	57,74	57,32	57,66	56,98	57,67	57,06	57,55	56,36	57,56	56,48	57,34
B737300	38,33	37,91	38,23	37,95	38,17	37,72	38,18	37,77	38,10	37,31	38,11	37,40	37,96
B737700	22,75	22,50	22,69	22,53	22,66	22,39	22,66	22,42	22,62	22,15	22,62	22,20	22,54
B747400	224,74	222,23	224,13	222,48	223,79	221,17	223,84	221,46	223,40	218,76	223,42	219,24	222,58
B757RR	17,81	17,61	17,76	17,63	17,74	17,53	17,74	17,55	17,70	17,34	17,71	17,38	17,64
B767400	52,41	51,82	52,27	51,88	52,19	51,58	52,20	51,64	52,10	51,02	52,10	51,13	51,91
F10062	11,90	11,77	11,87	11,78	11,85	11,71	11,85	11,73	11,83	11,58	11,83	11,61	11,79
EU 2022 50													
A30062	219,43	218,34	219,26	218,45	219,11	217,85	219,18	217,99	218,99	216,77	219,02	216,98	218,68
A320	96,53	96,05	96,45	96,10	96,39	95,83	96,42	95,89	96,33	95,35	96,35	95,45	96,20
A33034	142,21	141,51	142,10	141,58	142,00	141,19	142,05	141,28	141,92	140,48	141,94	140,62	141,72
A34030	196,46	195,49	196,31	195,59	196,18	195,06	196,24	195,18	196,07	194,08	196,10	194,27	195,79
B737300	130,07	129,43	129,97	129,49	129,88	129,14	129,92	129,22	129,81	128,49	129,83	128,62	129,63
B737700	77,21	76,83	77,15	76,87	77,10	76,66	77,12	76,70	77,05	76,27	77,07	76,35	76,95
B747400	762,57	758,80	761,98	759,18	761,48	757,10	761,70	757,59	761,04	753,32	761,16	754,06	759,97
B757RR	60,44	60,14	60,39	60,17	60,35	60,00	60,37	60,04	60,31	59,70	60,32	59,76	60,23
B767400	177,83	176,95	177,69	177,04	177,58	176,56	177,63	176,67	177,47	175,67	177,50	175,85	177,22
F10062	40,38	40,18	40,35	40,20	40,32	40,09	40,34	40,12	40,30	39,89	40,31	39,93	40,24

ACKNOWLEDGMENT

The authors want to thank the students participating in this study. Mr. Bähr especially wants to thank Mr. Schwanke for his unflappably help and support preparing this study.

REFERENCES

[1] Airbus S.A.S., *Global Market Forecast 2003 – 2022*, France, 2003
[2] S. Schwanke*, Effect of noise reduction technologies and operational procedures on external cost around a generic airport,* INTERNOISE 2005, 07.-10.08.05, Rio de Janeiro, Brazil IN05-1841
[3] C. Dimter, *Integration und Analyse von gegenwärtigen und zukünftigen lärmarmen vertikalen Flugprofilen anhand verschiedener Szenarien am Flughafen Frankfurt*, diploma thesis TU Munich, 2005
[4] T. Baehr, *Bestimmung ausgewählter externer Kosten des Fluglärms bei Flughafenanwohnern an einem generischen Flughafen*, diploma thesis TU Munich/TU Dresden, 2005
[5] P. Morell, C.H.-Y., *Aircraft noise social cost and charge mechanism – A case study of Amsterdam Airport Schiphol*, Air Transport Group, College of Aeronautics, Cranfield University, Cranfield 2000
[6] B. Pearce, D. Pearce, *Setting environmental taxes for aircraft:- a case study of the UK*, CSERGE working paper GEC 2000-26, 2000
[7] H.M.E. Miedema, C.G.M. Oudshoorn, *Annoyance from Transportation Noise: Relations with Exposure Metrics DNL and DENL and their Confidence Intervals*, in: Environmental Health Perspectives, Volume 109, Number 4, April 2001
[8] S. Navrud, *The State-of-The-Art on Economic Valuation of Noise*. Final Report to the European Commission DG Environment, http://europa.eu.int/comm/environment/noise/020414noisereport.pdf
[9] WG HSEA, *Valuation of Noise*, Draft Position Paper of the WORKING GROUP on HEALTH and SOCIO-ECONOMIC ASPECTS. Brussels, 2003
[10] S.A. Schmid, P. Preiss, A. Gressmann, R. Friedrich, *Ermittlung externer Kosten des Flugverkehrs am Flughafen Frankfurt/Main – Gutachten zur Ermittlung der externen Kosten des Flugverkehrs am Flughafen Frankfurt/Main*, Universität Stuttgart, Institut für Energiewirtschaft und Rationelle Energieanwendung, Stuttgart 2003
[11] Y.de Kluizenaar, W. Paschier-Vermeer, H.M.E. Miedema, *Adverse effects of noise exposure on health – a state of the art summary*, TNO report 2001.171, Leiden, Netherlands, 2001
[12] EU COM(2002) 683 final, *EU Directive on establishment of a community framework for noise classification of civil subsonic aircraft for the purposes of calculating noise charges*, Brussels, 2002
[13] R. Schlenker, The *Social Cost of Aircraft Noise – An Analysis of the Situation at Fourteen Airports Worldwide*, diploma thesis, TU Munich, 2004

Torsten Bähr, born on January13th 1978 in Rochlitz, Germany studied Transportation Engineering at the University of Technology Dresden, Germany. He holds the degree of a Diplom Engineer of Transportation since May 2005. His diploma thesis deals with the effect of aircraft noise on the social cost of aircraft noise.

He works as a part time research fellow at the Institute of Aeronautical Engineering at the Technical University of Munich, Germany since May 2005. Before he worked at the GfL, the Organization for Air Traffic Research, Berlin and at the Institute of Aviation at the Technical University of Dresden as a student research fellow.

Stefan Schwanke, born on August27th 1972 in Osterhofen, Germany studied Aerospace Engineering at the University of Stuttgart and at the Technical University Munich, Germany. He holds the degree of a Diplom Engineer of Aerospace and Aeronautics since July 2001.

He works as a research fellow at the Institute of Aeronautical Engineering at the Technical University of Munich, Germany since August 2001, where he is writing a PhD-Thesis dealing with the economic valuation of noise.

ABBREVIATION

A/C	Aircraft
AGL	Above Ground Level
AP	Auto Pilot
APP	Approach
CDA	Continuous Descent Approach
CVM	Contingent Valuation Method
DFS	Deutsche Flugsicherung (German air traffic control)
DLR	German Aerospace Center
DME	Distance Measurement Equipment
EPNdB	Decibels of Effective Perceived Noise Level
EPNL	Effective Perceived Noise Level
EU	European Union
FRA	Frankfurt Rhein- Main airport
GMF	Global Market Forecast
HPM	Hedonic Pricing Method
ICAO	International Civil Aviation Organization
INM	Integrated Noise Model
LDEN	Day Evening Night Level
LPLDA	Low Power Low Drag Approach
LTO	Landing and Take-Off cycle
m. p.	Market Penetration
MUC	Munich airport
n. t.	Noise Threshold
NAAP	Noise Abatement Approach Procedure
NADP	Noise Abatement Departure Procedure
NDI	Noise Depreciation Index
NPD	Noise Power Distance Curves
S/L	Side Line
SID	Standard Instrument Departure
STAR	Standard Terminal Arrival Route
TESCOAN	Tool for the Evaluation of Social Cost of Aircraft Noise
TAS	True Airspeed
T/O	Take Off
TSA	Two Segment Approach

Air Traffic Assignment As A Noise Abatement Measure Case Study: Zurich Airport, Switzerland

Fedja T. Netjasov

***Abstract*— One of the biggest problems of modern airports is the noise generated by air traffic, and the impact of this noise on those living near the airport. Noise is an unavoidable consequence of air traffic but it can be decreased by numerous measures: technical innovations in aircraft design, legislation, etc. This paper presents a suggested measure which has been developed for the needs of Zurich Airport (one of busiest airports in Europe). The proposed measure is based on the air traffic assignment model and takes into account Zurich Airport's basic goals: airport capacity increase with reductions in noise level in the airport surroundings. Although, these abovementioned goals are in apparent conflict, it is shown that the proposed model allows for decreases in noise level of, on average, 1dB(A) with a traffic volume increase of 20%. The model is based on categorization of aircraft according to engine type (jet and turbo prop) and wake turbulence category (heavy, middle, small and light) and the assignment of specific runways for take-off and landing for each of the mentioned categories.**

***Index Terms*—Airport, Noise Abatement Measures, Air Traffic Management, Air Traffic Control**

I. Introduction

ONE of the biggest problems of modern airports is the noise generated by air traffic, and the impact of this noise on those living near the airport. Noise is an unavoidable consequence of air traffic but it can be decreased by numerous measures: technical innovations in aircraft design, legislation, etc. These measures, however, often amount to discrimination against some airlines. In the judgment of numerous authors, the noise will probably remain the top concern in airport operation, planning and design in the future, primarily because of greater public sensitivity to noise and other environmental problems [1]. This justifies the level of attention given to finding solutions to this problem by numerous air transport system stakeholders (aircraft manufacturers, airports, airlines, air navigation service providers, etc). This paper presents a measure developed for the needs of Zurich Airport (one of busiest airports in Europe) which does not discriminate against certain airlines and which allows for noise level reduction in the airport surroundings.

Fedja T. Netjasov is with the Division of Airports and Air Traffic safety, Air Transport Department, Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade, Serbia and Montenegro, (e-mail: f.netjasov@sf.bg.ac.yu).

The proposed measure is based on the air traffic assignment model and takes into account the Zurich Airport basic goals: airport capacity increase with reductions in noise level in the airport surroundings. Although, these goals conflict, it can be shown that the proposed air traffic assignment model allows for decreases in noise level of, on average, 1dB(A) with a traffic volume increase of 20%. The proposed air traffic assignment model is based on categorization of aircraft according to engine type (jet and turbo prop) and wake turbulence category (heavy, middle, small and light) and the assignment of specific runways for take-off and landing for each of the mentioned categories.

II. Existing Ways of Addressing the Noise Exposure Problem

The noise exposure problem, generated by air traffic, can be addressed at different levels.

At the first level the problem is addressed during the aircraft design and production process – decreasing the noise at source (quieter engines, aerodynamic construction generating low drag, etc). Since 1960, in this way, the noise level has been decreased by around 10 dB(A) (on average 3 dB(A) per decade) [1] (Figure 1).

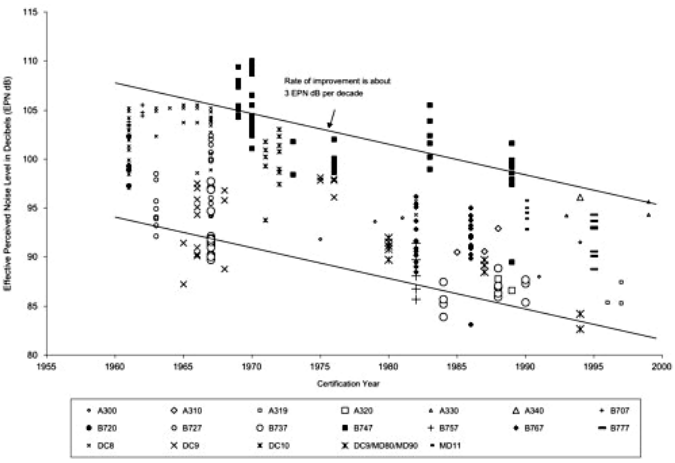


Figure 1. Trends in aircraft noise levels reduction during take-off for aircraft at the maximum take-off weight as a function of certification date [1]

The basic document for dealing with the noise problem at this level is ICAO Annex 16, Volume 1 – Aircraft Noise [2]. This document contains recommendations and guidelines for certification of aircraft intended for use in international air traffic, from a noise point of view. Modernization of the Aircraft Certification Scheme is the topic of the ICAO Committee on Aviation Environmental Protection (CAEP) Working Group WG1 – “Noise” [3].

The second level is related to changes in arrival and departure procedures, i.e. in flying techniques during the mentioned operations. Numerous departure procedures are known: Cutback, IATA, Climb-Cleanup-Cutback…, as well as arrival procedures: Low Drag - Low Power, Continuous Descent Approach,… [4]) and fundamentally all are based on two requirements [5]: to keep the aircraft further from the zones which they contaminate (higher climb and descent rates) and to generate lower noise at source (flying under lower engine power). The tendency in recent decades has been increased work on new arrival and departure procedures as well as on international harmonization and standardization of their development [6]. The European Union, through the Projects SOURDINE I and II is focusing on the development and assessment of new Noise Abatement Procedures (NAP) [7].

The third level is related to legislative measures of restriction introduced by airports and aviation authorities (restrictions related to airport usage and/or prohibition of airport usage at night for some or all aircraft types). Such measures penalize airlines whose aircraft exceed the permitted noise levels during arrivals or departures, and in some cases completely prohibit their operation. Figure 2 presents the number of airports during the past 30 years, which have introduced some kind of restriction. The number of restrictions as well as their type has become even higher, especially after 1996, because of the greater awareness about the air traffic noise among local communities [1] (Figure 2, values in brackets are for year 2001.).

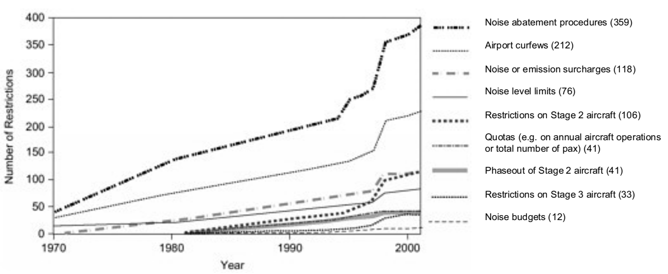


Figure 2. Number of airports worlwide imposing various constraints and charges as a function of time [1]

Recently the ICAO presented an Assembly Resolution A33-7 (September 2001), the Balanced Approach [8] to aircraft noise management around airports. The Balanced Approach is defined as a program for addressing aircraft noise at the individual airport level and considers four elements:

1) Reduction of noise at source;
2) Land-use planning and management;
3) Noise abatement operational procedures;
4) Operating restrictions on aircraft.

This Approach recommends that noise policy should not target single solutions but use any combination of solutions as the most appropriate option to solve the causes of problems [9]. The goal of implementing the above mentioned approach is to achieve the maximum environmental benefit in the most cost effective way. [8].

III. Factors Influencing Air Traffic Noise

The extent to which the airport surroundings will be exposed to air traffic noise depends on numerous factors, starting with the location of the airport, i.e. local topography, then with the traffic characteristics (number of take-offs and landings, as well as their distribution during the day), the structure of the fleet which uses the airport (aircraft types) and departure and arrival trajectories [5]. Other factors, contributing to the noise exposure, are: atmospheric characteristics (e.g. air temperature), noise from aircraft and handling vehicles on the airport maneuvering areas as well as other activities at the airport – aircraft maintenance, engine testing, etc [10].

IV. Air Traffic Noise Measurement

Noise generated by air traffic can be considered using numerous measures (more then 20 [11]) starting with those, which estimate the noise generated by one event (arrival or departure) to the cumulative measures taking into account all operations during the day. All estimation methods are characterized by weighting of the measured noise level, which can be by: frequency, duration or level [11]. Today, the most frequently used measure for a single event is SL (Sound Level) and for cumulative measures - Leq (Equivalent Sound Level), CNEL (Community Noise Equivalent Level), DNL (Day – Night Average Sound Level), NEF (Noise Exposure Forecast), etc [10]. The unit used for measuring the noise generated by air traffic is most commonly the dB(A).

Noise measurement is today an activity performed by specialized services (divisions) at the airports, which collect (at measuring points, with sound level meters), analyze and archive the data about the noise levels around the airport. This data is often used as evidence for penalizing airlines whose aircraft exceed the permitted noise level. Data collected in this way are used for the validation of new arrival and departure procedures, i.e. validation of noise abatement effects, for appropriate land usage (zoning) around the airport, for the acoustic insulation of buildings in airport vicinity, etc [6].

V. AIR TRAFFIC ASSIGNMENT AS A MEASURE OF NOISE EXPOSURE REDUCTION

It is the aim of modern airports to reduce noise in the airport's surroundings by applying various measures. From another perspective, it is in the airport's interest to serve increasing traffic and not to lose potential clients by applying some of the discriminatory measures for noise reduction. Taking all the mentioned facts into account, the idea has emerged that, for airports, which have already implemented some noise abatement measures, further, additional decreases of noise level could be achieved through various air traffic management measures. This paper presents an air traffic assignment model as well as the results of its application in the case of Zurich Airport [12].

The model, proposed in this paper was developed with the goal of decreasing noise levels while at the same time improving the usability of available airport resources (capacity is presented as an hourly number of operations). The presented model is designed for planning purposes.

The model consists of several steps and, for a given airport, contains the following assumptions:

1. Traffic volume for a given day (number of departures and arrivals and their daily distribution) is known;
2. A noise monitoring system (measuring points) is implemented at the airport;
3. Average noise over measuring points, generated by specific aircraft type is known;
4. Sets of departure and arrival routes are known, as well as their characteristics,
5. The ratio of specific aircraft types share in the total daily traffic is a constant (day, evening, night);
6. Sequencing of the arrival traffic as well as the location of parking position are not taken into account, and
7. Transition from en-route sectors to TMA and vice versa, is not taken into account.

Structure of the model:

STEP 1: Analysis of daily traffic characteristics (real or forecasted) with the aim of determining the aircraft fleet structure (aircraft mix) using the airport (aircraft types as well as number of aircraft of specific type (N) during the day); as well as classification of aircraft depending on engine type (turboprop or jet engine) and wake turbulence category (Heavy, Medium, Large and Small) [13];

STEP 2: Analysis of average (measured) noise values generated by different aircraft types, over M specific measuring points separately for departures and arrivals (average noise value is used because the noise value for the particular aircraft type over the particular measuring point is a random variable);

STEP 3: Distribution of aircraft of different classes on runways in use (heuristic model) based on the following criteria (by importance):

1. Average noise level for each aircraft type (departures and arrivals);
2. Available runway length;
3. Over-flying of densely populated areas.

This distribution is based on the air traffic controller's expertise and current practice in use. For example, instead of landing on a runway where, over a particular measuring point, a given aircraft generates a higher noise level, it is instructed to land on another runway (if its length is sufficient for landing of that aircraft type) where it will generate lower noise level over a given measuring point and it will over-fly a less populated area. For every aircraft type it was decided, based on a given criteria, on which runway they should land. Aircraft distribution on runways also depends on the current and forecasted meteorological situation (the runway condition as well as the runway configuration in use).

STEP 4: Calculation of the noise level using the proposed model, for each measuring point in the airport's surroundings separately (based on all departures and arrivals):

$$NL_i^{arr} = \frac{1}{m_i^{arr}} \sum_j (a_{ij}^{arr} \cdot n_{ij}^{arr}) \quad (1)$$

$$NL_i^{dep} = \frac{1}{m_i^{dep}} \sum_j (a_{ij}^{dep} \cdot n_{ij}^{dep}) \quad (2)$$

$$NL_i = \mathbf{max}\left\{NL_i^{arr}, NL_i^{dep}\right\} \quad (3)$$

$$m_i^{arr} = \sum_j a_{ij}^{arr}, \; m_i^{dep} = \sum_j a_{ij}^{dep} \quad (4)$$

where:
i – measuring point, i=1 to M;
j – aircraft type, j=1 to N;
NL_i^{arr} – noise level at measuring point i during arrival [dB(A)];
NL_i^{dep} – noise level at measuring point i during departure [dB(A)];
NL_i – maximal noise level at measuring point i [dB(A)] ;
n_{ij}^{arr} – average noise level which aircraft of type j generate at measuring point i during arrival [dB(A)];
n_{ij}^{dep} – average noise level which aircraft of type j generate at measuring point i during departure [dB(A)];
m_i^{arr} – total daily number of aircraft which over-fly measuring point i during arrival;
m_i^{dep} – total daily number of aircraft which over-fly measuring point i during departure;
a_{ij}^{arr} – daily number of aircraft of type j which over-fly measuring point i during arrival;
a_{ij}^{dep} – daily number of aircraft of type j which over-fly measuring point i during departure.

Equation (1) and (2) calculate noise level over a particular measuring point during arrivals and departures respectively. Equation (3) compares noise values calculated by equations (1) and (2) and takes the higher value as a critical one from the environmental perspective. Equations (4) present the additional condition that the total daily number of aircraft over-flying a particular measuring point during arrival/departure presents the sum of the daily number of aircraft of all types over-flying the same measuring point during arrival/departure respectively.

STEP 5: Comparison of the obtained noise level values NL_i (STEP 4) for each of the M measuring points, between various runway configurations as well as various time horizons.

VI. CASE STUDY: ZURICH AIRPORT, SWITZERLAND

The proposed air traffic assignment model is illustrated with reference to Zurich Airport. This airport is one of the major European hubs and is a pioneer in controlling the impact of aviation on the environment (aircraft noise and engine emission). It is characterized by rather complex infrastructure (three runways), by major spatial operational constraints (9 km from the city center, 15 km from the German border) and by high traffic volume (approximately 1000 operations per day).

In the year 2000, Zurich Airport defined a strategic goal related to the increase of traffic volume by 20% by year 2005 (from 297.000 operations in year 2000, to 358.000 operations in year 2005, Figure 3 [12]).

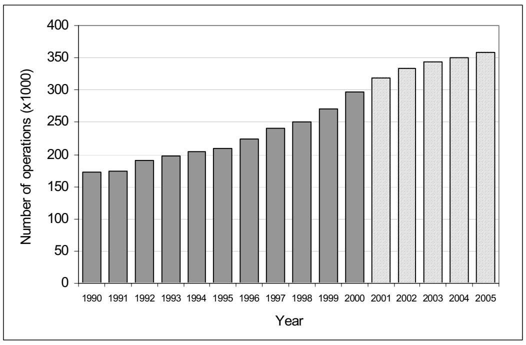


Figure 3. Forecasted traffic volume until 2005, at Zurich airport

Zurich Airport Authority determines the acceptable runway configurations, which will allow capacity to increase and thus service the planned traffic volume (Figure 4). The second goal, equally important, is related to the reduction of noise levels in the airport's surroundings.

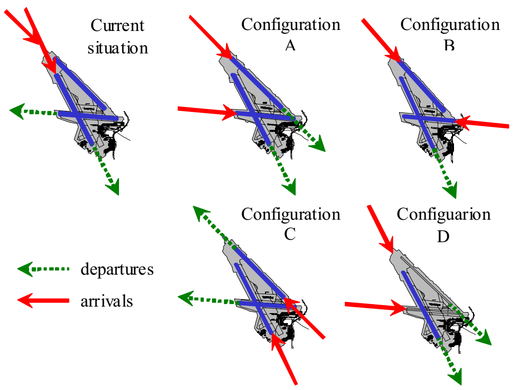


Figure 4. Current situation (year 2000) and available runway configurations

The "Average November 2000 day" was predicted based on Traffic data for November 2000 .(Figure 5). It was identified by traffic analysis that 34 aircraft types used the airport during that day (general aviation aircraft ware not taken into account) [12].

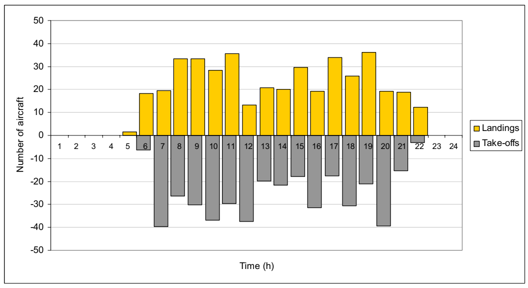


Figure 5. "Average November 2000 day" traffic volume

Based on this data, as well as data about the planned number of operations for 2005 (see Figure 3), the "Average November 2005 day" was predicted. This forecast took into account the planned technological development of aircraft as well as the plans of certain airlines, which are committed to introduce new aircraft.

There is a Division which monitors noise at Zurich Airport, and which collects data on measured noise levels for each aircraft type (separately for departures and arrivals) using a monitoring system containing 9 measuring points (sound level meters) in the airport vicinity (Figure 6, [14]). After statistical analysis, the Division makes this data available to the public.

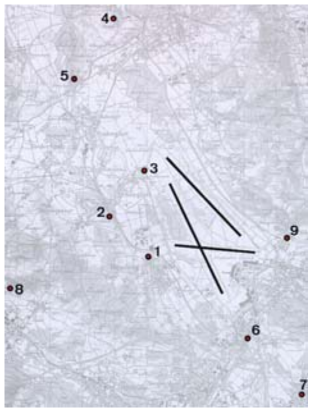


Figure 6. Locations of measuring points [14]

Analyzing the traffic from an "Average November 2000 day" (816 operations), the peak traffic period was identified as being between 08:00 and 11:00 hours with 194 operations in total (Figure 7). In this peak period 16% of operations were made by Medium Turbo Prop (MTP) aircraft, 71% Medium Jet (MJ) and 13% Heavy Jet (HJ) [12].

The noise level analysis for all 34 aircraft types shows that noise during aircraft take-off and departure is higher than noise during arrival and landing, for each aircraft type (Figure 8, e.g. measuring point No. 3, comparison of Medium Turbo Prop and Heavy Jet aircraft [12, 14]), as well as that Medium Turbo Prop aircraft generate lower noise than Medium Jet and Heavy

Jet aircraft (Figure 9, comparison by measuring points for both departures and arrivals, data for measuring point No. 8 were not available).

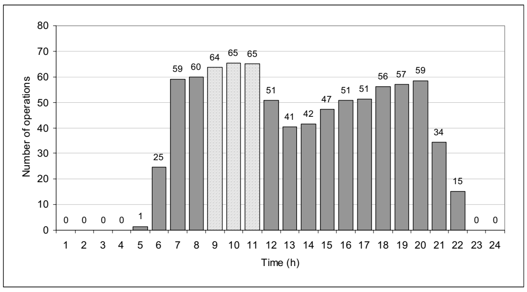


Figure 7. “Average November 2000 day” peak period (total operation)

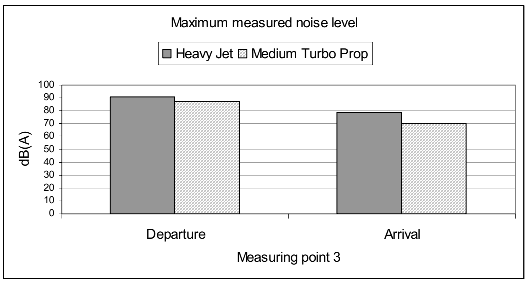


Figure 8. Comparison of noise levels during departure and arrival (measuring point 3)

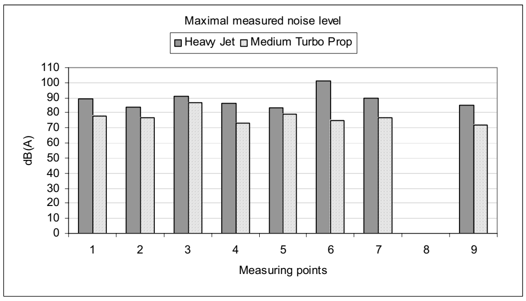


Figure 9. Comparison of noise levels by measuring points

The proposed model is firstly applied to the runway configuration currently in use (in year 2000) and traffic sample containing 816 operations, i.e. “Average November 2000 day”, and after that, on each of four runway configurations accepted by the airport authority (see Figure 4). The average noise levels, calculated by the model, are shown in Table 1. The results unequivocally show that the application of new runway configurations, in certain cases, influence noise reduction (relative to the current situation). Noise level values vary depending on the runway configuration in use as well as the measuring point location [12]. For some measuring points, either in current and other runway configurations, it was not possible to calculate noise levels or it was not necessary because of the measuring point location. E.g. in the current situation, measuring point 4 was written “<M.p.3” (see Table 1) which means that the value for measuring point 4 is less then value for measuring point 3 due to the fact that the point 4 is much further from the runway then point 3, see Figure 6. (The two decimal places for noise level values shown in the tables are used only for computational reasons).

TABLE 1. AVERAGE NOISE LEVELS IN dB(A) BY RUNWAY CONFIGURATIONS (YEAR 2000)

Measu-ring point (M.p.)	Runway configuration				
	Current situation	A	B	C	D
1	79.81	67.16	67.16	74.72	-
2	75.71	< M.p. 1	< M.p. 1	73.08	-
3	71.18	72.89	72.89	81.19	72.91
4	< M.p. 3	< M.p. 3	< M.p. 3	74.94	< M.p. 3
5	< M.p. 3	< M.p. 3	< M.p. 3	71.32	< M.p. 3
6	92.21	74.86	74.86	67.17	80.36
7	80.26	73.19	73.19	< M.p. 6	73.69
8	-	-	-	-	-
9	-	84.43	84.43	72.80	67.16

For the purpose of checking the effects of the proposed assignment model, noise level is also calculated for forecasted traffic (as it was mentioned before), i.e. for an “Average November 2005 day’s” traffic (984 operations, Figure 10, Table 2.) [12].

“Average November 2005 day” traffic sample, uses the same percentage ratio in the total traffic amount, between the three above mentioned aircraft classes, as in the year 2000 sample. Although there is a traffic increase of 20%, it was shown that the proposed model influenced noise level reduction. So, comparing the noise level values for two years (year 2000 and 2005), for the same runway configuration in use and for the same measuring point, the obtained values are, on the whole, lower (Table 3) [12].

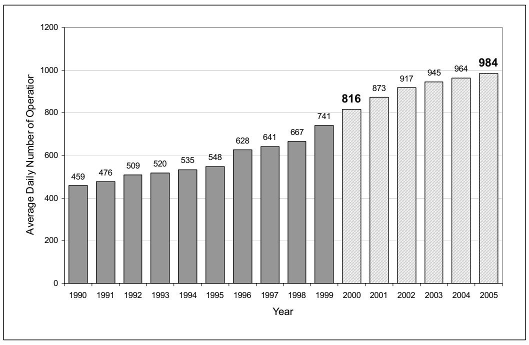


Figure 10. Predicted “Average November day” traffic volume until 2005, at Zurich airport

TABLE 2. AVERAGE NOISE LEVELS IN dB(A) BY RUNWAY CONFIGURATIONS (YEAR 2005)

Measuring point (M.p.)	Runway configuration			
	A	B	C	D
1	66.57	66.57	74.17	-
2	< M.p. 1	< M.p. 1	71.79	-
3	72.89	72.89	81.19	72.91
4	< M.p. 3	< M.p. 3	74.94	< M.p. 3
5	< M.p. 3	< M.p. 3	71.32	< M.p. 3
6	73.94	73.94	66.56	80.02
7	72.18	72.18	< M.p. 6	73.32
8	-	-	-	-
9	84.43	84.43	72.80	66.57

TABLE 3. NOISE LEVELS REDUCTIONS IN dB(A) BY RUNWAY CONFIGURATIONS (YEAR 2005. VS YEAR 2000.)

Measuring point (M.p.)	Runway configuration			
	A	B	C	D
1	0.59	0.59	0.55	-
2	0*	0*	1.29	-
3	0	0	0	0
4	0*	0*	0	0*
5	0*	0*	0	0*
6	0.92	0.92	0.61	0.34
7	1.01	1.01	0*	0.37
8	-	-	-	-
9	0	0	0	0.59

NOTE: reduction for measuring points for which noise level values are not calculated are depicted as 0*.

VII. DISCUSSION AND FURTHER RESEARCH

Comparing the average values for two years can show that at certain measuring points the resulting noise reduction is (roughly) between 0.5 and 1 dB(A), for a given aircraft fleet mix (16% MTP, 71% MJ, 13% HJ, Table 3). The percentage ratio of aircraft in the fleet mix fluctuate during the day, mostly increasing the share of quieter aircraft (MTP or MJ, Figure 11), so it can be expected that values of reduction could be increased to 2 dB(A) during the day for some measuring points.

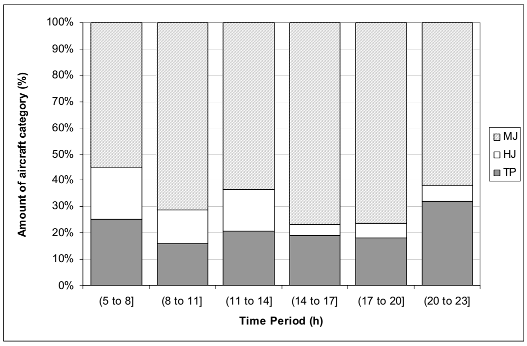


Figure 11. Fluctuation of Percentage ratio of aircraft in the fleet mix during the Zurich Airport working hours (from 5 to 23 o'clock)

According to obtained results, it could be concluded that promising potential for noise level decrease exist. But one question appears: how much the noise level decreases due to change of the percentage ratio of aircraft in the fleet mix and how much because of different runway configuration in use? Also, it would be very interesting to compare the obtained noise levels with the long-term noise levels recommended by the European Commission (2003/613/EC) [15]. It is expected that noise level values determined in such a way, will produce similar reductions to those presented in Table 3. Those are further research steps, together with comparison with results of some noise simulation models (such as FAA's Integrated Noise Model - INM).

VIII. CONCLUSION

The problem of air traffic noise exposure is solved with the application of various practical measures, often discriminating against certain airlines. This paper presents a model for noise level reduction, based on air traffic assignment, which is not discriminatory. The model is based on aircraft classification on two criteria: engine type and wake turbulence category. Distribution of aircraft classes (defined in such a way) on runways, is made using a heuristic algorithm with the aim of decreasing noise exposure. For illustration purposes, the model is tested on Zurich Airport, for different runway configurations as well as time horizons. It was shown that, in the case of a traffic volume increase of 20%, a significant noise level reduction (up to 1 dB(A)) would be possible to achieve.

REFERENCES

[1] National Research Council, *For Greener Skies - Reducing Environmental Impacts of Aviation*, Washington, D.C. USA, 2002.
[2] International Civil Aviation Organization, *Annex 16, Vol. 1, Aircraft Noise*, Montreal, Canada, 1993.
[3] Airport Council International, *Airports and Environmental Legislation, CAEP WG1 – Noise*, pp. 33-35, Brussels, Belgium, 2004.
[4] Uwe Rottman, *Reduction of Aircraft Noise in Civil Air Transport by Optimization of Flight Tracks and Take-Off and Approach Procedures*, Technical University of Berlin, Germany, 1988.
[5] Vojin Tosic, *Airports*, Federal Air Traffic Control Authority, Belgrade, 1978. (in Serbian)
[6] Robert de Muynck, "The SOURDINE Project: Developing Environmentally Friendly Arrival and Departure Procedures", *Air Traffic*, pp. 282-284, 2001.
[7] R. den Boer, C. Beers, H. Huisman, M. Roerdink, J. Scholte, "SOURDINE II: Some Aspects of the Assessment of Noise Abatement Procedures", *6th FAA/EUROCONTROL ATM R&D Seminar*, Baltimore, USA, 2005.
[8] Airport Council International, *Airports and Environmental Legislation, CAEP – The Balanced Approach*, pp. 32-33, Brussels, Belgium, 2004.
[9] A. Celikel, J. Hustache, I. de Lepinay, K. Martin, A. Melrose, "Environmental Tradeoffs Assessment Around Airports", *6th FAA/EUROCONTROL ATM R&D Seminar*, Baltimore, USA, 2005.
[10] Dusan Kalic, "Allowed L_{dn} and L_{eq} noise levels for day and night", *Proceedings of XXXV ETAN*, Belgrade, Vol. VI-VII, pp. 61-66, 1991. (in Serbian)
[11] *California Airport Land Use Planning Handbook - Chapter 6: Measuring Airport Noise*, 2002.
[12] Fedja Netjasov, *Implementation of New Zurich Airport Capacity Concepts, Volume 3: Noise Balancing Around Zurich Airport*, Zurich, Switzerland, 2000.
[13] International Civil Aviation Organization, *Document 8643/28 – Aircraft Type Designator*, Montreal, Canada, 2000.
[14] *Lärm Bulletin*, Zurich Airport, Zurich, 2000.
[15] European Commission, "Commission recommendation concerning the guidelines on the revised interim computation methods for industrial noise, aircraft noise, road traffic noise and railway noise, and related emission data (2003/613/EC)", *Official Journal of the European Union*, Brussels, 2003.

Fedja T. Netjasov is a Research and Teaching Assistant at the Division of Airports and Air Traffic Safety, Faculty of Transport and Traffic Engineering, University of Belgrade. He obtained his B.S. (Dipl. ing.) from the Faculty in 1999. He received his M.S. from the same Faculty in July 2003. Major fields of interest are: ATM, Airports, Modeling, Transport and Traffic System Analysis, Safety.

A framework for RNAV trajectory generation minimizing noise nuisances

Xavier Prats, Fatiha Nejjari, Vicenç Puig, Joseba Quevedo
Universitat Politècnica de Catalunya (UPC)
Av. Canal Olímpic s/n
08086 Castelldefels. Spain
Email: xavier.prats-menendez@upc.edu

Félix Mora-Camino
École Nationale de l'Aviation Civile (ENAC)
7, avenue Edouard-Belin
31055 Toulouse, France
Email: felix.mora@enac.fr

Abstract— **In this work it is presented a framework for a global optimization tool that will take into account aircraft dynamics and performances, noise nuisances and RNAV radionavigation requirements in order to assess an optimum flight depart or approach procedure. This strategy would be used as an optimization process performed by the corresponding authority in charge of the air traffic management of the involved airport or by an on-board optimization algorithm integrated in the Flight Management and Guidance System (FMGS). In both cases, the optimization framework is the same and the differences reside in the specific implementation of the optimization algorithms and the availability of the data in real time. In addition, aircraft's dynamic equations are developed in order to compute the flight trajectory from a set of flight guidance control variables and a first glance into a noise optimization criterion is given. Finally, the global optimization problem is properly formulated and the proposed solving utilities are presented.**

I. INTRODUCTION

Despite the substantial reduction of emited noise of recent aeronautic engines, the sustained growth of air traffic still makes noise reduction in the vicinity of the airports one of the main issues that airport authorities, air traffic service providers and aircraft operators may deal with. In this context, international and national regulations regarding noise exposure have been established by civil aviation authorities in order to cope with this problem but incurring, on the other hand, in higher operations costs for airlines.

Present noise abatement procedures around airports are based on avoiding overfly densely populated areas. This can be accomplished by either going around the concerned area (i.e. modifying the horizontal flight path) or assuring enough vertical distance between the aircraft and the populated area (i.e modifying the vertical flight path). Concerning horizontal flight path management, new Area Navigation (RNAV) concepts are about to offer great advantages in defining more flexible procedures avoiding noise sensitive areas and reducing as well flight path dispersion (due to the better accuracy of RNAV systems regarding conventional ones). In Europe, the RNAV introduction is settled as an objective for all phases of flight, and in this context, EUROCONTROL (the European organization for the safety of air navigation [1]) has defined RNAV concept and satellite navigation systems as the key enablers for future improvements in terms of safety, efficiency and/or economy of flight, provided that their implementation is based on a fully co-ordinated, harmonized, evolutionary and flexible planning process [2], [3].

It is also possible to define noise abatement procedures acting in the vertical domain of a given flight trajectory. In this case, the methodology consist in the definition of optimal climb profiles, acting on the climb speeds and thrust configurations in such a way to increase the vertical distance of the aircraft and a sensitive area located in a certain region under the flight path. For instance, the most widely used noise abatement procedures for take-off are the so called *ICAO NADP departure procedures* defined in [4]. The NADP-1 procedure is designed to protect areas located *close* to the airport, while the NADP-2 procedure is designed to protect *distant* areas to the airport. Each procedure specifies the airspeed profile that should be maintained during the initial climb as well as the points (altitudes) where thrust/power reduction may be done. The difference between NADP-1 and NADP-2 procedures resides in the fact that NADP-1 gives more importance to climb as fast as possible and then accelerate and gain airspeed while NADP-2 tries to accelerate first and then climb. However, the main problem of these kind of procedures are that they are generic procedures and not always fit into the specific problems or environment that a certain airport may suffer.

In SOURDINE I project [5] an effort to improve take off procedures was done and some simulations with specific aircraft types were carried out. In this study optimal take-off procedures were obtained involving a progressive increase in thrust (which is not feasible with present technology) and low airspeeds during the whole departure which could be a problem regarding airport capacity. Those procedures were derived from former NADP-1 and NADP-2 procedures and optimization involved changes in values for engine cut-off, acceleration and climb points (altitudes). Further work performed in SOURDINE II project [6] dealt with refined take-off procedures selecting a grid of speed/thrust combinations and altitudes where thrust cut-off was performed. In this case, more than 30 different simulations were carried out, but not global optimization was done.

Concerning the approach procedure, there exist very simple procedures such as to intercept the Instrument Landing System (ILS) glide-slope at higher altitude, higher ILS glide-slope angles or Low Drag-Low Power Approach (LPLD) procedures. More complex and efficient approach procedures deal with Continuous Descent Approaches (CDA) which, in turn, are being tested in some airports. During a CDA the aircraft performs a thrust-idle flight until a point before ILS-Localizer interception, reducing considerably the emitted noise. CDA procedures reduce significantly noise levels during the approach, but have an important impact on air traffic control operations and airport capacity being useful only in certain circumstances while keeping airport capacity with acceptable levels [7]. Clarke et al. proposed an improvement of conventional CDA approaches by defining the Three Degree Decelerating Approach (TDDA) showing improvements on better noise abatement, while maintaining acceptable capacity levels [8].

As it is seen, present noise abatement procedures are far from being the optimal ones minimizing noise nuisances. This is due to several factors like the impossibility to define a general criterion fitting all airports necessities, the limitations of nowadays technology on-board, and the constraints imposed by airport capacity or air traffic control issues. Nevertheless, some research in theoretical optimum trajectories minimizing the noise impact in depart/approach procedures is also found in the literature. For instance, Visser et al. show in [9] and [10] a technique to obtain optimal noise depart and approach procedures respectively. This tool combines the noise computations of the FAA's Integrated Noise Model (INM) [11] a Geographical Information System (GIS) and a dynamic trajectory optimization algorithm. Similar methodology is proposed by Clarke et al. in [12], and an adaptative algorithm for noise abatement can be found in [13]. An other study (see [14]) empathizes that most current noise abatement procedures (like those explained above) are local adaptations of generic procedures aimed at optimizing aircraft noise footprint and do not generally take into account the actual population density and distribution at a specific airport site. Thence, a noise performance trade-off is presented between arrival trajectories that are optimized according to different types of noise abatement criteria. Typically, these different criteria are not compatible and the variables that optimize one objective may be far from optimal for the others, pointing out the difficulty to properly identify the absolute minimal trajectory among all the local minimal ones.

This paper firstly presents the framework for a global optimization tool, developing the different components that are involved in the process. Then, a flight guidance model is presented in order to describe the dynamic constraints that will apply to the flight trajectory. Fourth section of this paper is devoted to define the optimization criterion, which will take into account not only the noise nuisances but also some airlines considerations such as time and fuel consumptions and finally the global optimization problem is formally formulated.

II. Framework of the optimization strategy

The framework proposed to optimize depart or approach trajectories is presented in this section and it is summarized in figure 1. The involved airport, with its surrounding cartography, geography and meteorological data, will define a *scenario* which will be used to compute a given *noise nuisance* in function of the emited aircraft noise along its trajectory. This value, jointly with some fuel/time economy considerations, will define a global *optimization criterion*. Then, an *optimization algorithm* will compute the best departing or approaching trajectory which minimizes the *optimization criterion* and satisfies a set of *trajectory constraints* which, in turn, will depend on the dynamics of the aircraft, RNAV design constraints and the airspace configuration.

This strategy would be used as an optimization process performed by the corresponding authority in charge of the air traffic management of the involved airport or by an on-board optimization algorithm integrated in the Flight Management and Guidance System (FMGS). In both cases, the optimization framework is the same and the differences reside in the specific implementation of the optimization algorithms, which in the case of being integrated on-board will require some real time performance as well as the possibility of the on-board system to obtain the scenario, airspace and meteorological data. In a further study it is proposed to assess the specific requirements of both implementations.

The following sections give more details of each component involved in the trajectory optimization.

A. Input data

Different kind of input data are required for the proper model of each component. As shown i figure 1 the input data sources are:

- **Airspace**: containing the airspace characteristics of the studied area, restricted areas, airspace structure organization as well as the the departing or arriving points that will define the procedure final or initial conditions respectively.
- **RNAV Navigation**: containing the RNAV design procedure criteria that specify the constraints that should satisfy any procedure in order to meet the required level of safety.
- **Cartography**: containing cartographic data, including terrain elevations and obstacle identification, that will be needed to safely define the flight procedure according with the RNAV procedure design criteria. In addition, this information will be used by the noise model.
- **Aircraft performances**: including the aerodynamic and power plant related data of the studied aircraft that will be needed to build up the aircraft dynamic model as well as the noise model.
- **Meteorology**: containing meteorological data that will affect available runway configuration, the noise propagation model and the aircraft ground trajectory.
- **Airport**: containing the location of the airport, the type of procedure, available runway configuration etc.
- **Geography**: containing geographic data of airport's surrounding areas such as the location and characteristics of the inhabited areas.

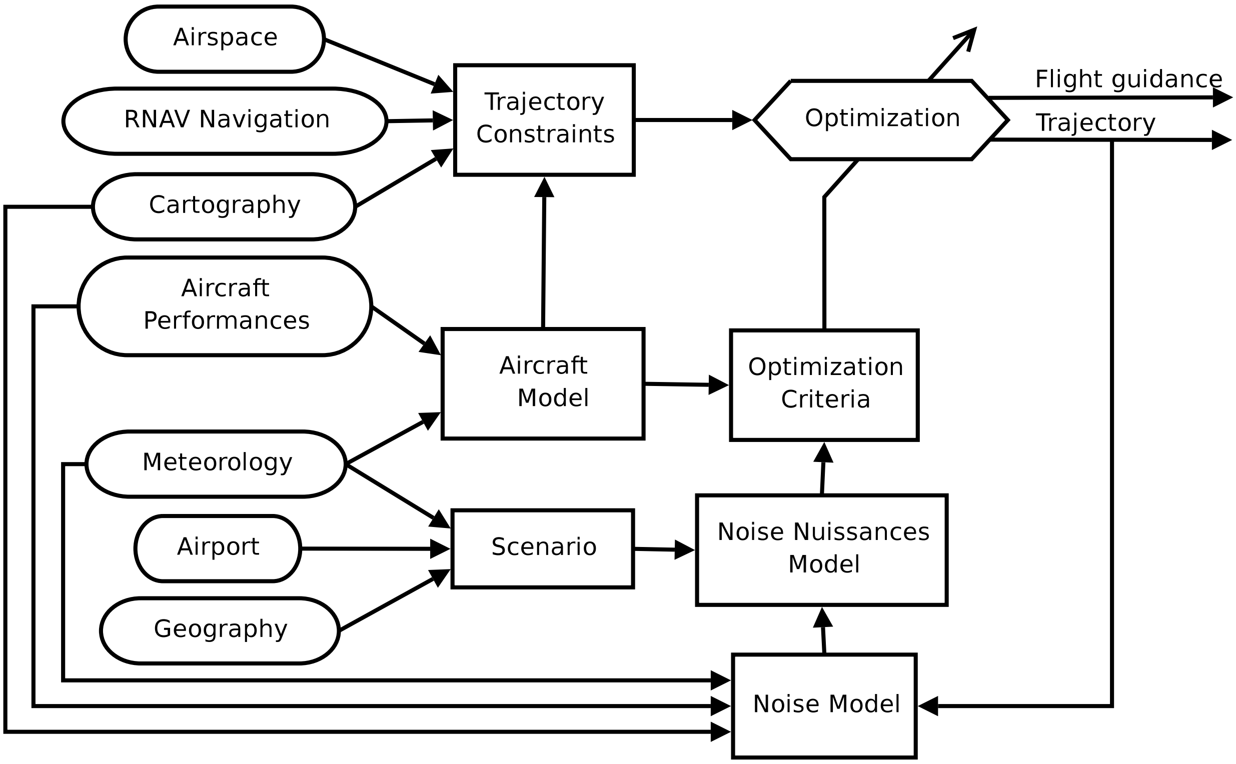


Fig. 1. Noise nuisances optimization framework

B. Trajectory constraints

This component will define a set of rules that will restrict the amount of possible trajectories into a valid domain where the optimization will take place.

First of all, airspace organization will be taken into account, regarding prohibited, dangerous and restricted areas as well as and particular airspace sectorization focusing on the compatibility with other existing flight procedures in the same airspace. This analysis will finally identify a set of *usable airspace* portions where the obtained trajectories should be contained.

In addition, this component will take into account procedure safety issues. As it is well known, ICAO Document 8168: *Procedures for Air Navigation Services - Aircraft Operations (PANS-OPS) - Volume II, Construction of Visual and Instrument Flight Procedures* [15], contains all the rules and methodology for designing flight procedures. All these information will be transformed into the form of trajectory constraints, restricting even more the *usable airspace* defined above (for instance, in order to take into account obstacle clearance etc.) and bounding the trajectory design variables (for instance, in order to consider maximum climb gradients etc.).

Finally, trajectory boundary conditions will be also specified. If a departing procedure is studied, the final departure point (or points) location and altitude will be included as the trajectory final boundary condition. On the other hand, if an arriving procedure is studied, the initial arriving point (or points) location and altitude will be included as the trajectory initial boundary condition. It should be noted that it is not necessary to define fixed (known) boundary points since the optimization algorithm will be able to deal with not fixed boundary conditions.

C. Aircraft Model

The dynamic equations of the aircraft trajectory and flight guidance are contained in this component, which uses the required parameters from an aircraft dependent performance data base. These equations will set the relations between the actual aircraft trajectory and the variables required to control the aircraft's trajectory. In addition, a cost function expression regarding the airplane operator objectives, will be modeled in order to be taken into account into the global optimization criterion giving a specific weight to the time and fuel consumed during the whole procedure.

D. Scenario definition

The airport configuration, the present meteorology and the geographic data around the airport will define the scenario where the trajectory should be optimized. Thence, this scenario will be used to determine the noise nuisances caused by the noise generated by the departing/approaching aircraft along its trajectory. Therefore it should provide information about:

- the runway used and the type of procedure (depart or approach)
- the time of the year and hour of the day that the procedure will be flown

- the procedure's intended frequency of use.
- the population density and distribution of the inhabited areas surrounding the airport.
- the type of activities developed in the inhabited areas (industrial zone, residential zone...)
- the location and characteristics of other existing noisy areas surrounding the airport, such as motorways, harbors...
- the location of possible sensitive areas such as environmental protected zones.

E. Noise model

This component will contain a noise model of propagation and will compute the perceived noise level at each point of the trajectory. Therefore, information about the cartography and the meteorology will be used in order to compute the noise propagation and attenuation as well as some aircraft data which will enable the model of the emited aerodynamic and propulsive noise.

F. Noise nuisances model

The aircraft perceived noise in a given point of the trajectory will be important in assessing the optimum trajectory. Nevertheless, what is intended to be minimized is the noise nuisances and not only the perceived noise. For instance, the same noise is more penalizing if it is heard in an inhabited residential zone at midnight than in an industrial zone at noon. Therefore, this component will compute a given value of nuisance in function of the perceived noise and in function of the scenario characteristics. A set of fuzzy inference rules will be defined in order to give a weight to all factors that will play a significant role in the noise nuisance build up.

G. Optimization criterion

As it has been commented before, this component will define a global optimization criterion that will take into account the noise nuisances of the flight trajectory and the aircraft operator considerations. This criterion, will be used as performance factor in the optimization process.

H. Optimization algorithms

Optimal control techniques will be applied here in order to find the optimum trajectory (and flight guidance parameters) that minimizes the global criterion of nuisance and aircraft operator economy under a set of constraints.

III. FLIGHT GUIDANCE MODEL

Starting from a dynamic and cinematic analysis of an aircraft's motion, the goal of this section is to obtain a state representation of the flight guidance equations that are needed to describe the trajectory of the aircraft. This state representation deduction will assume some initial hypotheses and will define the dynamic relationship between the state variables and a set of flight control variables. This kind of representation will enable the use of future optimization methods for dynamic systems.

A. Reference frames definition

Three different reference frames are needed to describe the aircraft's equations of motion. A *Ground* reference frame which will be used as inertial frame, an *Air* reference frame where the aerodynamic forces are easily expressed, and finally a *Body* reference frame used as an intermediate frame to convert *Air* magnitudes to *Ground* magnitudes. These three reference frames are defined as:

- **Ground** $G = [O; n, e, d]$ **reference frame**: North-East-Down conventional right handed frame on the surface of the Earth with a given origin O. The d axis points downwards following the local vertical direction (i.e the direction of the local gravity vector, $\boldsymbol{g}$) and the n-e plane is tangent to the Earth's surface at O. The e axis points to the East and therefore the n axis points to the North.
- **Body** $B = [P; x, y, z]$ **reference frame**: Conventional right handed set of body fixed axes with origin P at the center of mass of the airplane. The x axis is forward aligned, y axis starboard aligned and z axis down in the aircraft.
- **Air** $A = [P; x_A, y_A, z_A]$ **reference frame**: Conventional right handed frame with origin P at the center of mass of the airplane. The x_A axis is always aligned with the relative velocity vector between the air and the plane.

Three consecutively rotations are defined to describe the instantaneous attitude of the aircraft (Body reference frame) with respect to the Ground reference frame. Starting from G these rotations are:

- First rotation about the d axis, nose right (*yaw* angle ψ)
- Second rotation about the new e axis, nose up (*pitch* angle θ)
- Third rotation about the new n axis, right wing down (*roll* angle ϕ)

In the same way, two consecutively rotations define the frame A from the frame B:

- First rotation about the y axis, downwards (*angle of attack* angle α)
- Second rotation about the new z axis, rightwards (*sideslip* angle β)

If $\boldsymbol{a}_G$ and $\boldsymbol{a}_B$ are the position vectors of a given point in frames G and B respectively the coordinate transformation which relates both vectors is:

$$\boldsymbol{a}_G = \boldsymbol{\mathcal{R}}_{GB}\boldsymbol{a}_B \tag{1}$$

with:

$$\boldsymbol{\mathcal{R}}_{GB} = \begin{bmatrix} \mathrm{c}\psi\mathrm{c}\theta & -\mathrm{s}\psi\mathrm{c}\phi + \mathrm{c}\psi\mathrm{s}\theta\mathrm{s}\phi & \mathrm{s}\psi\mathrm{s}\phi + \mathrm{c}\psi\mathrm{s}\theta\mathrm{c}\phi \\ \mathrm{s}\psi\mathrm{c}\theta & \mathrm{c}\psi\mathrm{c}\phi + \mathrm{s}\psi\mathrm{s}\theta\mathrm{s}\phi & -\mathrm{c}\psi\mathrm{s}\phi + \mathrm{s}\psi\mathrm{s}\theta\mathrm{c}\phi \\ -\mathrm{s}\theta & \mathrm{c}\theta\mathrm{s}\phi & \mathrm{c}\theta\mathrm{c}\phi \end{bmatrix} \tag{2}$$

Where, for the sake of simplicity, c(.) and s(.) represent $\cos(.)$ and $\sin(.)$ respectively. Similarly, if $\boldsymbol{a}_A$ is the position vector of a given point in frame A, the coordinate transformation yields:

$$\boldsymbol{a}_B = \boldsymbol{\mathcal{R}}_{BA}\boldsymbol{a}_A \tag{3}$$

with:

$$\boldsymbol{\mathcal{R}}_{BA} = \begin{bmatrix} \mathrm{c}\alpha\mathrm{c}\beta & -\mathrm{c}\alpha\mathrm{s}\beta & -\mathrm{s}\alpha \\ \mathrm{s}\beta & \mathrm{c}\beta & 0 \\ \mathrm{s}\alpha\mathrm{c}\beta & -\mathrm{s}\alpha\mathrm{s}\beta & \mathrm{c}\alpha \end{bmatrix} \tag{4}$$

It is also possible to define three new angular rotations which led us from *Ground* reference frame directly to *Air* reference frame. In this case, starting from G, these rotations are:

- First rotation about the d axis, nose right (aerodynamic *yaw* angle χ)
- Second rotation about the new e axis, nose up (aerodynamic *pitch* angle γ)
- Third rotation about the new n axis, right wing down (aerodynamic *roll* angle μ)

In terms of coordinate transformations:

$$\boldsymbol{a}_G = \boldsymbol{\mathcal{R}}_{GA}\boldsymbol{a}_A \tag{5}$$

with:

$$\boldsymbol{\mathcal{R}}_{GA} = \begin{bmatrix} \mathrm{c}\chi\mathrm{c}\gamma & -\mathrm{s}\chi\mathrm{c}\mu + \mathrm{c}\chi\mathrm{s}\gamma\mathrm{s}\mu & \mathrm{s}\chi\mathrm{s}\mu + \mathrm{c}\chi\mathrm{s}\gamma\mathrm{c}\mu \\ \mathrm{s}\chi\mathrm{c}\gamma & \mathrm{c}\chi\mathrm{c}\mu + \mathrm{s}\chi\mathrm{s}\gamma\mathrm{s}\mu & -\mathrm{c}\chi\mathrm{s}\mu + \mathrm{s}\chi\mathrm{s}\gamma\mathrm{c}\mu \\ -\mathrm{s}\gamma & \mathrm{c}\gamma\mathrm{s}\mu & \mathrm{c}\gamma\mathrm{c}\mu \end{bmatrix} \tag{6}$$

B. Dynamic analysis

At this point Newton's Second Law will be applied to a given flying airplane. First of all, the adopted notation and some basic hypothesis that will ease our study will be introduced.

1) Basic hypothesis and notation: Let $\boldsymbol{a}$ be a given vector, then $\boldsymbol{a}_X$ will be vector $\boldsymbol{a}$ expressed in X reference frame coordinates. On the other hand, if $\boldsymbol{b}$ is a velocity vector, it will be noted $\boldsymbol{b}_X^Y$ as the velocity $\boldsymbol{b}$ seen from reference frame Y, expressed in X reference frame coordinates.

Finally, the operator $\frac{d}{dt_X}(.)$ stands for the time derivative as seen from reference frame X.

HYPOTHESIS 1 : The wind components are constant
It is worth to assume that the wind velocity vector is constant over a region much larger than the size of the aircraft, so wind shearing effects and torques will be neglected.

HYPOTHESIS 2 : The total mass of the airplane remains constant with time
This problem will consider the take-off or approach maneuvers of a conventional commercial airplane. For example, a typical airplane of 180 passengers will consume around $1000\,Kg$ during a climb to cruise altitude, being its total mass of about $70000\,Kg$. Therefore the mass change over the considered time period will be about 1.5% which is considered negligeable [16].

HYPOTHESIS 3 : The mass distribution is also constant with time
Passenger movements, fuel sloshing and shifting payloads effects are neglected and therefore the center of gravity of the airplane will be supposed to stay in the same place during the time period of consideration.

HYPOTHESIS 4 : The reference frame G is supposed to be an *inertial* frame
In fact, Ground *reference frame is both accelerating and rotating, however the accelerations associated with the Earth's movement are negligible compared to the accelerations that will produced by a maneuvering aircraft.*

2) Dynamic equations: Taking into account all hypothesis and considerations above, Newton's Second Law can be written in the *Ground* reference frame as:

$$\frac{1}{m}\sum \boldsymbol{F}_G = \frac{d}{dt_G}\boldsymbol{v}_G^G = \frac{d}{dt_G}\left(\boldsymbol{v}_G^A + \boldsymbol{w}_G^G\right) = \frac{d}{dt_G}\boldsymbol{v}_G^A \tag{7}$$

were $\sum \boldsymbol{F}$ is the sum of all external applied forces, m is the total mass of the aircraft, $\boldsymbol{v}$ is the velocity of the aircraft's center of mass and $\boldsymbol{w}$ is the local wind velocity.

As it will be seen later, aerodynamic forces are much simpler if represented in the *Air* reference frame. Therefore, it is interesting to rewrite equation 7 and express all magnitudes in A frame [17]:

$$\frac{1}{m}\sum \boldsymbol{F}_A = \frac{d}{dt_G}\boldsymbol{v}_A^A = \frac{d}{dt_A}\boldsymbol{v}_A^A + \boldsymbol{\omega}_A^{GA} \times \boldsymbol{v}_A^A \tag{8}$$

where $\boldsymbol{\omega}^{GA}$, is the angular velocity vector of frame A relative to frame G.

The sum of all external applied forces will be decomposed as:

$$\sum \boldsymbol{F}_A = \boldsymbol{F}_A^a + \boldsymbol{F}_A^p + m\boldsymbol{\mathcal{R}}_{AG}\boldsymbol{g}_G \tag{9}$$

were $\boldsymbol{F}^a$ represents the sum of all aerodynamic forces, $\boldsymbol{F}^p$ the sum of all propulsive forces and $\boldsymbol{g}$ is the local gravity vector.

Lets expand now all vectors used in last equations. As commented before, in the *Air* reference frame aerodynamic forces can easily written as [16]:

$$\boldsymbol{F}_A^a = \begin{bmatrix} -D \\ Y \\ -L \end{bmatrix} \tag{10}$$

Where D and L are the *Drag* and *Lift* aerodynamic forces and Y the aerodynamic sideforce component along the *Air* y axis.

Concerning the propulsive forces, it is assumed:

HYPOTHESIS 5 : The sum of all propulsive forces is a vector directed as the body x axis
In modern jet aircraft, with the engines under the main wings, there exist typically a small thrust component in the vertical z axis that will be neglected in this study. In addition it is assumed that in our study all aircraft's engines are operative leading to a symmetrical thrust force regarding x-z plane.

Thus:

$$\boldsymbol{F}_A^p = \boldsymbol{\mathcal{R}}_{AB}\boldsymbol{F}_B^a = \begin{bmatrix} T\cos\alpha\cos\beta \\ -T\cos\alpha\sin\beta \\ -T\sin\alpha \end{bmatrix} \tag{11}$$

where T is the total thrust forces developed by all aircraft's engines.

The local gravity vector $\boldsymbol{g}$ in G reference frame is simply:

$$\boldsymbol{g}_G = \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \quad \text{with } g \simeq 9.81\ ms^{-2} \tag{12}$$

and, by definition, $\boldsymbol{v}^A$ in A reference frame is written as:

$$\boldsymbol{v}_A^A = \begin{bmatrix} v \\ 0 \\ 0 \end{bmatrix} \tag{13}$$

where v is the module of the relative air to aircraft velocity, also known as the *True Airspeed* (TAS).

Finally, vector $\boldsymbol{\omega}_A^{GA}$ components are defined as:

$$\boldsymbol{\omega}_A^{GA} = \begin{bmatrix} p_A \\ q_A \\ r_A \end{bmatrix} \tag{14}$$

In our study it will be perfectly reasonable to assume:

HYPOTHESIS 6 : The sideslip angle β is considered to be zero

This hypothesis assumes that the flight is always symmetrical and turns are always coordinated, which is perfectly reasonable in civil transport aircraft when all engines are operative.

Last hypothesis leads to $Y = 0$ [16].

With all considerations above, equations 8 and 9 can be finally expanded as:

$$\begin{aligned} \dot{v} &= \tfrac{1}{m}T\cos\alpha - \tfrac{1}{m}D - g\sin\gamma \\ 0 &= vr_A - g(\cos\gamma\sin\mu) \\ 0 &= -vq_A + \tfrac{1}{m}L + \tfrac{1}{m}T\sin\alpha - g\cos\gamma\cos\mu \end{aligned} \tag{15}$$

C. Cinematic analysis

The determination of the flight path of the airplane relative to the *Ground* reference system will be done by numerical integration of the airplane's *Ground* coordinates, which in turn, are expressed in function of the velocity of the aircraft's center of mass as:

$$\dot{\boldsymbol{p}}_G = \boldsymbol{v}_G^G = \boldsymbol{\mathcal{R}}_{GA}\boldsymbol{v}_A^A + \boldsymbol{w}_G^G \tag{16}$$

if it is assumed that the local wind has north, east and down velocity components as:

$$\boldsymbol{w}_G^G = \begin{bmatrix} w_n \\ w_e \\ w_d \end{bmatrix} \tag{17}$$

the expansion of equation 16 leads to:

$$\begin{bmatrix} \dot{n} \\ \dot{e} \\ \dot{d} \end{bmatrix} = v \begin{bmatrix} \cos\chi\cos\gamma \\ \sin\chi\cos\gamma \\ -\sin\gamma \end{bmatrix} + \begin{bmatrix} w_n \\ w_e \\ w_d \end{bmatrix} \tag{18}$$

On the other hand, in order to perform this integration, the time evolution of the angles $\boldsymbol{\Psi} = [\chi\ \gamma\ \mu]^T$ must be known. It can be easily shown that the angle rates $\dot{\boldsymbol{\Psi}}$ can be expressed in function of the angular velocities $\boldsymbol{\omega}_A^{GA} = [p_A\ q_A\ r_A]^T$ as [17]:

$$\begin{bmatrix} p_A \\ q_A \\ r_A \end{bmatrix} = \begin{bmatrix} \dot{\mu} - \dot{\chi}\sin\gamma \\ \dot{\gamma}\cos\mu + \dot{\chi}\cos\gamma\sin\mu \\ \dot{\chi}\cos\gamma\cos\mu - \dot{\gamma}\sin\mu \end{bmatrix} \tag{19}$$

D. Space state representation

By substituting equations 19 into 15 yields:

$$\begin{aligned} \dot{v} &= \tfrac{1}{m}T\cos\alpha - \tfrac{1}{m}D - g\sin\gamma \\ \dot{\chi} &= \tfrac{1}{mv\cos\gamma}L\sin\mu + \tfrac{1}{mv\cos\gamma}T\sin\mu\sin\alpha \\ \dot{\gamma} &= \tfrac{1}{mv}L\cos\mu + \tfrac{1}{mv}T\cos\mu\sin\alpha - \tfrac{g\cos\gamma}{v} \end{aligned} \tag{20}$$

Aerodynamic forces L and D can be modeled as:

$$\begin{aligned} L &= \tfrac{1}{2}\rho(d)Sv^2C_L \\ D &= \tfrac{1}{2}\rho(d)Sv^2C_D \end{aligned} \tag{21}$$

where $\rho(d)$ is the air density, which can be considered only altitude $(-d)$ dependant and S is the total surface of the airplane's wings. Aerodynamic coefficients C_L and C_D are usually modeled as:

$$\begin{aligned} C_L &= C_{L_0} + C_{L_\alpha}\alpha \\ C_D &= C_{D_0} + \tfrac{1}{\pi Ae}C_L^2 \end{aligned} \tag{22}$$

where C_{L_0}, C_{L_α} and C_{D_0} are known aerodynamic parameters, A is the aspect ratio of the wing $A = \frac{b}{c} = \frac{b^2}{S}$ (being b the total wing span and c the mean chord of the wing's airfoils) and e the *Oswald factor* of the wing.

Concerning the propulsive force T, considering a turbojet engine type, in a first approximation this force can be modeled as:

$$T = n_e\rho(d)S_ev(V_e - v) \tag{23}$$

where n_e is the number of engines, S_e is the effective engine inlet area and V_e is the exhaust gas velocity. It can be shown that V_e can be fitted in a linear form depending on the engine's low pressure rotor speed $N1$ provided that $50\% \lesssim N1 \lesssim 100\%$:

$$V_e = V_{e_0} + \lambda(N1 - N1_0) \tag{24}$$

Therefore, in a more generic way, aerodynamic and thrust forces can be written as:

$$\begin{aligned} L &= L(d, v, \alpha) \\ D &= D(d, v, \alpha) \\ T &= T(d, v, \alpha, N1) \end{aligned} \tag{25}$$

Finally, merging equations 20 and 18 a state representation of the flight guidance equations is obtained:

$$\begin{bmatrix} \dot{v} \\ \dot{\chi} \\ \dot{\gamma} \\ \dot{n} \\ \dot{e} \\ \dot{d} \end{bmatrix} = \begin{bmatrix} -g\sin\gamma \\ 0 \\ -\frac{g}{v}\cos\gamma \\ v\cos\chi\cos\gamma + w_n \\ v\sin\chi\cos\gamma + w_e \\ -v\sin\gamma + w_d \end{bmatrix} + \\ + \begin{bmatrix} \frac{1}{m}\left[-D(d,v,\alpha) + T(d,v,\alpha,N1)\cos\alpha\right] \\ \frac{\sin\mu}{mv\cos\gamma}\left[L(d,v,\alpha) + T(d,v,\alpha,N1)\sin\alpha\right] \\ \frac{\cos\mu}{mv}\left[L(d,v,\alpha) + T(d,v,\alpha,N1)\sin\alpha\right] \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{26}$$

the state $\boldsymbol{x}$ and control $\boldsymbol{u}$ vectors are identified as:

$$\begin{aligned} \boldsymbol{x} &= [v\ \chi\ \gamma\ n\ e\ d]^T \\ \boldsymbol{u} &= [\alpha\ \mu\ N1]^T \end{aligned} \tag{27}$$

and the flight guidance equations can be summarized in the following expression:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{u}) \tag{28}$$

E. The final guidance model

As it has been shown, the guidance parameters adopted in last model are $\boldsymbol{u} = [\alpha\ \mu\ N1]^T$. Engine's parameter $N1$ is often the main control variable of the *auto/thrust* system but, on the other hand, the typical variables for attitude guidance are pitch and roll angles $[\theta, \phi]$ or even, in recent fly-by-wire auto-flight systems, the roll rate and the vertical load factor $[p, n_z]$. However, the use of $[\chi, \gamma, \mu]$ angles instead of $[\psi, \theta, \phi]$ angles, and the possibility of expressing all magnitudes into the *Air* reference frame simplifies considerably the state equations given in 26, playing an important role for easing further optimization algorithms. Thus, the optimization process will yield to the required guidance control variables $\boldsymbol{u}(t) = [\alpha(t)\ \mu(t)\ N1(t)]^T$ needed to achieve the optimum trajectory. From equations 1, 3 and 5 it appears that θ and ϕ angles can be obtained, through a complex trigonometric expression, from α , μ and γ angles. This expression is here summarized by :

$$[\theta, \phi] = \boldsymbol{\Gamma}(\alpha, \mu, \gamma) \tag{29}$$

Therefore, the optimum values of angles $[\theta, \phi]$ can be computed in real time using equation 29 and supplied as reference values for the basic longitudinal and lateral modes of a classical autopilot (see figure 2). Nevertheless, it should be noted that thanks to the wide use of high performance inertial reference systems and integrated auto-pilots architecture, it is worth to suppose that in a future an autopilot may deal directly with $[\alpha, \mu]$ control guidance inputs.

IV. Definition of the optimization criterion

The main cost components, related with the evolution of an aircraft during a time interval $[t_0, t_f]$ and considering either a climb, cruise or descent flight phases, are the fuel cost, the delay cost and the noise effect.

A. Flight costs model

Fuel cost C_f can be computed as:

$$C_f = \int_{t_0}^{t_f} \pi_c FF(t)dt \tag{30}$$

where π_c is the fuel price and FF is the fuel flow which can be estimated as a smooth function of aircraft's speed and engine's thrust:

$$FF = f_f(v, T) \tag{31}$$

For instance, the BADA aircraft performance database from EUROCONTROL [18] provides the formula:

$$FF = c_{f1}(1 + c_{f2}v)T \tag{32}$$

where c_{f1} and c_{f2} are parameters whose values are characteristic of the aircraft type.

The delay cost represents the different constant rate costs associated with aircraft operations (insurances, traffic control fees, crew salaries, etc). Total delay cost is given by:

$$C_d = \pi_d(t_f - t_0) \tag{33}$$

where π_d is the cost attached to one unit of time of delay. Here the value of π_d could be a source for controversy, so this cost will be merged in a global noise index, NI, representative of the traffic management policy with respect to noise.

Aircraft noise is mainly composed of an omnidirectional source related with the aerodynamic noise and a directional source related with the engine noise. The level of noise P received at point $\boldsymbol{p}_G = [p_n, p_e, p_d]$ can be modeled as [19]:

$$\begin{aligned} P(p_n, p_e, p_d) = (P_a(d, a) + P_e(d, V_e)\cdot \\ \cdot\, \varpi_d(n - p_n, e - p_e, d - p_d, \gamma, \psi))\cdot \\ \cdot\, \varpi_m(n - p_n, e - p_e, d - p_d, w_n, w_e, w_d) \end{aligned} \tag{34}$$

where P_a is the power of aerodynamic noise, P_e is the power of engine noise, ϖ_d is related with the directional effect of jet noise and ϖ_m is related with the noise distortion resulting from the wind. The impact of instant noise over an individual location $\boldsymbol{p}_G = [p_n, p_e, p_d]$ is then given by:

$$\frac{\Omega(P(p_n, p_e, p_d))}{((n - p_n)^2 + (e - p_e)^2 + (d - p_d)^2))} \tag{35}$$

where Ω is a logarithmic function.

Then an aggregated measure of the noise nuisances over the ground area surrounding the aircraft can be computed by:

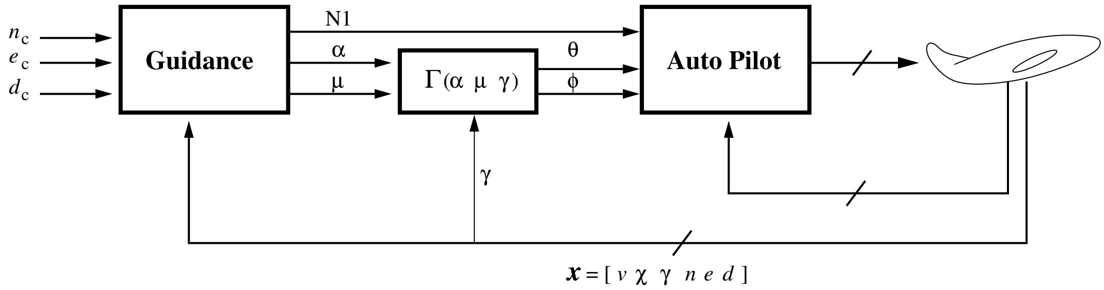


Fig. 2. Flight guidance schema

$$\Phi_n(n,e,d) = \int_{A_n}\int_{A_e}\int_{A_d} \sigma(p_n,p_e,p_d)\cdot \\ \cdot \frac{\Omega(P(p_n,p_e,p_d))}{((n-p_n)^2+(e-p_e)^2+(d-p_d)^2))} dndedd \tag{36}$$

where σ is a density function which takes into account all noise nuisance factors and $A_n \times A_e \times A_d$ is the surrounding area. Considering Φ_n a penalty, a proportional cost (π_n) to this measure could be applied to the operator of the aircraft. Then, the total cost associated with noise is given by:

$$C_n = \pi_n \int_{t_0}^{t_f} \Phi_n dt \tag{37}$$

And finally, the total cost for the operator of the aircraft is:

$$C = \int_{t_0}^{t_f} [\pi_c FF + \pi_d + \pi_n \Phi_n]\, dt \tag{38}$$

B. Cost Index and Noise Index

In order to optimize aircraft trajectories over space and time, a cost index is provided to the flight management system of modern aircraft. The cost index relates the cost of delay to the price of the fuel:

$$CI = \frac{\pi_d}{\pi_c} \tag{39}$$

Economy flights are associated with low values of the cost index while more direct and faster flights are associated with high values of the cost index.

In the same way, a noise index NI can be also introduced:

$$NI = \frac{\pi_n}{\pi_c} \tag{40}$$

Therefore, trajectories with higher Noise Index will minimize noise impact over surrounding populations, while lower Noise Index will give priority to fuel and/or delay considerations.

V. Trajectory Optimization

A canonical formulation of the optimization problem associated with the minimum of noise nuisance flight trajectories is presented in this section and some very preliminary optimization results are also shown.

A. Optimization problem formulation and assessment

Let $\mathbf{x} \in \mathbb{R}^{\mathbf{N_x}}$ and $\mathbf{u} \in \mathbb{R}^{\mathbf{N_u}}$ be the state and control variables respectively as defined in relation 27. The optimal value of control and state variables would be determined by solving the following dynamic optimization problem in the time interval $[t_0, t_f]$:

$$\text{minimize} \quad J(\boldsymbol{x}(t), \boldsymbol{u}(t), t_0, t_f) = \\ = E(\boldsymbol{x}(t_0), \boldsymbol{x}(t_f), t_0, t_f) + \int_{t_0}^{t_f} [F(\boldsymbol{x}(t), \boldsymbol{u}(t), t)]\, dt \tag{41}$$

subject to:

- the dynamic constraints derived in section III:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{u}) \tag{42}$$

- end point or event constraints (i.e. the initial and final boundary conditions):

$$\boldsymbol{e}_L \leq \boldsymbol{e}(\boldsymbol{x}(t_0), \boldsymbol{x}(t_f), t_0, t_f) \leq \boldsymbol{e}_U \tag{43}$$

- mixed state-control path constraints:

$$\boldsymbol{h}_L \leq \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{u}(t), t) \leq \boldsymbol{h}_U \tag{44}$$

- box constraint on the state and control variables:

$$\begin{aligned} \boldsymbol{x}_L \leq \boldsymbol{x}(t) \leq \boldsymbol{x}_U \\ \boldsymbol{u}_L \leq \boldsymbol{u}(t) \leq \boldsymbol{u}_U \end{aligned} \tag{45}$$

Function E can optionally model an extra cost, depending on the initial and final values of the trajectory, that would be associated to airspace capacity and efficiency criteria. On the other hand, function F contains the minimization criterion (as outlined in section IV) as:

$$F(\boldsymbol{x}(t), \boldsymbol{u}(t), t) = FF + CI + NI\Phi_n \tag{46}$$

The vectorial functions $\boldsymbol{e}$ and $\boldsymbol{h}$ will define the event and path constraints respectively and vectors $\boldsymbol{e}_L$, $\boldsymbol{e}_U$, $\boldsymbol{h}_L$, $\boldsymbol{h}_U$, $\boldsymbol{x}_L$, $\boldsymbol{x}_U$, $\boldsymbol{u}_L$ and $\boldsymbol{u}_U$ are respectively the *L*ower and *U*pper values which will bound all constraints. All these functions and variables will be defined in the trajectory constraints component, as explained in section II.

In this problem $\boldsymbol{u} = [\alpha, \mu, N1]^T$ is the input (control) vector and drive the space-time aircraft trajectory. Of course the above problem is far from being trivial and cannot today be solved with accuracy by current on-board computers. Dynamic programming (see, for instance [20]) is one of the possible optimization techniques able to cope with this general optimization problem, but this implies the discretization, either over time or over space, of the whole problem.

As a first step, the optimization problem associated to the presented framework will be solved by using DIDO, a MATLAB application package. DIDO provides tools for solving a broad class of *Smooth and Non-smooth Hybrid Optimal Control* problems defined over a time interval $[t_0, t_f]$ that may be fixed or free. The basic idea behind the solution method is an adaptive algorithm based on a pseudospectral approximation theory [21]. The pseudospectral approach is significantly different from prior methods used to solve such problems [22] and hence the code is a realization of a fundamentally different way [23] of rapidly solving dynamic optimization problems. Currently, DIDO implements approximations of state and control functions in Hilbert spaces and employs the NLP solver SNOPT [24], through TOMLAB [25]. In particular, the type of optimal control problems that can be solved with this package, corresponds with the formulation presented above in equations 41 to 45.

B. Preliminary optimization results

A extremely basic problem has firstly used in order to test the optimization techniques presented above. A hypothetic straight take-off of a four engine aircraft was considered. The trajectory optimization starts at a point where the aircraft reaches V_2 safety operational speed and it is supposed to be at $400\,ft$ above the departing runway elevation. Only the final altitude is fixed to $3000\,ft$, leaving all the other variables free (final position, speed, flight path angle...). Finally some obvious constraints are also implemented in order to restrict for example speed and flight path angle to realistic values.

Aircraft performance data was obtained from BADA data base ([18]) provided by EUROCONTROL

In this first simulation no noise abatement constraints were considered and a mixed fuel-time minimum trajectory was obtained. Figure 3 show the vertical flight profile and figure 4 the speed profile. As it can be easily seen, the obtained optimal trajectory consist in to perform a initial horizontal segment allowing the speed to increase to a optimal value. Then a constant speed climb is performed until the final trajectory altitude is achieved. This result is in perfect accordance with the typical climb profiles performed by airline operators.

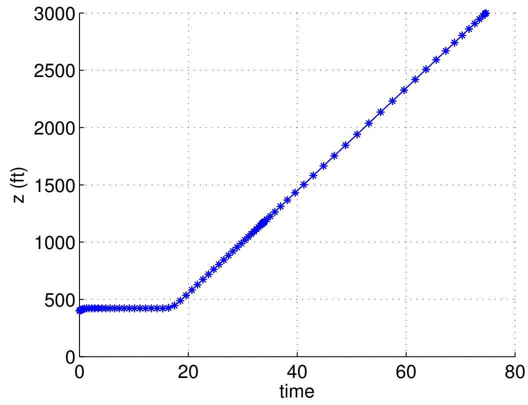


Fig. 3. Vertical flight profile for minimum fuel-time optimization

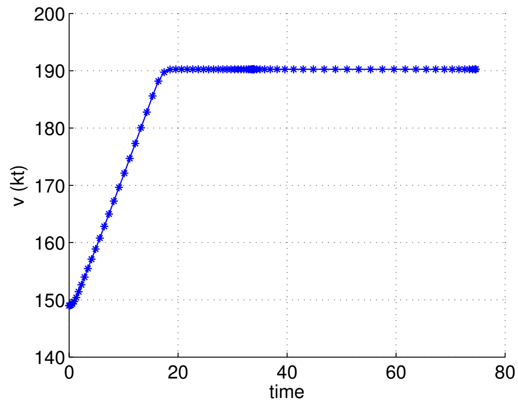


Fig. 4. Speed profile for minimum fuel-time optimization

VI. Conclusion

A framework for a global optimization tool that will take into account aircraft dynamics and performances, noise nuisances and RNAV radionavigation requirements in order to assess an optimum flight depart or approach procedure is presented in this work. The involved scenario defines a given amount of noise nuisance, in function of the emited aircraft noise along its trajectory and geographical and sociological considerations. This value, jointly with some fuel/time economy considerations, conduces to a global optimization criterion that will be minimized obtaining the best flight trajectory satisfying the procedure constraints. In addition, a complete guidance model is obtained, in a state representation form in order to be integrated into the envisaged optimization methods for dynamic systems. The formulation of the problem is formally presented as an optimal control problem, appearing to be a complex issue and requiring specific aerodynamics, propulsive and noise emission models. Further work may deal with the specific noise nuisances model and with the viability of solving the problem by using DIDO MATLAB package and the assessment of the potential use of another additional optimization techniques. First attempts to solve a simplified version of the initial problem are also presented, encouraging further work using the presented methodology.

References

[1] "http://www.eurocontrol.int."
[2] "http://www.ecacnav.com/navigation_strategy.htm."
[3] "http://www.eurocontrol.int/corporate/public/standard_page/cb_atm_strate%gy.html."

[4] *Procedures for Air Navigation Services - Aircraft Operations (PANS-OPS) - Volume I, Flight Procedures*, 4th ed., International Civil Aviation Organisation (ICAO), Montreal (Canada), 1993, doc. 8168-OPS/611.

[5] S. I. Consortium, "Establishment of noise abatement solutions," Report D3, Tech. Rep., 2000.

[6] ——, "Definition of new noise abatement procedures," Work Package 3, Tech. Rep., Mar. 2003.

[7] V. F. Gomez Comendador, "Determinacin de condiciones de compatibilidad en un TMA de tamao medio de procedimientos ACDA de empuje mnimo, y perfil vertical variabla o perfil de velocidad veriable, con procedimeintos convencionales, miniizando su impacto sobre la capacidad," Ph.D. dissertation, Escuela Tcnica Superior de Ingenieros Aeronuticos. Universidad Politcica de Madrid, Madrid (Spain), Feb. 2004.

[8] L. Ren, J.-P. Clarke, and N. T. Ho, "Achieving low approach noise without sacrificing capacity," in *22nd Digital Avionics Systems Conference*, Oct. 2003.

[9] R. Wijnen and H. Visser, "Optimal departure trajectories with respect to sleep disturbance," *Aerospace Science and Technology*, vol. 7, pp. 81–91, 2003.

[10] H. Visser and R. Wijnen, "Optimisation of noise abatement arrival trajectories," *The aeronautical journal*, vol. 107, no. 1076, pp. 607–615, 2003.

[11] *Integrated Noise Model (INM) Version 6.0 Technical Manual*, Office of Environment and Energy, Washington, DC. USA, Jan. 2002, fAA-AEE-02-01.

[12] J.-P. Clarke and R. J. Hansman, "A systems analysis methodology for developing single event noise abatement procedures," MIT Aeronautical Systems Laboratory, Cambridge, Massachusetts (USA), Tech. Rep., 1997, report No ASL-97-1.

[13] K. Feng Zou and J.-P. Clarke, "Adaptative real-time optimization algorithm for noise abatement approach procedures," in *AIAA's 3rd Annual Aviation Technology, Integration, and Operations (ATIO) Technology Conference*, vol. 1, Nov. 2003, AIAA paper No 2003-6771.

[14] H. Visser, "Generic and site spacific criteria in the optimization of noise abatement procedures," *Transportation Research Part D: Transportation and Environment*, vol. 10, pp. 405–419, Sept. 2005.

[15] *Procedures for Air Navigation Services - Aircraft Operations (PANS-OPS) - Volume II, Construction of Visual and Instrument Flight Procedures*, 4th ed., International Civil Aviation Organisation (ICAO), Montreal (Canada), 1993, doc. 8168-OPS/611.

[16] J. Roskam, *Airplane flight dynamics and automatic flight controls*, 3rd ed. Lawrence, Kansas (USA): DARcorporation, 2001, vol. 1.

[17] B. L. Stevens and F. L. Lewis, *Aircraft control and simulation*. London, UK: John Wiley and Sons, INC., 1992.

[18] *User manual for the base of aircraft data (BADA). Revision 3.6*, EUROCONTROL Experimental Centre, Bretigny (France), Sept. 2004, aCE-C-E2.

[19] F. Nejjari, X. Prats, V. Puig, J. Quevedo, M. Polit, B. Ouattara, K. Achaibou, and F. Mora-Camino, "Containing aircraft noise levels at take-off: a mathematical programming approach," in *The 2005 Congress and Exposition on Noise Control Engineering*, Rio de Janeiro (Brazil), Aug. 2005.

[20] R. Bellman, *Dynamic Programming*. Courier Dover Publications, 2003.

[21] I. Ross and F. Fahroo, "A direct method for solving nonsmooth optimal control problems," in *Proceedings of the 2002 IFAC World Congress*, Barcelona (Spain), July 2002.

[22] J. T. Betts, "Survey of numerical methods for trajectory optimization," *Journal of Guidance, Control and Dynamics*, vol. 21, no. 2, 1998, 193-226207.

[23] J. Strizzi, I. Ross, and F. Fahroo, "Towards real-time computation of optimal controls for nonlinear systems," in *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Monterey, California (USA), Aug. 2002, invited Paper no. AIAA 2002-4945.

[24] P. Gill, W. Murray, and M. Saunders, "Snopt: An sqp algorithm for large-scale constrainted optimization," *SIAM Journal of Optimization*, vol. 12, no. 4, pp. 979–1006, 2002.

[25] K. Holmstrom, A. Goran, and M. M. Edvall, *User's guide for TOMLAB 4.0.6*, Tomlab Optimization, Sweden, Aug. 2003.

Exhaust Emissions of Small Aircraft Engines

Paul Arentzen, Ray Taghavi

***Abstract*—This work is a preliminary investigation of the exhaust emission profile of small aircraft engines. Some data has been acquired for the NO_X, CO_2, and CO emissions of two small gas turbine engines and one aircraft reciprocating engine at the Mal Harned Propulsion Laboratory of the University of Kansas. The small gas turbine engines data demonstrate that the NO_X emission factors (g/kg fuel) are considerably lower than of the state of the art larger turbofan engines reported in the literature. This was the case for all operating conditions except idle. The variation in the NO_X emission factor from idle to the maximum load is quite small for the engines tested. The carbon monoxide content in the exhaust from the two engines is low compared to the range of the instrument that measured this content in the experiments. For the aircraft reciprocating engine, the measured fractions of NO_X and CO in the exhaust, during the highly transient engine operating conditions, fell well within the predicted range for emissions from similar engines reported in the literature. The consequence of the high CO emission from this engine is a relatively low CO_2 content in the exhaust, as expected.**

***Index Terms*—Aircraft Engines, Exhaust Emission, Gas Turbines, Pollution.**

I. Introduction

AIRCRAFT engines produce emissions that are similar to other emissions resulting from fossil fuel combustion. However, aircraft emissions are unusual in that a significant proportion is emitted at altitude. These emissions give rise to important environmental concerns regarding their global impact and their effect on local air quality at ground level. Aircraft emit gases and particles, which alter the atmospheric concentration of greenhouse gases, trigger the formation of condensation trails and may increase cirrus cloudiness, all of which contribute to climate change. In this regard, the Kyoto Protocol (1997), which entered into force on 16 February 2005, requires industrialized countries to reduce their collective emissions of six greenhouse gases [1]. The aviation also contributes to the emissions both locally and globally, and intensified efforts to improve the technology and the operational procedures are recommended [2].

Aircraft are required to meet the engine certification standards adopted by the ICAO (International Civil Aviation Organization). As a consequence, they establish limits for emissions of oxides of nitrogen (NO_x), carbon dioxide (CO_2), carbon monoxide (CO), and unburned hydrocarbons. There are also provisions regarding smoke and vented fuel. Of particular relevance is the standard for NO_x, a precursor for depletion of ozone, which at altitude is an ultraviolet blocker. In the direct vicinity of airports, emissions of nitrogen oxides, unburned hydrocarbons, carbon monoxide, and visible smoke contribute to local air quality concerns [3], [4].

In most research and regulatory work the large turbine-based propulsion engines are considered the major sources of exhaust emissions from aircraft, although there are constantly more of the smaller airplanes in operation. Most small gas turbines run at much lower temperature, and therefore the fuel specific NO_X emission should be low, if at all significant, compared to the large turbofans. If the combustion efficiencies of smaller, less advanced engines are lower, their CO emissions would be expected higher than those from bigger engines. We have initiated a research program to study the emissions of small aircraft engines, and our focus is the engine emissions from general aviation aircraft (GA), business jets, unpiloted air vehicles (UAVs), uncertified kitplanes, and homebuilts. Majority of these aircraft have either piston, or small gas turbine engines. Typical fuels consumed by these engines are avgas (leaded gasoline), unleaded gasoline, diesel, or jet fuel. We have acquired some preliminary data on NO_x, CO_2, and CO emissions of an aircraft piston engine and two small gas turbine engines at the Mal Harned Propulsion Laboratory of the University of Kansas. The purpose of this paper is to report the results of the analysis of this data and explain the future plans for these efforts.

II. Experimental Facility, Engines & Instrumentation

The tests were conducted in the Mal Harned Propulsion Laboratory of the University of Kansas. This facility is located in a 360 m^2 hangar space area at Lawrence Municipal Airport. The test cell is equipped for testing turbojet, turbofan, turboprop, and aircraft reciprocating engines. It is a concrete structure, 3.7 m wide, 7.3 m long and 3.2 m high and could be opened at each end to the atmosphere.

A. Engines

Three types of engines were tested for emissions. These are a Teledyne Continental O-470M Aircraft Reciprocating Engine, a Williams International WR-24-6 turbojet engine, and a Garrett GTC85 series APU gas turbine engine.

1) Aircraft Piston Engine: The engine used for this test was a 178 kW, Teledyne Continental O-470M, 6-cylinder,

Manuscript received February 13, 2006.

Ray Taghavi is with the Department of Aerospace Engineering, University of Kansas, Lawrence, KS 66045 USA (corresponding author to provide phone: 785-864-2973; fax: 785-864-3597; e-mail: rtaghavi@ku.edu).

Paul Arentzen is with the Faculty of Engineering and Science, Agder University College, N-4879 Grimstad, NORWAY (paul.arentzen@hia.no).

opposed, air-cooled aircraft engine. The engine has a total piston displacement of 7,700 cm^3 and is naturally aspirated. It is equipped with a Bendix Energy Systems, model PSD-5c single throttle valve, down-draft, pressure injection carburetor. The ignition system consists of two Bendix Energy Systems, model S6RN-20 impulse coupled magnetos. An aluminum 2-bladed Hartzell propeller having a diameter of 1.96 m is installed on the engine. The engine on the test stand is shown in Fig. 1.

2) The Jet Gas Turbine Engine: The gas turbine engine used for this laboratory is a 0.45 kN thrust Williams Research WR 24-6 gas turbine engine. The engine was originally designed to be used on military drones and missiles. It has a single stage centrifugal compressor and single stage axial turbine. The inlet is an ASME standard bellmouth with a pressure recovery of about 99%. The engine performance parameters are measured by various static and total pressure taps, thermocouples, fuel flow meter, sound pressure level meter, etc. Maximum compressor exit pressure was 270 kPa_{abs} during the present tests while the exhaust gas temperature peaked at around 1000 K. The engine on the test stand is shown in Fig. 2.

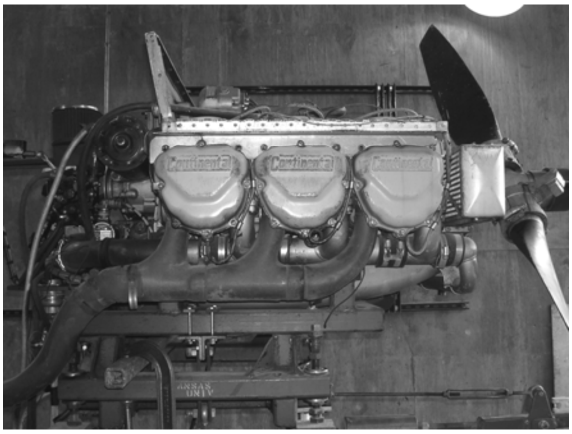


Fig. 1. Teledyne Continental O-470M aircraft piston engine on the test stand.

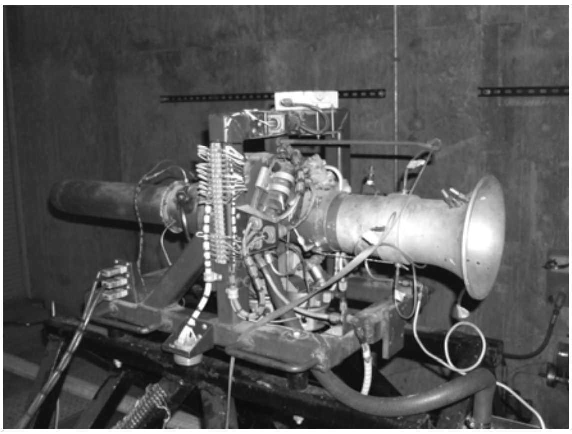


Fig. 2. Williams Research WR 24-6 gas turbine engine on the test stand.

3) The APU Gas Turbine Engine: The APU (auxiliary power unit) engine used for this test was a 149 kW GTC85 series gas turbine engine manufactured by Garrett Turbine Engine Company, a division of Garrett Corporation. This engine is designed to provide power in the form of compressed bleed-air for the operation of the main jet engine starters, air conditioning systems, and other consumers of compressed air. The engine is equipped with a two-stage centrifugal compressor assembly. The compressor section is the source of compressed air for the turbine section and of bleed-air for pneumatic power. Maximum exhaust gas temperature at 880 K was observed at maximum air delivery (valve fully open) while the compressor exit pressure was at it lowest, only 395 kPa_{abs}. The APU engine is shown in Fig. 3.

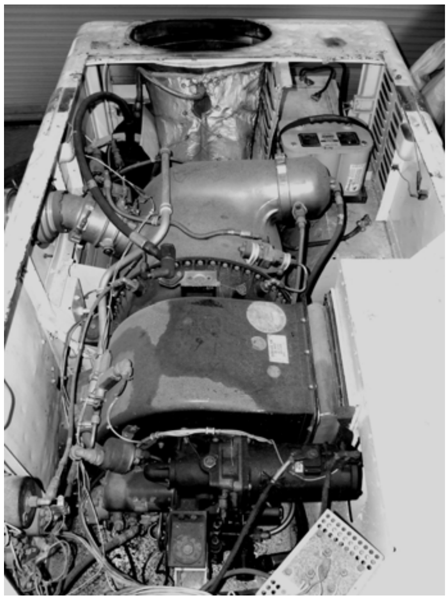


Fig. 3. Garrett GTC85 Series APU engine.

B. *Instrumentation*

Engine exhaust gas, called the sample gas, was taken form the immediate exit of the exhaust pipe/nozzle through a stainless steel probe. A vacuum pump pulled the sample gas through a stainless steel heat exchanger and into the sample gas manifold. The sample gas was then led to the three separate instruments that measured carbon dioxide, carbon monoxide, and nitric oxides concentrations (volume fractions) in the gas.

Two Beckman Industrial Model 870 Non-Dispersive Infrared Analyzers were used for CO_2 and CO measurements. The carbon dioxide instrument has a maximum range of 30 %, and the carbon monoxide range is 19.2 %. The concentrations of the particular components are detected by differences in absorption of infrared light. The volume percentages are displayed on the front of the instruments.

The content of nitric oxides was measured in a Thermo Environmental Instruments Inc. Model 10A Chemiluminescent

Fig. 4. Emission measuring instrumentation.

NO-NO_X Gas Analyzer. This instrument can operate at full-scale ranges from 2.5 to 10,000 ppm. The light emission from the chemical reaction between NO and ozone is used to detect the NO_X concentration in the sample gas. The concentration is displayed on the front of the instrument. A Model 800 Heated NO_X Sampler from Thermo Electron Corporation was also used for sample gas treatment.

All three instruments were calibrated using calibrating gases as prescribed in the instrument manuals. The instrumentation rack is shown in Fig. 4.

III. Results

1) Gas Turbine Engines: The two engines that were tested in these experiments are small gas turbine engines compared to modern turbofan engines on civil transport aircraft. The engines are designed for different applications and engine internal pressures and temperatures are not as high. These two engines are older technology.

Carbon dioxide in the exhaust was measured primarily as a reference figure to quantify the fuel/air mixture ratio in the combustion chamber. Assuming that the un-burnt hydrocarbon content is insignificantly low, the exhaust CO_2 and CO gases together contain all the fuel carbon. By mass balance calculation based on a known fuel compositions [9], the fuel to air ratio was quantified.

The present experiments and preliminary measurements done on these small engines demonstrate that the NO_X emission factors (g/kg fuel) are considerably lower than for state of the art larger turbofan engines in operation [5], [6], [7], only 20 – 25 % at full load/take-off (Fig. 5). This is the case for all conditions except idle. The variation in the NO_X emission factor from idle to the maximum load is quite small for the engines in the test.

The Williams turbojet engine is throttled by increasing the engine RPM, whereas the Garrett engine, a constant RPM engine, takes more load as the valve for usable high-pressure air is opened. In the turbojet the airflow through the combustion zone increases with the engine load, and in the Garrett APU engine this airflow, calculated from our measurements, appears to be constant or decrease with higher engine load. Hence, in the jet an overall fuel leaner combustion is maintained with increasing temperature, and for the APU engine the mixture in the combustion chamber becomes successively richer with the higher temperature. This may be the explanation why the NO_X emission factor increases with load for the jet whereas it drops off for the APU.

The carbon monoxide content in the exhaust from the two engines is low compared to the range of the instrument that measured this content in the experiments. This had an impact on the accuracy of the CO preliminary measurements for these engines. It is still apparent from the graph in Fig. 6 that the carbon monoxide content is of the same order of magnitude for the small and old engines as for a modern turbofan engine [7]. Further experiments are necessary to quantify exact figures for the carbon monoxide emissions from small gas turbine engines.

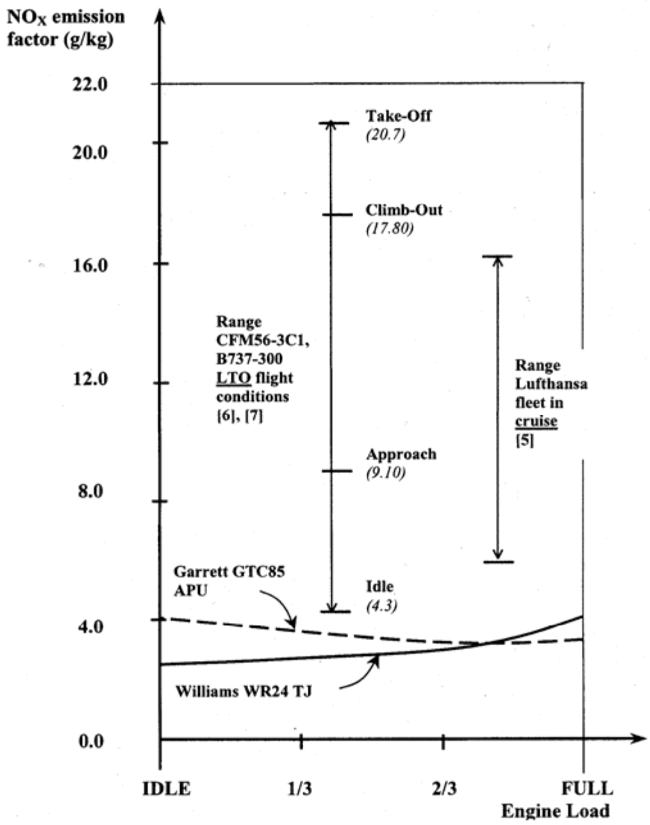


Fig. 5. NO_X emission in gram pr. kilogram fuel for two small gas turbine engines compared to similar for representative civil transport aircraft engines.

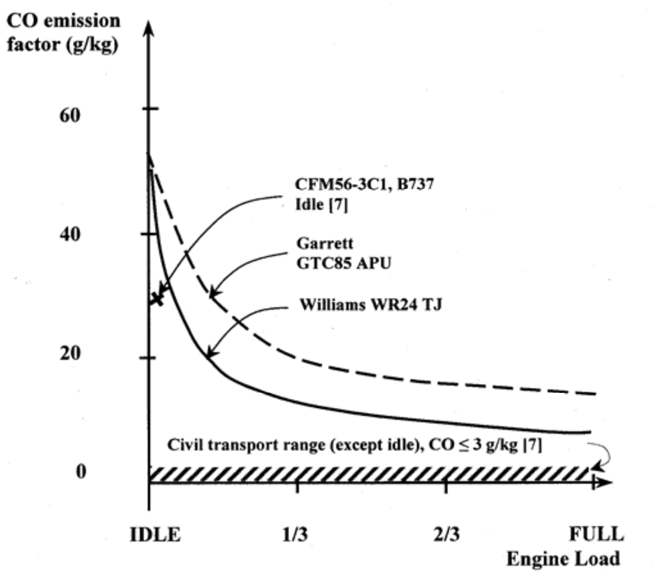


Fig. 6. CO emission in gram pr. kilogram fuel for two small gas turbine engines compared to similar for representative civil transport aircraft engines.

2) Reciprocating Engine: The reciprocating engine was operated in two starting and warm up sequences, the first a cold engine start at ambient temperature at 10 – 12 ºC and the second when engine had cooled for about one hour since last run. Each operation (start to stop) lasted a few minutes. In both cases the engine was set at the highest RPM and stepwise slowed down to the specific speeds while the engine manifold pressure was reduced accordingly.

The measured ratios of NO_X and CO in the exhaust for these transient engine conditions, fall well in the predicted ranges for emissions from a similar engine category and operation [8], as described in Fig. 7.

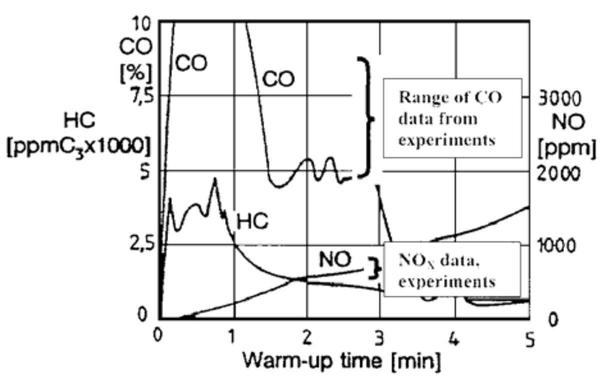


Fig. 7. NO_X and CO emission (volume fractions) from current measurement on reciprocating engine compared to similar data in fig. 2.11 of [8].

The calculated NO_X and CO emission factors (g/kg fuel) are shown in the graphs of Fig. 8. With a cold start the CO emission factor increases with time, and as the engine speed is

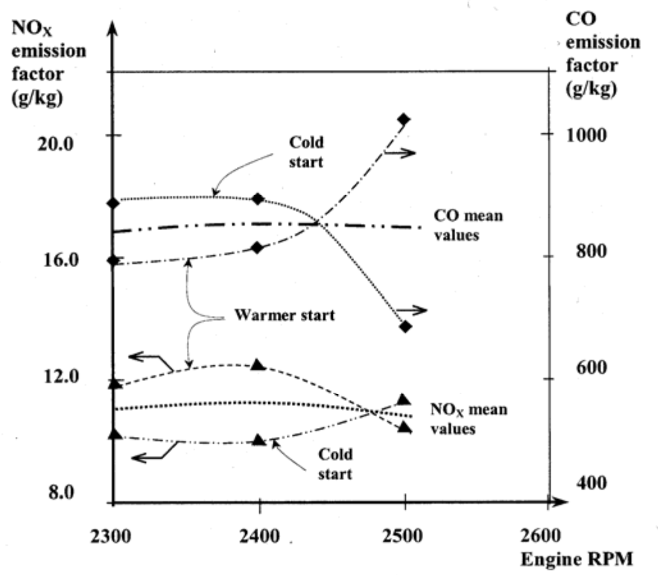


Fig. 8. NO_X and CO emission in gram pr. kilogram fuel for Continental O-470M reciprocating engine in the start up sequence.

reduced. Starting the engine warmer, the opposite trend is observed.

The NO_X emission factor decreases with time at cold start and while the engine speed is reduced. It increases with time after the engine was started slightly warmer.

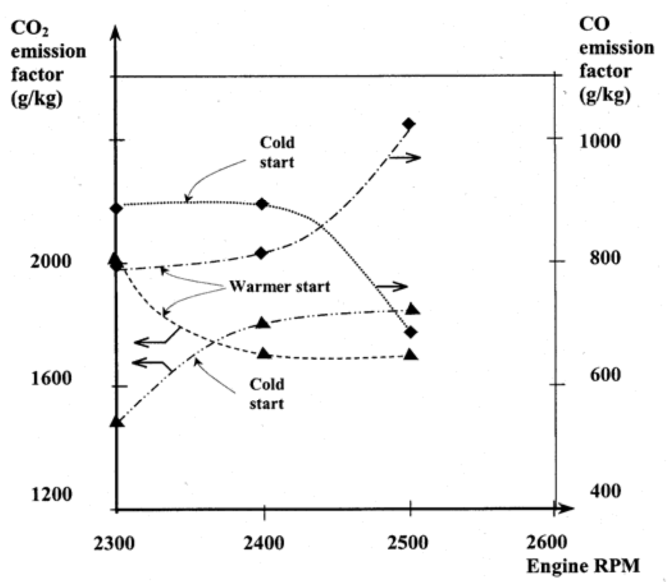


Fig. 9. CO_2 and CO emission in gram pr. kilogram fuel for Continental O-470M reciprocating engine in the start up sequence.

Unlike for the gas turbine engines, the carbon monoxide content is high, at the same order of magnitude as the content of carbon dioxide. The CO_2 is therefore much lower than if a complete combustion had taken place. Corresponding figures for CO_2 and CO emission factors are shown in Fig. 9.

IV. Further Work

The main points of focus in the continuation of this work on small aircraft engine emissions will be to improve the carbon monoxide measurements, to include modern technology engines in the experiments, and to study variations with more engine and ambient parameters. The capability to also measure water vapor, un-burnt hydrocarbons and particulates in the exhaust will be accommodated, and the variation in the emission concentrations outside the exhaust exit area, in the flow direction and across the exhaust plume, will be measured. Change of fuel, particularly into natural gas and hydrogen, and how that can influence the exhaust emissions of small gas turbine engines will be investigated. Measurements of the emissions from aircraft reciprocating engines using leaded fuel, diesel fuel and jet fuel will also be performed.

V. Conclusion

Exhaust emissions from two small gas turbine engines and one reciprocating engine were measured. The gas turbine engines are small engines compared to modern turbofan engines on civil transport aircraft. These experiments demonstrate that the NO_X emission factors from the small gas turbine engines are considerably lower than for state of the art larger turbofan engines, except at idle condition. They also show that the variation in the NO_X emission factor between idle and maximum load is quite small for the gas turbines in this test. For increasing engine loads the NO_X emission factor increases for the jet and decreases for the APU, constant RPM engine.

The carbon monoxide content in the exhaust from the two engines is very low in the range of the instrument that was used for the measurements. It is concluded however, that the carbon monoxide content is of the same order of magnitude for the small and older engines as for a modern turbofan engine. Further experiments are necessary to determine more precise numbers for the carbon monoxide emissions from small gas turbine engines.

The reciprocating engine was tested in transient conditions, warming up immediately after start. The measured fractions of NO_X and CO in the exhaust for these conditions fall well in the predicted ranges for emissions from similar experiments that are documented. Unlike for gas turbine engines the carbon monoxide content is high, at the same order of magnitude as the content of carbon dioxide.

References

[1] http://www.icao.int/icao/en/env/aee.htm
[2] Committee on Aeronautics Research and Technology for Environmental Compatibility, "For Greener Skies, Reducing Environmental Impacts of Aviation", National Academy Press, Washington, D.C. 2002, ISBN 0-309-08337-0
[3] http://www.iata.org/whatwedo/environment/aircraft_emissions.htm
[4] http://www.epa.gov/otaq/aviation.htm
[5] H.-G. Nüßer, A. Schmitt, "The Global Distribution of Air Traffic at High Altitudes, Related Fuel Consumptions and Trends", Proceedings of a DLR International Colloquium, Editor U. Schumann, Bonn, Germany, November 15/16, 1990, Springer-Verlag.
[6] CFM International publications (Company Brochures)
[7] Australian Government, Department of the Environment and Heritage, NPI: "Emission Estimate Technique Manual for Airports" Version 1.1, May 2001, ISBN: 06 4254 6878
[8] H. P. Lenz, "Mixture Formation in Spark-Ignition Engines", Springer-Verlag, SAE, Wien-New York 1990
[9] http://www.chevron.com

Session 2

Airport Operations

Airlines’ point of view as a new approach to measuring quality of service at Airport Bratislava

Benedikt Badánik, *University of Žilina*

Department of Air Transport, Žilina, Slovak Republic

***Abstract*— The purpose of this article is to provide airport professionals and those which do not have experience in the field of service quality at airports with information on new approach to measuring quality from airlines’ point of view. It is a brand new approach due to previous approaches from the customers’ side. This new approach has been tested at Airport Bratislava. Research results will be published. The information included in this article was partly provided by the University of Žilina survey. Other important information was obtained from representatives of the Airport Bratislava, Air Slovakia, Air Traffic Services of the Slovak Republic, Czech Airlines, SkyEurope Airlines and Slovak Airlines.**

***Index Terms*—Airport service, Quality control, Airport operation**

I. Introduction

It is obvious that customers of airports are divided into five main groups: airlines; passengers; concessionaires; meeters, greeters, visitors, personnel and non-travellers. Passengers are the largest group of all, but sometimes are considered by the airlines as their customers, and therefore only indirect customers of the airport. But the airlines can be still considered as the primary customer of the airport: the major facilities (runways, taxiways, apron, terminal facilities…) have been built for their use. They pay for the services provided. They also use and pay for the office and technical space required for their staff and operations.

Therefore, the aim of the article is to bring new vision; airline’s point of view as a new approach to measuring quality of service at airports.

Manuscript received February 26, 2006.

Benedikt Badánik is an Assistant Lecturer at the University of Žilina, Faculty of Operation and Economics of Transport and Communications, Department of Air Transport, Univerzitná 1, 010 26 Žilina, Slovak Republic (corresponding author to provide phone: +421-41-513 3457; fax: +421-41-513 1517; e-mail: Benedikt.Badanik@fpedas.utc.sk).

II. CURRENT APPROACH TO MEASURING QUALITY OF SERVICE AT AIRPORTS (PASSENGER’S POINT OF VIEW)

An airport is perceived as a key point in the air transport system. The efficiency and speed of the processes at an airport are critical. This is usually summed up by the term “facilitation”, i.e. giving free and unimpeded passage to aircraft, passengers, freight and mail, including all clearance and handling processes. One of the facets of facilitation is quality of service. In order to satisfy the airport’s customers, it is important to provide the best service possible, according to customer needs. In order to verify that the desired service quality has been achieved, it has to be measured, evaluated and also anticipated.

ACI quality of service survey uses two kinds of measurement regarding quality of service. Objective measurement, which is provided by the measurement of defined criteria, with indicators which help in achieving objective measures (objective criterion is one which is measured objectively, e.g. a time measurement) and subjective measurement which depends on the subjective value attributed to quality of service by passengers (given by surveys, comment cards, or complaints).

Example of objective service quality criteria:

Item: Aircraft turn-around process

Indicator: aircraft turn-around time
Measurement: difference between Actual Time of Arrival (ATA) and Actual Time of Departure (ATD) (computerized data)
Target/objective: depends upon the type of aircraft and the airline service choice

Example of subjective service quality criteria:

Item: Overall customer satisfaction at the airport/overall attractiveness/convenience of airport/overall quality of service

Based upon ACI quality of service survey (realized by ACI world headquarters Geneva – Switzerland in 1998), 95 % of objective criteria and more than 99 % of subjective criteria used to measure quality of service at airports worldwide is related to passengers.

Just a little of the following services which are found in an airport related to airlines have been criterions of measuring quality:

- offices and desks (or generally surface areas)
- terminal resources: check-in desks and baggage belts, gate allocation (aircraft stands: at contact or remote)
- information technologies and telecommunications
- ground-handling services
- movement areas (runways and apron areas)
- technical facilities and services
- signage and guidance, and way-finding
- announcements
- information (including flight information)
- comfort (architecture, volumes, temperature, visual environment, smoking areas)
- provision of washrooms and toilets
- cleanliness
- staff courtesy, empathy, contact, accuracy (appropriate staff) and efficiency capacity
- “walking” times (connecting flight flow, embarkation or disembarkation flow)
- availability of lifts, escalators, moving walkways, people-movers, etc.
- provision for the disabled
- special services (business lounge, VIP, facilities for religious observance, emergency medical services, etc.)

Therefore, most of the airports worldwide currently apply passenger approach to measuring quality of service. The weak point of measuring quality of service at airports has been revealed.

III. NEW APPROACH TO MEASURING QUALITY OF SERVICE AT AIRPORTS (AIRLINES’ POINT OF VIEW)

New approach means that information related to quality of service at airports will not be obtained from passengers, but from pilots – first representatives of airlines.

Due to the fact that the previous list of services doesn’t include the areas of services, quality of which can easily be measured (subjectively) by pilots, it is needed to specify those areas (see Fig.1). Quality of service at airports will be measured from the airlines’ point of view.

This part of research has been done with support of research partners: Airport Bratislava, Air Slovakia, Air Traffic Services of the Slovak Republic, Czech Airlines, SkyEurope Airlines and Slovak Airlines.

.

A. QUESTIONNAIRE DRAFT

After the areas of research on measuring quality of service at airports from the airlines’ point of view had been selected, questionnaire draft with questions related to the specific area was prepared (Fig. 2 below). Questionnaire is the best way of collecting data within the surveys. Its main advantage compared to other ways of collecting data is transparency of data, which enables its easy usage in the future. Four levels of satisfaction (quality marks) with quality of service have been determined.

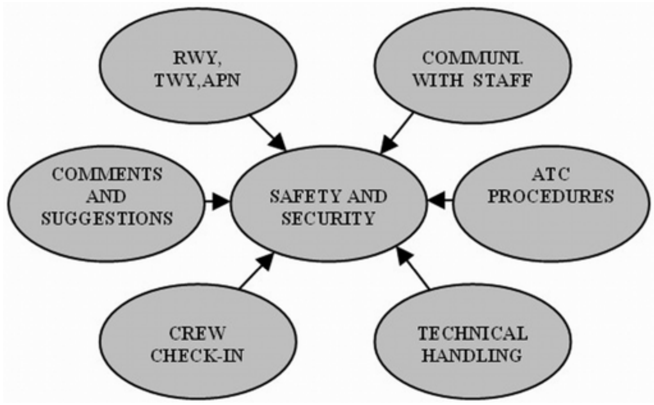


Fig. 1 Areas of research on quality of service measuring from airlines’ point of view

4	Very good
3	Good
2	Fair
1	Poor

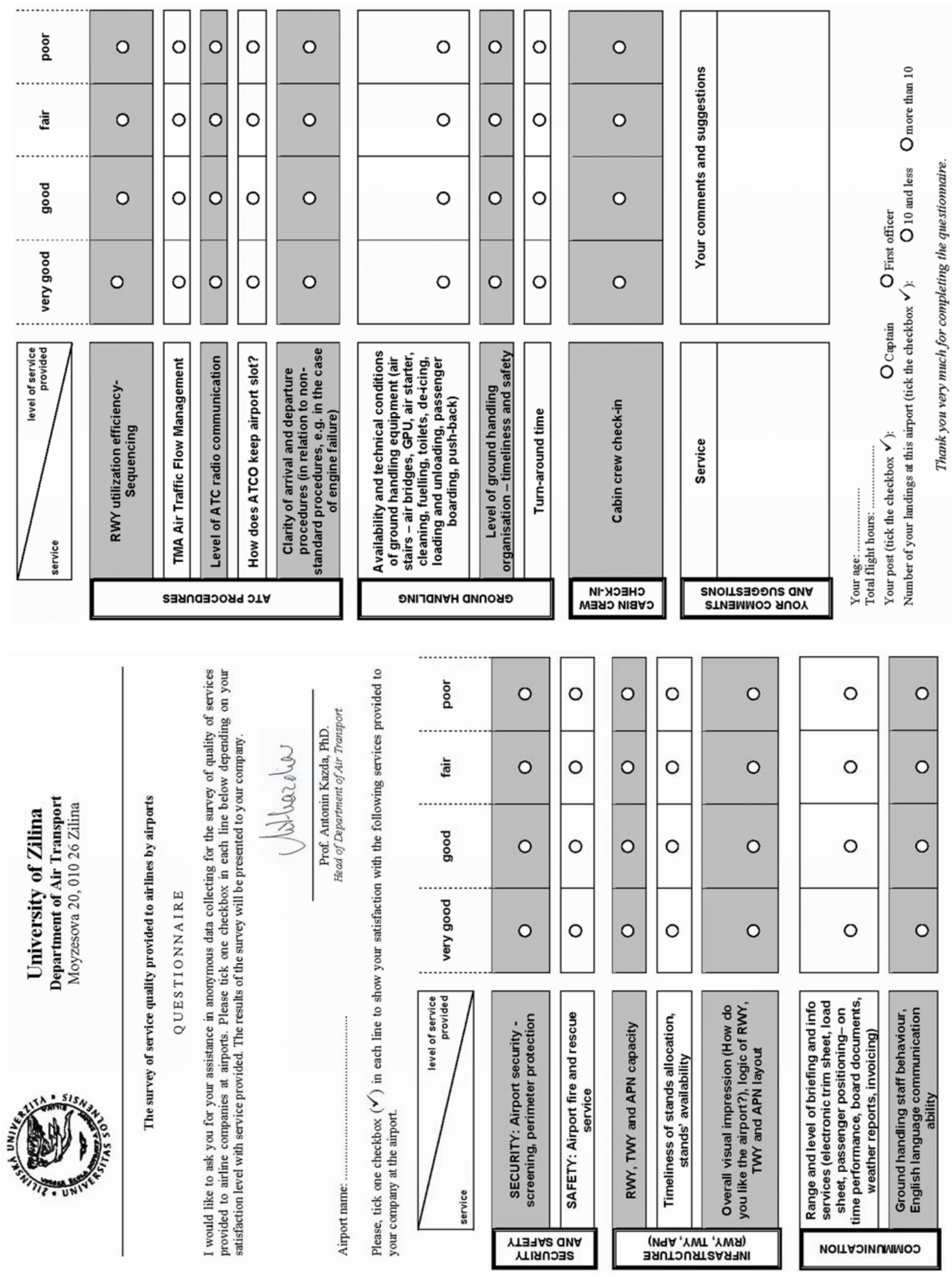
There are only subjective criterions applied in the questionnaire. We couldn’t manage such an objective criterions within the research partners that would suit every partner as well as airport researched.

IV. QUALITY OF SERVICE AT AIRPORT BRATISLAVA

A. RESEARCH FACTS AND FIGURES

56 pilots measured quality of service at Bratislava Airport. Their age ranged from 23 to 58 years and their total flight hours ranged from 500 to 20000. They had been selected as representatives of three of the five strongest airlines (according to passenger transported in 2004) at Airport Bratislava (as shown in the Fig. 5) – SkyEurope Airlines, Slovak Airlines and Air Slovakia. 14 questionnaires had to be rejected due to incompleteness or discrepancies.

Figure 3 shows indifference within the measuring quality of service at Airport Bratislava by pilots. Level of satisfaction is in the range from 1, 8 to 3, 6. According to Fig. 4, visual impression was “the best” average rated service at Airport Bratislava. Level of English language proficiency was “the worst” average rated service at Airport Bratislava.



University of Zilina
Department of Air Transport
Moyzesova 20, 010 26 Zilina

The survey of service quality provided to airlines by airports

QUESTIONNAIRE

I would like to ask you for your assistance in anonymous data collecting for the survey of quality of services provided to airline companies at airports. Please tick one checkbox in each line below depending on your satisfaction level with service provided. The results of the survey will be presented to your company.

Prof. Antonin Kazda, PhD.
Head of Department of Air Transport

Airport name: ....

Please, tick one checkbox (✓) in each line to show your satisfaction with the following services provided to your company at the airport.

	service / level of service provided	very good	good	fair	poor
SECURITY AND SAFETY	SECURITY: Airport security - screening, perimeter protection	O	O	O	O
	SAFETY: Airport fire and rescue service	O	O	O	O
INFRASTRUCTURE (RWY, TWY, APN)	RWY, TWY and APN capacity	O	O	O	O
	Timeliness of stands allocation, stands' availability	O	O	O	O
	Overall visual impression (How do you like the airport?), logic of RWY, TWY and APN layout	O	O	O	O
COMMUNICATION	Range and level of briefing and info services (electronic trim sheet, load sheet, passenger positioning– on time performance, board documents, weather reports, invoicing)	O	O	O	O
	Ground handling staff behaviour, English language communication ability	O	O	O	O

	service / level of service provided	very good	good	fair	poor
ATC PROCEDURES	RWY utilization efficiency-Sequencing	O	O	O	O
	TMA Air Traffic Flow Management	O	O	O	O
	Level of ATC radio communication	O	O	O	O
	How does ATCO keep airport slot?	O	O	O	O
	Clarity of arrival and departure procedures (in relation to non-standard procedures, e.g. in the case of engine failure)	O	O	O	O
GROUND HANDLING	Availability and technical conditions of ground handling equipment (air stairs – air bridges, GPU, air starter, cleaning, fuelling, toilets, de-icing, loading and unloading, passenger boarding, push-back)	O	O	O	O
	Level of ground handling organisation – timeliness and safety	O	O	O	O
	Turn-around time	O	O	O	O
CABIN CREW CHECK-IN	Cabin crew check-in	O	O	O	O
YOUR COMMENTS AND SUGGESTIONS	Service	Your comments and suggestions			

Your age: ....
Total flight hours: ....
Your post (tick the checkbox ✓): O Captain O First officer
Number of your landings at this airport (tick the checkbox ✓): O 10 and less O more than 10

Thank you very much for completing the questionnaire.

Fig. 2 Quality of service measuring questionnaire

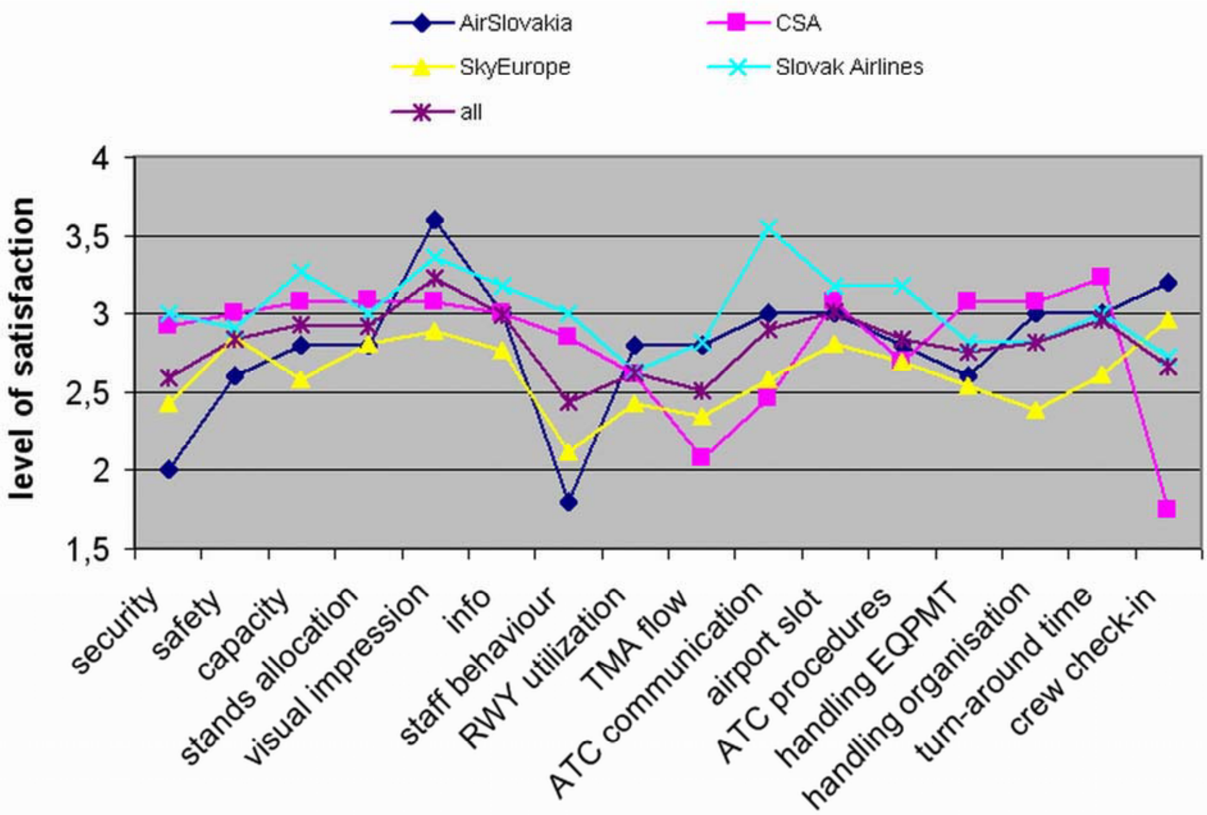


Fig. 3 Level of satisfaction with quality of service at Airport Bratislava measured by pilots

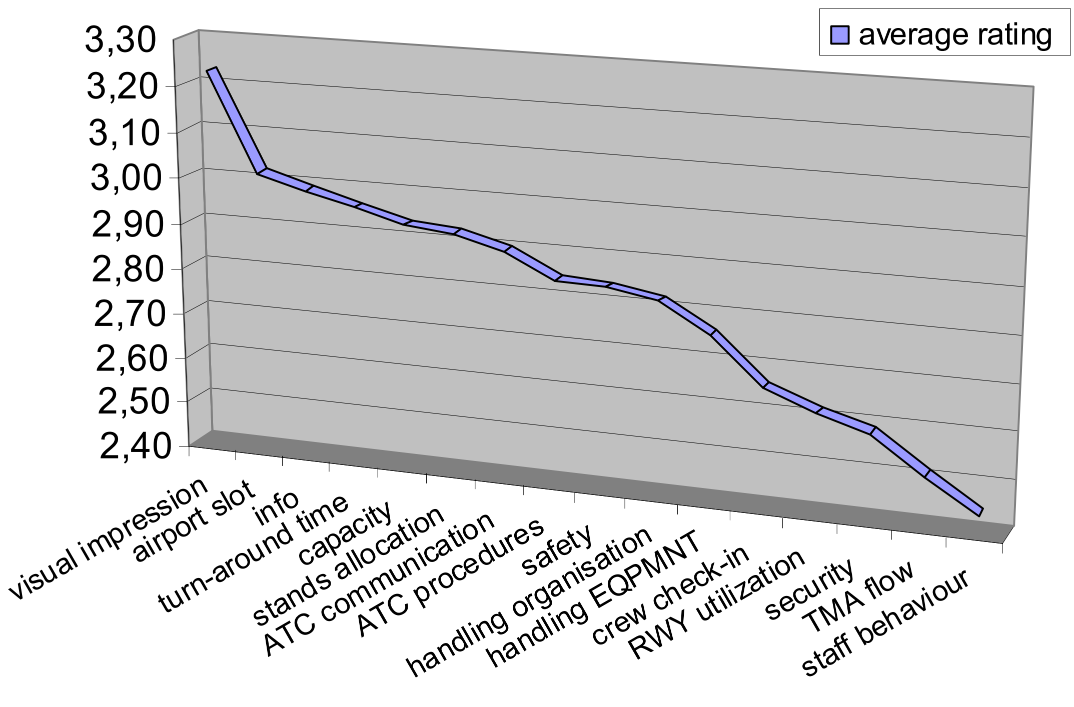


Fig. 4 “The best” and “the worst” average rated service at Airport Bratislava

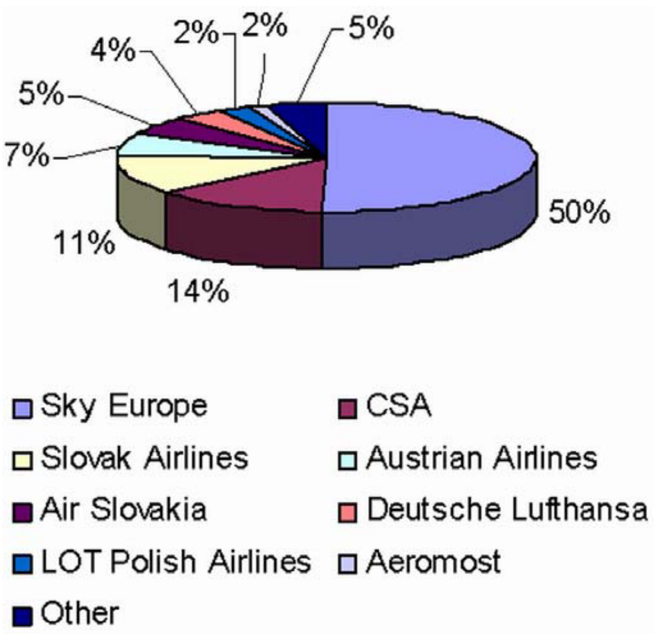


Fig. 5 Percentage of passenger transported at Airport Bratislava (2004)

B. QUALITY OF SERVICE MEASURING INDEPENDENCE (from pilot's age)

Older pilots or more experienced (according to total flight hours) are sometimes considered to be more sceptical in evaluation quality of service at airports. It can be concluded that evaluating quality of service is dependent on pilot's age or total flight hours. So that it is needed to prove independence of such an evaluation.

According to the fact above, research results have been scanned to prove that evaluating quality of service at Airport Bratislava by pilots is independent from their age and total flight hours. Calculating of statistical data (needed to determine independence described above) is demonstrated on example of security service at Airport Bratislava. To demonstrate calculation, any other service can be used. The same methodology would be applied.

Tab. 1 shows absolute number of quality marks within the ranges of pilot's age. It means how many pilots in each range of age put quality mark within the range from 1 to 4.

	pilot's age				
marks	**23-31**	**32-40**	**41-49**	**50-58**	**total**
I.	1	2	0	0	**3**
II.	8	4	4	1	**17**
III.	7	7	3	2	**19**
IV.	1	1	1	0	**3**
total	**17**	**14**	**8**	**3**	**42**

Tab. 1 Absolute numbers of quality marks

Computed theoretical values of random quantity shows Tab. 2 .

a_i	3	17	19	3
b_j	17	14	8	3
t_{ij}	**1,21**	**5,66**	**3,62**	**0,21**

Tab. 2 Computed theoretical values of random quantity

T_1 = 10, 71 (according to standard chi-square test methodology). Number of degrees of freedom has been found consequently $\nu = (4-1)\cdot(4-1) = 9$ (it applies for Tab. 3). Critical value of χ^2 distribution (if $\nu = 9$, α/2=0,025) is $\chi^2_{\alpha/2}(\nu)$=19. T_1 (10, 71) < 19 then, consequently $T_1 > \chi^2_{\alpha/2}(\nu)$ is not valid. Hypothesis tested (that measuring quality of service at Airport Bratislava by pilots is independent from their age) is accepted. Because of level of importance α = 0, 05, it is concluded that measuring quality of service by pilots is 95 % independent from their age. All the data obtained from the research are usable without restriction. and it is not needed to correct it due to dependency on pilots' age.

C. QUALITY OF SERVICE MEASURING INDEPENDENCE (from pilot's total flight hours)

According to the previous independence testing, the same methodology is applied. The aim is to prove independence of measuring quality by pilots from their total flight hours.

Tab. 3 gives absolute number of quality marks within the ranges of pilots' total flight hours.

	total flight hours (thousands)				
marks	**0,5 - 3**	**3 - 6**	**6 - 9**	**9 more**	**total**
I.	1	1	1	0	**3**
II.	9	6	2	0	**17**
III.	6	9	1	3	**19**
IV.	1	2	0	0	**3**
total	**17**	**18**	**4**	**3**	**42**

Tab. 3 Absolute number of quality marks (total flight hours related)

a_i	3	17	19	3
b_j	17	18	4	3
t_{ij}	**1,21**	**7,29**	**1,81**	**0,21**

Tab. 4 Theoretical numbers of random quantity

T_1 = 10, 52 (according to standard chi-square test methodology). Number of degrees of freedom has been found consequently $\nu = (4-1)\cdot(4-1) = 9$ (it applies for Tab. 5). χ^2 distribution critical value (if $\nu = 9$, α/2=0,025) is $\chi^2_{\alpha/2}(\nu)$=19. T_1 (10, 52) < 19 then, consequently $T_1 > \chi^2_{\alpha/2}(\nu)$ is not valid. Hypothesis tested (that measuring quality of service at Airport Bratislava by pilots is independent from their total flight hours) is accepted. Because of level of importance α = 0, 05, it is concluded that measuring quality of service by pilots is 95 % independent from their total flight hours.

	age dependency	total flight hours dependency
safety	9,02	7,21
capacity	9,31	7,88
stands allocation	8,79	6,79
visual impression	7,62	6,29
info	8,52	6,43
staff behaviour	11,79	12,45
RWY utilization	10,55	9,50
TMA flow	10,57	9,81
ATC communication	8,81	8,05
airport slot	8,43	6,14
ATC procedures	9,05	7,52
handling EQPMT	10,19	9,81
handling organisation	10,40	9,55
turn-around time	9,29	7,19
crew check-in	8,36	6,93

Tab. 5 Testing characteristics values table

Pilots measured 16 services provided to airlines at Airport Bratislava. It has been researched that no measurement has been dependent on total flight hours or age of pilots (as it is proved in the Tab. 5, where testing characteristic values T_1 (age dependency and total flight hours dependency) are smaller than 19 (critical value of χ^2 distribution within the level of importance α = 0, 05)).

V.CONCLUSIONS

Present fast growth of number of passengers transported turns Airport Bratislava into a stronger position on “small regional airports” market in Europe. Average growth of passenger transported is more than 30 % annually. Therefore Airport Bratislava will face increasing demand for quality of service in the near future. More low-cost carriers will consider Airport Bratislava great hub and gate to Eastern Europe.

Should Airport Bratislava stay attractive for low-costs it will need to improve most services: first of all security and safety (as research results proved). Quality of service needs to be measured periodically. An airline and its needs and expectations needs to be the midpoint.

Methodology described should be the handbook to perform measuring quality of service at Airport Bratislava within more airlines. According to that methodology comparison between Airport Bratislava and similar airport should be done.

REFERENCES

[1] ACI Airports Council International: Quality of service at airports: Standards & measurements, First edition – 2000, Published by ACI World Headquarters Geneva – Switzerland

[2] ASHFORD, Norman, STANTON, Martin P H, MOORE, A Clifton: Airport operations, Loughborough, Leicestershire Crowthorne, Berks Llano, California 1983, ISBN 0 273 03229 1

[3] AZZOLINI, Cesare: Can Quality and Profitability Go Hand In Hand? Airports Council International 1 st CIVIL AVIATION FORUM 2 nd WINTER SERVICE WORKSHOP PAPER, Budapest 1993

[4] BADÁNIK, Benedikt: Aspekty kvality v prevádzke letísk, Písomná práca k dizertačnej skúške Žilinská univerzita v Žiline, f PEDaS KLD. Žilina 2002

[5] DOGANIS, Rigs: The airport business., 1992, ISBN 0-415-08117-3

[6] FRANCIS, Graham, HUMPHREYS, Ian, FRY, Jackie: An international survey of the nature and prevalence of quality management systems in airports ISSN 1478-3363, 2003

[7] GRAHAM, Anne: Managing Airports. An International perspective, Chapter IV – Service quality and its measurement. 1. Title ISBN 0 7506 4823 6

[8] GRAHAM, Anne: Managing Airports. An International perspective, Chapter V – The airport-airline relationship. 1. Title ISBN 0 7506 4823 6

[9] HEINZELMANN, Utz: Can Quality and Profitability Go Hand In Hand? Airports Council International 1 st CIVIL AVIATION FORUM 2 nd WINTER SERVICE WORKSHOP PAPER, Budapest 1993

[10] IVANCEVICH, John M., LORENZI, Peter, SKINNER, Steven J. with CROSBY, Philip B.: Management quality and competitiveness, III. Title ISBN 0-256-12453-1, 1994

Benedikt Badánik was a Member of Slovak Airport Authority Operational Transformation Team (2004) and CEO of Airport Žilina (2005).

He holds MSc. Degree in Air Transport and PhD. Degree in Airport Operation. Main research interests concern surveys on quality of service at airports, airport operation and airline marketing practices.

He has published in air transport journals and conference proceedings.

Will Intermodal Transport Create Additional Passenger Bottlenecks at Various Choke-Points in the Airport Terminal?

Antonia Cokasova

***Abstract*—Number of studies exists that propose transferring some short haul flights to high-speed train in aim to release runway and ATC resources, while offering immediate relief to congestion, reduced negative environmental impact, and finally improved ground access to airports [7, 8, 9].**

In Europe there are very few examples of transport shift from short haul air traffic to high-speed rails, therefore little proof exists demonstrating the benefits derived from intermodal way of transport resulting into increased airport available resources.

Our approach consists of several assumptions; firstly passenger perspective is the key element. It is rather impossible to develop a well-organized intermodal node if it does not work towards passenger satisfaction. Secondly the layout of intermodal terminal can strongly influence the outcome of integrated intermodal services' efficiency.

Despite some potential benefits the intermodal passenger flow could possibly create additional bottlenecks at different choke-points, since passenger movement will not be smoothened by check-in services. Movement of passengers at intermodal airport terminal is a rare phenomenon; the little information gathered is based on guess work and assumptions.

***Index Terms*—Intermodality, passenger movement simulation, intermodal airport terminal, transport shift**

I. Introduction

With increasing integration in regional and High Speed Train networks, airports are becoming focal points of landside transport – multimodal interchange nodes. Their network position is a strategic advantage that makes airports new development poles, airports are transforming into centres of activity within them, especially due to increasing passenger traffic.

Drastic traffic growth has several downsides, notably airport delays and congestion are one of the most important threats to the future growth of civil aviation. Travel distances in Europe are such that more than 50 % of European flights are of less than 370 N.M.; a statistic heavily influenced by airlines' use of a hub and spoke operation. European airlines are operating fragmented networks with approximately 50% of the traffic being concentrated to only 10% of the airport pairs [1]. Both factors; a significantly large number of short-haul traffic in Europe and airlines' use of hub & spoke, are important drivers implying a potential mode shift. We assume that some high speed rail links may improve the capacity of airports for higher-earning long-distance flights, by providing a good alternative to short haul flights.

The layout of intermodal terminal can strongly influence the outcome of integrated intermodal services' efficiency, for this reason detailed analysis is needed allowing the airport to benefit from its intermodal premises and provide user friendly services to passengers. One of the ACARE 'challenges' is to ensure that passengers spend only 15 minutes in the airport terminal for short-haul flight and 30 minutes for long-haul flights. If this target is to be approached then we need quantum improvements in passenger flows in the terminal and in particular baggage check-in, passport control and security screening. An intermodal facility can offer enhancements to all of these processes and aim to look specifically at improved passenger flows in order to achieve assessments of the potential benefits.

Antonia Cokasova, antonia.cokasova@eurocontrol.int, is a PhD candidate at the EUROCONTROL Experimental Centre and EPHE Sorbonne, France and at the University of Zilina, Slovakia.

A. *Study objective*

The missing element in intermodal transport research is passenger perspective and willingness to switch between two transport modes. The objective of the study into 'Intermodality from the passenger perspective' is to analyze passenger perspective to air-rail intermodality and its impact on air transportation in general. The study will focus an analyzing passenger movement at an intermodal terminal. Despite the fact that some barriers of intermodality are highly visible, there are plenty of reasons to support the idea of system integration, such as high potential growth of air travel, congestion in air and on the ground, delays and their rising cost to operators. Important element of intermodal transport is the airport changing node. The station where people catch an airport train or high-speed train is where the ambience needs to start the transition from railway to airport to airline. In Europe there are very few examples of mode shift, hence there is no clear demonstration of the benefits derived from intermodal way of transport, resulting into growth of airport available resources or other benefits such as shorter minimum connection time for intermodal passengers. Only Frankfurt airport has experienced intermodal passenger movement while offering intermodal travel features such as curbside luggage check-in at the train station and minimum connection time as little as 45min. Due to the possibility of remote check-in at the railway station, intermodal passengers tend to behave in a different way as regular air passengers. Intermodal passengers can avoid long queues at the check-in counters, they are spared from carrying their own luggage and they can proceed straight to the security checks. These potential benefits might however imply drastic changes to passenger movement at the airport terminal

The impact of passenger movement tends to bear certain duality. Despite some potential benefits the intermodal passenger flow can possibly create additional bottlenecks at the security checks as passenger movement is not smoothened by check-in services. Movement of passengers at intermodal airport terminal is a rare phenomenon; the little information gathered is based on guess work and assumptions, as opposed to tangible results obtained by detailed analyses resulted from intermodal passenger movement simulation. Additional difficulty consists in the fact that human behaviour is often thought to be hard to define in models; simulations of processes with people involved are less common than industrial simulations [2].

The missing element in intermodal transport is passenger behaviour and its impact to intermodal transport. The analysis was conducted following an assumption, based on which passenger behaviour represents a decisive factor which will further determine the success of intermodal services. As little knowledge is available in the field of intermodal passenger movement it was necessary to gather a large amount of data from various passenger categories. Thorough research of intermodal passenger behaviour was conducted among passengers in order to obtain valid input information for further analyses on intermodal passenger movement in the airport terminal.

II. PASSENGER MOVEMENT SIMULATION

A. Simulation objective

The objective of the research was to simulate passenger flow at an intermodal air/rail terminal in order to evaluate the reduction or increase of demand on airport resources and to determine passenger movement in terms of bottlenecks and constraints in relation to passenger flows. Furthermore we examined the impact of intermodal passenger movement on minimum connection time, assuming that an intermodal transfer passenger should be able to spare significant amount of time.

The benefits identified for our study may be decomposed into several categories, as indicated below:

Reduction of demand on the airport resources considering passenger preferences – one of the main causes of delay at airports at the moment (in the absence of bad weather) is the saturation at the various choke-points (passport control, security screening and check-in). Intermodal way of transport can help to shift passengers from short haul flights to railways, improve airport access, offer available runway capacity for longer, more beneficial flights and among all, reduce the demand on airport resources inside the terminal or on the apron.

Shorter minimum connection time – the study will focus on the potential impact that an efficient layout may have on minimum connection time, assuming that a transfer passenger travelling by air/rail or rail/air instead of air/air should be able to spare significant amount of time due to shorter connection time that can be as little as 30 min

Passenger flow bottlenecks – intermodal way of transport may cause several bottlenecks at congested areas. This study will identify the impact of intermodal passenger flow on these critical airport resources.

Successful operation of a passenger building requires that travellers can flow through many different areas and processes in time and with enough space to meet expected levels of services. All too often, unfortunately, relatively insignificant aspects of the configuration or organization of elements of the airport building cause bottlenecks that jam a significant portion of the operation. In order to identify the hot spots, it is necessary to focus on the dynamics of the processes. An example of a hot spot is an underground station, where an overall design of platform for the train serving a major airport allows enough space for a trainload of people. If a detailed design placed a staircase somewhere in the middle of the platform the travellers descending on the platform naturally cluster at the platform level in front of the stairs and when the train comes in too many people try to board its cars next to the area near the staircase. Their perception is that there is not enough train capacity [2].

The analyses need to carefully consider how the flows of different streams of traffic might interact – for example the flows of arriving, departing, transfer and intermodal passengers. Particular attention is required on the peak instances when traffic surges over short periods. These may occur over much smaller intervals than the peak hours of traffic.

The analyses should also consider the psychological behaviour of individuals and crowds. People normally, for example, proceed to the first line of services they encounter, rather than turn to the side to look for servers who have shorter lines [2, 5].

B. Simulation tool and model description

1) Model description

For long time only very few analogous, low-level-of-detail models were available for passenger buildings primarily due to the complexity of the processes and interactions that take place in these buildings. Moreover, the few models that exist are very recent, and little practical experience has been obtained with their use. In airfield modelling, a few simulation models dominate the field internationally, whereas there are no dominant models of passenger building operations. Part of the reason is that passenger buildings and associated processes differ in many respects from one location to another. It is thus very difficult to develop a ‘standard’ simulation model that can be adapted to any given local conditions in a simple way. A

striking feature of models of passenger building operations is their requirement for enormous quantities of often difficult-to-obtain input data [2].
As air-rail intermodality is a well known phenomenon only at very few European airports, there is limited information available that describes passenger movement at intermodal terminals. A general unknown in the field of intermodality is passenger behaviour. In aim to analyze the impact of intermodal passenger movement on airport services' efficiency we propose a model reflecting realistic passenger and luggage movement thanks to various inputs from several sources. To avoid result bias caused by large number of inputs only five elements are taken into consideration, with possibility of further update in the future if needed. The model serves as a starting point for more complex intermodal simulations in the future. Each input will be further analysed in chapter C.

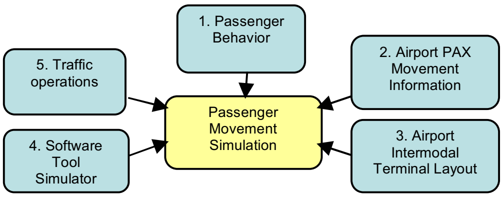


Fig. 1. Passenger simulation model input

The starting point for the simulation was a detailed drawing specifying the layout of the terminal building. All areas through which objects move had to be drawn and their function specified. In addition aircraft parking bays and points where passengers and baggage queue had to be identified. So the elements required were as follows:

- Layout of airport terminal and parking bays
- Schedule (Flight)
- Aircraft Information
- Physical airport features and passenger flows
- Earliness arrival distribution profile of departing passengers
- Baggage profile of passengers
- Counter schedule
- Dwell Area profiles and Schedule
- Passenger entrances and exits
- Passenger loads
- Check-in counter classes

The terminal consists of 4 airport levels; L2 corresponds to arrivals with 8 baggage carousels and a public arrival hall with dwell areas; L3 consists of public arrival area with 13 check-in counters for domestic movement and several dwells; L4 contains 128 check-in counters with passport control and 3 departure gates with security checks; and L5 is the departures level with dwell areas, security checks, gates and associated aircraft stands. The baggage detection processing is both at L4 and L5. The airport has 29 different passengers entry points with corresponding passenger arrival percentages. Entry points represent various doors from bus arrivals at L3, car parking at L2, elevators from underground parking at L1 which is not displayed in the layout and other entries. Approximately 70% of passengers arrive to the airport by car or taxi, to the front main entrance at L2.

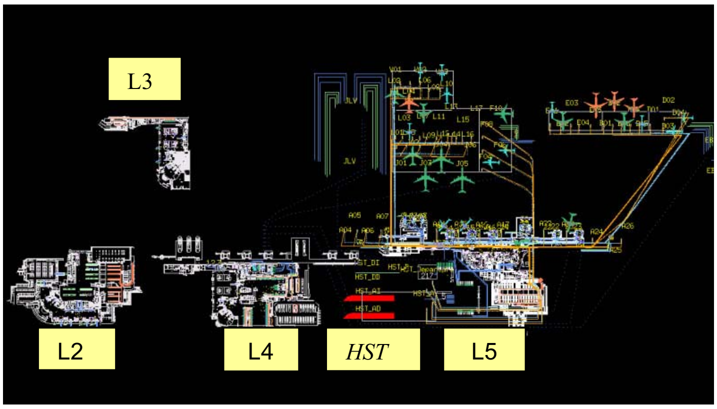


Fig. 2. Airport layout with 4 levels and railway station (A2 scenario)

The airport terminal, in second scenario, represents an international transport node (A2 scenario), connecting airline services to high-speed rail operations. Passengers can connect from rail to air transport, from air to rail transport, from air to air connection naturally and from rail to rail connection. Traffic sample analyses of transferred flights to railway services will be explained in chapter C.5.

The high-speed rail (HST) station with 4 platforms, in the second scenario, is situated to the left of the airport terminal. The railway station has separate platforms for arrival, departure, domestic and international passengers. The concept of passenger movement is similar for railway passengers as for airline passengers. All railway passengers have to go through check-in services, passport control, security checks and the check-in luggage has to be separated from both arriving and departing passengers. Passenger movement was designed as follows:

- separate flow for departing and arriving passengers
- separate flow for transit and terminating passengers
- all passengers are separated from their luggage at check-in services
- all passengers go through security check and passport control
- all passengers are able to visit dwell areas

Arriving passengers using platform HST_AI and HST_AD (international arrivals and domestic arrivals) are further separated to terminating and transit passengers. The arrivals gate at the railway station is connected to the airport terminal by people movers. There are 4 connections; a people mover for terminating passengers connected to the arrival area with baggage carousels by stairs leading to L2; a people mover for transit passengers connecting to gates at the right wing of the departures; a people mover for transit passengers connecting to gates at the left wing of the departures; and a people mover for transit passengers connecting to a departing train at gates HST_DI and HST_DD (international departures and domestic departures). The goal was to optimize the walking distance in the terminal.

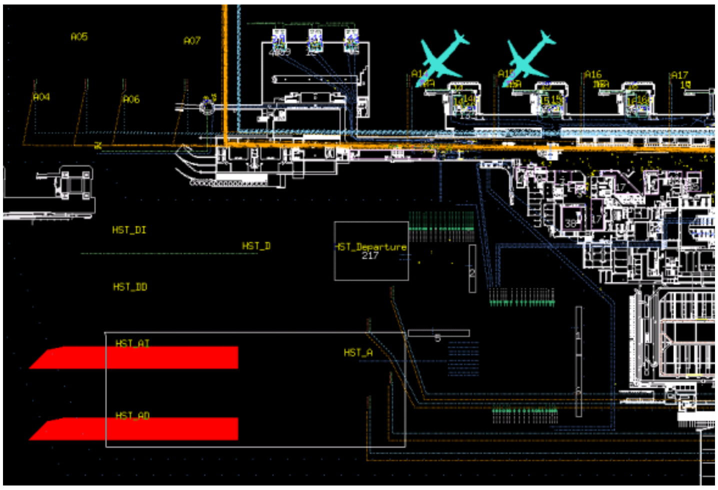


Fig. 3. High-speed rail station (A2 scenario)

2) Passenger movement in the terminal

Paths are directional open polygons used to represent passenger and baggage flows where it is very predictable, such as waiting in a customs queue [6]. The model contains over 820 paths, all with different speed, spacing and priority. The higher the priority the more likely the passenger will follow a certain path.

Grids are areas that allow for the free movement of passengers. The default destination of passengers in the grid is determined by searching all transfer areas that overlap or lie within the grid for exit routes that are valid for that particular passenger. To be valid the tag expression must be legal and any destination type and destination data values provided match the passengers [6]. The model contains 9 grid areas. The movement of passengers through the grid is determined by a probability function that is highly weighted to the direction in which the calculated default exit lies. If a grid edge, grid exclusion area or other passengers block the passenger's route to its destination for example, then those movement directions are deleted from the probability calculation. The passenger may choose to move in a direction around the obstacle and recalculate a valid option.

Before entering the departures area each passenger has to be checked-in and separated from his check-in baggage. This is ensured by 128 check-in counters at the airport level L3 and L4. Each counter has its corresponding baggage separation area and service time parameters with associated distribution.

Each passenger is obliged to pass through a detection processing system (DPS) at least once before entering the corresponding gate. The model contains 31 detection systems with either passenger or carry-on baggage detection. If a passenger happens to fail at the DPS check he is further checked at the hand-check-body or hand-check-baggage control stand. The probability of DPS failure varies from 10% to 28%. Higher for passengers than for carry-on bags; according to passenger movement analyses at our sample airport.

3) Arrival earliness and boarding profile

The arrival of departing passengers at the airport is determined by the values set in the Arrival Earliness profile [6]. Multiple profiles, each with a unique name, can be provided to cover the 24 hours of the day. We have spread the arrival percentage according to results obtained from passenger survey undertaken at Lisbon Airport, Roissy CDG in Paris, and on-board of Thalys and Eurostar trains. In the survey passengers were asked to indicate their estimated arrival earliness to the airport [3, 4]. The total of all percentages entered was equal to 100% and different profiles were used for domestic, international and intermodal flights.

Passenger boarding was controlled by the boarding profile which allowed setting up different passenger profiles. Boarding of international flights commenced 45 minutes prior to departure; in case of domestic flight it was 30 minutes prior to departure and 20 minutes for boarding a high-speed train. The model contains 27 gates and lounges for aircraft and 4 gates for high-speed rail. Passengers move towards these areas and wait until the boarding has commenced. The high-speed train gates are considered as platforms for arriving and departing, domestic and international trains.

Boarding time for transit passengers was calculated as follows: - If transit time was greater than 30 minutes, then if passenger numbers exceeded 200, boarding started 30 minutes prior to flight departure. If passenger numbers were less than 200, boarding commenced 20 minutes prior to departure time. - If transit time was less than 30 minutes passengers were boarding after 20% of the transit time has expired, e.g. if transit time was 25 minutes boarding has commenced 20 minutes prior to departure [6].

4) Baggage transfer

The model considers the impact that passenger baggage has on passenger movement through the airport. Baggage, be it check-in or carry-on, effects the ultimate efficiency of terminal operations. The time taken for baggage to be screened at various security points, such as major detection areas (for check-in baggage) and areas for the X-ray of hand-held bags can be modelled thanks to the simulation tool used in our project. Each piece of baggage (of either type) is uniquely identifiable with each passenger. There are two types of baggage considered for this functionality: carry-on and check-in to be associated with passengers, with support for multiple items of each type. The number of bags can be set via the Baggage profile and apply to either an individual or a party [6].

Baggage transfer has a significant role in intermodal transport representing one of the main differences between regular passenger and intermodal passenger movement. Based on passenger movement analyses undertaken at different European hub airports, passengers without luggage tend to move 10% faster than passengers carrying their own luggage.

C. Input elements

The following input elements were used in the passenger movement simulation

1. Passengers behavior
2. Airport passenger movement information
3. Airport intermodal terminal layout
4. Simulation tool
5. Traffic sample

1) Passenger behavior and requirements

Passenger shift to high-speed rail depends on the level of satisfaction, not necessarily providing better services than air transport but fulfilling certain needs that stand in priority to others. Individuals choose to travel by a mode of transport that offers a preferred bundle of levels of attributes which are important in making the choice between available alternative transport modes. In determining travel preference rules, individuals implicitly attach weights to a set of attributes that influence their choice, and make a choice based on the available set. To complete the set of items needed to derive a demand function a questionnaire was designed to identify the homogeneity of passengers; main passenger groups and major travel attributes that most passengers find crucial when deciding between air and rail transport.

After evaluating several methodologies we have decided to undertake a self-administered questionnaire. This method of data collection ensured a high response rate, accurate sampling and minimum of interview bias, while providing necessary explanations (but not the interpretation of questions) and giving the benefit of a degree of a personal contact.

The target sample consists of respondents open to undertake their journey both by high-speed train or airplane. To ensure correct sampling the questionnaire has been collected both at airport terminals and on-board of high-speed train. Regarding railway transport we have approached passengers travelling on Paris-Brussels-Amsterdam-Brussels-Paris line and Paris-London-Paris line, using Thalys International and Eurostar services. Questionnaires at the airport have been collected both in boarding gates and check-in areas. All together 913 valid samples have been collected, out selecting incomplete and incorrectly filled out samples.

Passenger travel preferences vary depending on nationality, purpose of travel and even age. About 60% of all passengers agree that travel time, ticket price and access to airport or station are the three main decision makers. 40% of all passengers consider on-board comfort, schedule & frequency, punctuality & reliability and walking/waiting time crucial when choosing between different transport modes, while on-board services and luggage handling have little or no influence on passengers. Some factors proved to be important only to a small group of people; however they are very influential; fear of flying or fear of crossing the Channel Tunnel (Eurostar) will certainly strongly influence passengers' inclination towards the competitive mode. The results of the analyses show strong relationship between passengers' gender, age, purpose of travel and their choice of transport mode. Among other questions related to general travel behaviour, passengers have been asked to indicate their arrival earliness to the airport and railway station. Several results concerning passenger travel behaviour have been used in the passenger movement simulation [3, 4].

2) Passenger movement information

There is very little information available concerning airport passenger movement; their speed, their personal security space, way they move on the escalators and people movers, movement of parties of people and many other examples. Most of the information is based on observations of target passengers moving along the terminal. It has been observed by Frankfurt Airport Expansion Program team that passengers without luggage move approximately 10% faster than passengers carrying their own luggage (source: Frankfurt Airport Expansion Program - undergraduate thesis). In general passengers move about 60 to 120 meters per minute [5]. In order to verify these data we have observed passenger movement at the sample airport for four days, with aim to gain better knowledge of passenger movement, in particular their speed, walking direction, spacing, preference for used doors and corridors, and dwell area occupancy time. The airport operation team at the sample airport has provided valuable feedback to our observations and data used for the simulation purposes.

3) Airport Intermodal Terminal

Out of six International European airports only one airport was ready to co-operate and provide us with additional support needed in later period of the study. The airport chosen for simulation is a main national hub; in 1998 it had the world's highest passenger growth rate. In 2003 the airport processed a total of 9, 63 million passengers and was served by 58 scheduled services airlines.

Airport layout in reality does not include a railway station; in reality the airport is served by regular bus shuttle. The layout of railway station is fictive and passenger movement at railway station has been designed based on conditions mentioned in chapter B.1. The aim was to propose the same conditions and advantages for railway passengers as airline passengers experience in the airport terminal. The railway station has both commercial facilities and sanitary areas.

For comparison purposes we have designed two airport layout scenarios:

Scenario A1: Original layout without high-speed rail station

Scenario A2: Original airport layout with high-speed rail station

In scenarios A2 arriving passengers have the advantage of remote check-in at their departure railway station before boarding the train. When arriving to the platform HST_AI or HST_AD (international arrivals or domestic arrivals) passengers are free to proceed straight to terminal without stopping at the check-in counter. The luggage is transferred by tugs and dollies from the railway station to the main baggage collection area for further distribution.

4) Simulation tool

The most suitable choice of tool was ensured by using a product developed by the Preston Aviation Solutions, called PaxSim. It is a fast-time simulation tool including passenger behaviour modelling and enabling a statistically significant number of runs to be performed for a particular set of input parameters representing all the various possible scenarios.

PaxSim is a graphics-based application used for the fast-time simulation of passenger and associated baggage flow at an airport terminal building. PaxSim can be used to simulate flows through discrete areas of a terminal or the entire terminal area. The analyst can view the actual flow of passengers, any potential bottlenecks and the ripple effects of passenger delays in other areas of the terminal.

5) Traffic sample

We have obtained valid traffic data from 29th of July 2005 delivered by airport operations team of the sample airport. The data contains detailed traffic information about the type of aircraft; each flight is allocated to a stand, gate and luggage carousel. The number of terminating and transit passengers was calculated based on passenger sample information obtained from the sample airport's operation team. The load factor varies from 70 to 85% depending on the destination, representing a busy summer day. The sample airport is located close to the coast line, therefore a large number of domestic flights could not be considered for potential mode shift. Airport's traffic distribution by segments is the following:

- International Movements: 82%
- Domestic Movements: 18%
- International Passengers: 78%
- Domestic Passengers: 22%

We will consider three traffic samples:

Sample B1: original traffic sample – 207 arrivals with 25345 arriving passengers on-board and 205 departures with 25295 departing passengers on-board.

Sample B2: intermodal traffic sample for domestic flights – Continental domestic short-haul flights are replaced by high-speed rail services. It corresponds to 7% of arriving passengers and to 8% of departing passengers.

Sample B3: intermodal traffic for domestic and international flights – Continental domestic and short-haul international flights are replaced by high-speed rail services. It corresponds to 18% of arriving passengers and 19% of departing passengers.

In sample B2 and B3 we have decided to merge several flights into train arrivals and departures based on the following criteria:

- flights arriving and departing from and to the same destination airport
- short-haul flights
- continental flights without natural boundaries
- frequent flights to ensure large number of passengers feeding train carriages
- flights with no more than 90 min difference between their scheduled time of arrival or departure

In scenario B2 only two domestic destinations have fulfilled all criteria; however the frequency of these two flights ensured 7 to 8% of passenger shift to high-speed rail. In scenario B3 five domestic and international flights correspond to the criteria, resulting into 24 arriving high-speed train services and 28 departing high-speed train services. The number of passengers has remained the same in all scenarios; no modification has been applied to the types of passengers or their destination.

D. Simulation scenarios

There are 3 scenarios; composed of 3 traffic samples and airport layout with 2 different configurations. The main difference between the scenarios is the connection to a railway station at airport terminal and the flow of intermodal passengers. Starting from the original airport layout and traffic flow without railway station, to a transition where an airport is connected to a high-speed rail station with remote check-in facilities and 19% intermodal flow.

Base scenario – A1B1

Sample airport as it is known today, without a direct connection to high - speed rail station and with 26% of traffic less than 800 km's.

Air/Rail Domestic – A2B2

Sample airport with high-speed train connection, serving two domestic destinations with 7 to 8% intermodal passenger flow

Air/Rail International– A2B3

Sample airport with high-speed train connection, serving two domestic and three international destinations with 18 to 19% intermodal passenger flow

III. ANALYSES AND CONCLUSION

Similar to most international hub airports, our sample airport registered its peak traffic at 6H45, 8H35 and 11H30. The maximum number of passengers in the terminal was at 11H30 with 4342 arriving and departing passengers for base scenario and 4187 arriving and departing passengers for air/rail international scenario (A2B3), which is 3.6% less passengers in case of air/rail scenario.

In base scenario 16 flights had their boarding delayed. Due to ongoing disembarkation their assigned gates were occupied by previous arriving flights with remaining passengers on-board. Lack of stand availability has resulted into 121 passengers and luggage missing their connection flights. In air/rail international scenario only 5 flights announced delayed boarding, resulting into 39 passengers missing their flights or train connection.

Intermodal passenger movement has improved the stand occupancy time. The average stand occupancy time for each stand was 7.6 min shorter in intermodal scenario. While average occupancy time of the stand per flight was 64.6 min, it has been reduced to 57 min with high-speed train correspondence. We have achieved 11.8% decrease in stand occupancy time per flight which has resulted into fewer passengers missing their flights.

The largest difference between the base scenario and air/rail international scenario (A2B3) concerned arrival passenger movement. In base scenario arriving passengers are transferred to airport arrival gate for passport control while terminating intermodal passengers can proceed straight to the luggage carousel which shortens their time spent in the airport terminal. The passport control for railway passengers is processed at the railway station, which significantly increased arrival passenger services and therefore decreased the number of passengers in the terminal.

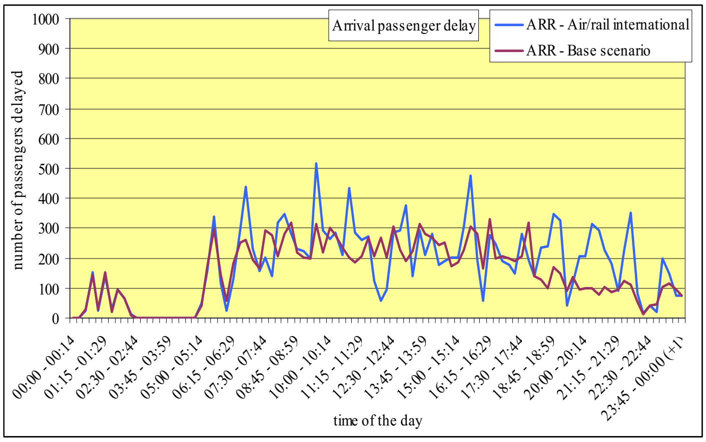


Fig. 4. Arrival passenger delay (comparison between A1B1 and A2B3)

Arrival intermodal passenger movement however registered higher passenger delay peaks. Each train arrival with significantly larger number of passengers than it is in case of aircraft arrival represents a higher workload for passport control and baggage handling facilities. As only short haul flights were replaced by high-speed rail, operated by smaller airliners, train arrival with the number of passengers representing several short-haul flights was followed by large number of passengers being released into the terminal at the same time, resulting into more queues for shorter period of time, as demonstrated in the chart above.

The choice of the simulation tool proved to be very successful. Thanks to its user friendly interface, sophisticated simulation editor and flexibility, PaxSim enabled the design of intermodal connection with transport features corresponding to intermodal air-rail terminal.

Intermodal connection has demonstrated possible improvements in passenger flow at the airport terminal. More results are required before understanding the impact of intermodal air-rail transport on the airport operations. Passenger throughput, service utilization and queue length analyses are on the way for all three simulation scenarios. The relationship between the numbers of intermodal passengers and its impact on airport resources will be further measured by regression analyses.

Based on preliminary results obtained from the simulation scenarios, transport shift in the future will be suitable only for airports with a large number of short-haul and frequent traffic. The possibility of a potential shift will be strongly influenced by the flight schedule; the more short-haul traffic arriving or departing to the same destination in short space of time, the more successful intermodal shift will be. Train frequency and the choice of destination will be the strongest asset of intermodal travel.

Furthermore the success of intermodality and the possibility of easing congestion will depend not only on airports efficiency but also on passengers' willingness to experience new way of travelling.

ACKNOWLEDGMENT

Many thanks to Mr. Vu Duong, Mr. Louis SILLARD (EUROCONTROL Experimental Centre) and the Preston Aviation Solutions for their valuable comments, operational help and support

REFERENCES

[1] EUROCONTROL Network, Capacity and Demand, Experimental Centre, Bretigny sur Orge

[2] Neufville, R and Odoni, A (2003), Airport Systems Planning Design and Management

[3] Cokasova, A. (2004), Passengers' Choice Between High-Speed Train and Air Transport, ICRAT2004, Zilina (SK), November 2004

[4] Cokasova, A. (2005), Analyses of Passenger Preferences in Air-Rail Intermodal Transport, COST340, Paris (FR), 13-15June

[5] Fruin, J.J. (1971), Pedestrian Planning and Design

[6] PaxSim Solutions User Manual, Preston Aviation Solutions 2006, Melbourne, Australia

[7] Givoni, M. (2001). Airline and railway co-operation, a new approach to intermodality – a research perspective. April 2001, University College of London, London

[8] Moore, P. (2001). High speed rail as a solution to airport congestion. Sierra Club – Loma Prieta Transportation Committee, California

[9] Priebe, M. (2002). High speed rail and air transport: Intermodality – the moment of truth (Presentation paper), ATAG, October 2002, Eurailspeed Conference, Madrid

A. Cokasova Is in her final year of PhD studies at EUROCONTROL Experimental Centre in Bretigny sur Orge, France. She has an engineering degree in Civil Aviation, majoring in airport operations and intermodal transport.

Antonia has spent several years acquiring knowledge and experience in the field of civil aviation. She undertook several internships at the EUROCONTROL Headquarters in Brussels, in the Experimental Centre and at the Slovak Airport Authorities. She was an active member of the Rail Air Facilitation Forum organized by the European Commission.

Operational Concept for a complete A-SMGCS

J. Jakobi

Abstract - **The basic idea of this paper is to identify a comprehensive A-SMGCS (Advanced Surface Movement Guidance and Control System) concept that allows incorporating the full scope of surveillance, guidance, routing and control services as well as new onboard related A-SMGCS services including their air-ground interoperability. This concept approach supports the stepwise implementation of a complete A-SMGCS at a specific airport with its specific needs. These results are an output of the European project *EMMA* (European Airport Movement Management by A-SMGCS), an Integrated Project launched by the European Commission in its sixth framework programme. The paper identifies all A-SMGCS services, their evolutionary implementation steps, and gives recommendation to the composition of *implementation packages (IP)* to meet the airports' specific needs.**

Index Terms - **A-SMGCS, operational Concept, EMMA**

I. Introduction

THE basic idea of this paper is to identify a comprehensive A-SMGCS (Advanced Surface Movement Guidance and Control System) concept that allows to incorporate the full scope of surveillance, guidance, routing and control services as well as new onboard related A-SMGCS services including their air-ground interoperability. This concept approach supports the stepwise implementation of a complete A-SMGCS and shall give recommendations for A-SMGCS "implementation packages" that are tailored to the user's needs. These results are an output of several workshops with A-SMGCS users, industry- and R&D organisations who participate in the European project *EMMA* (European Airport Movement Management by A-SMGCS), an Integrated Project launched by the European Commission in its sixth framework programme.

The finally introduced *implementation packages* go beyond the already existing EUROCONTROL, EUROCAE, and ICAO A-SMGCS *implementation level* definitions because they also consider <u>equipment</u> and <u>procedures</u> of a specific service. The new term *packages* has been chosen to distinguish from the term '*implementation level',* which proved as insufficient in meeting the needs of stakeholders during a stepwise implementation of a full scope A-SMGCS. EUROCONTROL's and EUROCAE's definition with its four implementation levels focuses on the main four A-SMGCS functions: *surveillance, control, guidance,* and *routing*, which works fine with *surveillance* and *control* because these two functions depend on each other logically in a successive way and do not consider the onboard part. Moreover, these two services are mainly used to assist the users, thus procedures do not have to be changed fundamentally, and interoperability to other services is not a critical issue. The implementation of automated routing or guidance services, however, would increase the complexity of the A-SMGCS system and their operational use still lacks maturity. The concept has to give careful consideration on changing operational procedures, shifting responsibilities from human to equipment (e.g. visual reference vs. electronic display), introducing onboard automated services, and equipments, as well as on the interrelations between the A-SMGCS functions.

The EUROCONTROL and EUROCAE level III and level IV concepts do not help here anymore as these concepts do not adequately address the full scope of ASMGC-S operational use. ICAO goes a step further and considers the responsibility shift between controllers, pilots, and equipment for all A-SMGCS services but does not give sufficient information to procedures with which the system is used nor does it describe what technical enablers would be needed and what service performance level the users can expect.

II. EMMA Implementation Steps for A-SMGCS Services

The EMMA operational concept approach commenced to extend EUROCONTROL's level I&II concept [3]-[4] by a detailed description of all A-SMGCS services that include guidance, routing, planning, and onboard services, as well an extensions of surveillance and control services. This has been done for each of the three main users of an A-SMGCS: Air traffic controllers (ATCOs), flight crews, and vehicle drivers. The second step was to look for appropriate technical enablers that are needed to bring the service to live. The third step was to derive successive implementation steps for each A-SMGCS service to provide recommendations for a stepwise implementation. To achieve this, the services and technical enablers have been assessed with respect to their:

- Development status of the technical enabler (e.g. standardised, on the market or to be developed yet)
- Development status of the operational service (e.g. already validated by operational life trials or under investigation through simulation or only at the stage of a concept)
- Degree of interrelations to other functions (in terms of its complexity)
- Quality of the technical enabler (needed reliability)
- Impact on current operational procedures and size of the changes
- Cost/benefit considerations

J. Jakobi is with the Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), member of the Helmholtz-Community, Institute of Flight Guidance, 38108 Braunschweig, Germany (e-mail: joern.jakobi@dlr.de).

The following tables show the implementation steps for each A-SMGCS service that have been identified by the EMMA partners[1] (cf. EMMA OSED document, §2, [1]).

ATCO Service – SURVEILLANCE

Service Steps	Characteristics	Comments
S1	• Detection and accurate position of all aircraft in the manoeuvring area • Detection and accurate position of all vehicles in the manoeuvring area • Detection and accurate position of obstacles in manoeuvring area • Identification of all cooperative aircraft in manoeuvring area • Identification of all cooperative vehicles in manoeuvring area	All movements on the manoeuvring area have to be authorised by aerodrome controller (§7.5.3.2.1 [8]). With EMMA all authorised movements shall be properly equipped to enable automatic identification. All other movements are intruders or obstacles. There might be authorised aircraft on the manoeuvring area that are not properly equipped to be identified automatically (e.g. in case of transponder failure). Procedures to cover such exceptional cases have to be derived.
S2	• Step1 + • Detection and identification of all aircrafts in <u>movement</u> area	There might be authorised aircraft (e.g. VFR) on the movement area that are not properly equipped to be identified automatically. Those aircraft are an exception but procedures have to be derived to cope with those aircraft too.
S3	• Step2 + • Detection and identification of all vehicles in movement area (where manoeuvring aircraft may come into conflict with each other or with vehicles) • Detection of Obstacles in movement area	ICAO [7] §3.5.16.3 "It is not practicable to exercise total control over all traffic on the movement area. On the apron, an A-SMGCS applies only to those areas where manoeuvring aircraft may come into conflict with each other or with vehicles. Therefore, one requirement is to restrict the movement of vehicles on the apron to designated areas and routes. It is also necessary to keep service vehicles away from an active stand. This can be achieved by having painted lines that outline the areas to be left clear when a stand is active. Alternate means of protecting an active stand might become available as a result of technology." Those restrictions to apron areas where manoeuvring aircraft may come into conflict with each other or with vehicles are particularly needed in low visibility conditions, when movements are not able to avoid each other. Movements, which uses those apron areas, should be co-operative to get identified automatically on the ATCO surveillance display and should be equipped with an onboard display showing the own position and position of other aircraft to avoid conflicts.

ATCO Service – CONTROL

Service Steps	Characteristics	Comments
C1	• <u>Runway</u> Conflict/Incursion detection and alerting of: 1. Aircraft arriving to, or departing on, a closed runway; 2. Arriving or departing aircraft with traffic on the runway (including aircraft beyond the runway-holding positions); 3. Arriving or departing aircraft with moving traffic to or on a converging or intersecting runway; 4. Arriving or departing aircraft with opposite direction arrival to the runway; 5. Arriving or departing aircraft with traffic crossing the runway; 6. Arriving or departing aircraft with taxiing traffic approaching the runway (predicted to cross the runway-holding position); 7. Arriving aircraft with traffic in sensitive area (when protected); 8. Aircraft exiting the runway at unintended or non-approved locations 9. Unauthorized traffic approaching the runway; 10. Unidentified traffic approaching the runway 11. Aircraft on a closed taxiway[2]; 12. Aircraft taxiing with excessive speed; 13. Crossing of a lit stop bar;	The conflicts in this step address mainly runway conflicts which are the most critical in safety terms. Conflicts are initially defined by ICAO ([7], §3.4.5.7).
C2	• <u>Taxiway</u> Conflict/Incursion detection and alerting of: 1. Arriving[3] aircraft exiting runway at high speed with converging taxiway traffic; 2. Aircraft approaching stationary traffic; 3. Aircraft overtaking same direction traffic; 4. Aircraft with opposite direction traffic; 5. Aircraft approaching taxiway intersections with converging traffic; 6. Aircraft exiting the taxiway at unintended or non-approved locations; 7. Unauthorized traffic on the taxiways, 8. Unidentified traffic on the taxiways	The conflicts in step 2 address mainly taxiway conflict. Conflicts are initially defined by ICAO ([7], §3.4.5.7).
C3	• Detection of <u>route/plan deviations</u> • Support to Ground Clearances and ATCO coordination	The Support to Ground Clearances and ATCO handover coordination involves Controller Pilot data link Communication (CPDLC) and Electronic Flight Stripes (EFS)

[1] Beside industry and R&D representatives there were representatives from ANSPs (ANS_CR, ENAV, DSNA, AENA, and DFS) and Airlines (CSA and DLH).

[2] Step 1 conflicts 11), 12), and 13) are taxiway conflicts, but because their less complexity EMMA decided to implement them with step 1 already.

[3] ICAO doc. 9830 considers this as a runway conflict

Service Steps	Characteristics	Comments
C4	• Conflict/Incursion detection and alerting of apron/stand/gate conflicts 1. Aircraft movement with conflicting traffic; 2. Aircraft movement with conflicting stationary objects; 3. Aircraft entering/exiting the apron / stand / gate area at unintended or non-approved locations; 4. Unidentified traffic in the apron / stand / gate area.	The conflicts in step 4 address mainly apron/stand/gate conflicts which are the most challenging conflicts in terms of their complexity. Conflicts are initially defined by ICAO ([7], §3.4.5.7). Only the movements in the apron which could be threats to aircrafts movements shall be covered.

Note: The provision of automatic *conflict resolution advisories* may be initiated at any step of the control service depending on the complexity of the resolution possibilities.

ATCO Service – ROUTING

Service Steps	Characteristics	Comments
R1	• Manual Routing	Manual input of a route supported by the computation of the shortest taxi route w.r.t. to local standard routes
R2	• Semi-automatic Routing	Routing service proposes a most suitable route, taking into account control and flight plan information.
R3	• Automatic Routing	Routing service provides route (track) and time information by aid of a planning function.
R4	• Automatic Routing + Optimisation of runway resource[4]	Planning support is further increased by a departure manager providing optimal departure times and sequences. Procedures are supposed to be changed.

ATCO Service – Guidance

Service Steps	Characteristics	Comments
G1	• Manual Operation of Ground based Guidance Means	Equipment available on the market.
G2	• Automatic Operation of Ground based Guidance Means	Automatic generation of guidance information, based on the cleared route and the actual position of the aircraft.

Note: Ground based guidance gives direct visual information to the pilot by viewing outside the cockpit windows, and thus is independent on on-board services. Ground guidance means are not a real service to the ATCO but are operated by her/him – that is the reason why it was grouped under *ATCO services*.

Flight Crew Service – Aircraft

Service Steps	Characteristics	Comments
A1	• Airport Moving Map • Surface Movement Alerting (initial, incl. proximity of obstacles – runway) • Braking and Steering Cue (for landing roll)	Own ship position on a moving airport map including a monitoring and alert service referring fixed obstacles or proximity alerts.
A2	• Ground Traffic Display • CPDLC Ground Clearance and Taxi Route Uplink • Ground-Air Database Upload • Surface Movement Alerting (taxi route deviation) • Traffic Conflict Detection • Braking and Steering Cue (landing roll and taxi)	Own ship position including the indication of the surrounding traffic and a monitoring service to detect surface or traffic conflicts. Airport map upload service to update recent layout changes. Communication is supported by a point to point data link. Taxi routes can be transmitted and represented graphically on the AMM display. Braking and steering cues are extended to taxiways.
A3	• HUD Surface Guidance	A Head Up Display provides the pilot flying with a scene linked outside view, means; her/his outside view is supported by scene linked navigation cues that are presented on a transparent screen in front of her/his cockpit front window.
A4	• Automated Steering	This service can be compared to an auto pilot for taxi operations. The nose wheel is automatically kept on the centre lines, whereas the throttles will be further under control of the pilot flying.

Vehicle Driver Service – Vehicle

Service Steps	Characteristics	Comments
V1	• Airport Moving Map • Surface Movement Alerting (vehicle alone)	Own vehicle position on a moving airport map including a monitoring and alert service referring fixed obstacles infringements into restrictions areas.
V2	• Surface Movement Alerting (complete) • Ground Traffic Display Traffic Conflict Detection	Own ship position including the indication of the surrounding traffic and a monitoring service to detect surface or traffic conflicts.
V3	• Dispatch and Guidance by data link	Communication is supported by a data link. Information or Taxi routes can be transmitted and represented graphically on the AMM display.

III. FROM SERVICE STEPS TO IMPLEMENTATION PACKAGES

Having defined implementation steps for each A-SMGCS service the next step was to give recommendations to the stakeholders how to group them to *implementation packages* (IPs). Not all combinations, however, are feasible. Therefore, the services and their technical enablers cannot be considered in isolation. The services interact and thus depend on each other. For instance, there is no sense in implementing a *route deviation monitoring* service when the taxi route is not known to the *conflict monitoring* service, thus, a routing function has to be implemented first.

Table I attempts to depict the arrangement of implementation steps for each A-SMGCS service in a logical order. The services are arranged to the main users and are identical to the introduced service steps above.

[4] Planning services, like a DMAN, may also be implemented without a pre-existing "routing" function.

TABLE I
LOGICAL INTERDEPENDENCIES BETWEEN EMMA SERVICE STEPS

		Expected Steps to each Service				
ATCO	Surveil.	S1 id/pos everything on RWYs	S2 S1 + id/pos a/c in the complete movement area	S3 S2 + id/pos vehicles movement area		
ATCO	Control	C1 Conflict RWY	C2 Conflict TWY	C3 Conflict Route Deviation / CPDLC / EFS	C4 Conflict Apron	
ATCO	Guidance	G1 Manual switched ground guidance	G2 Auto switch			
ATCO	Routing	R1 Manual	R2 Semi-auto	R3 Auto (planning)	R4 ROP	
Flight Crew	Aircraft	A1 AMM	A2 Ground traffic / CPDLC	A3 HUD	A4 Auto steering	
Vehicle Driver	Vehicle	V1 AMM	V2 Ground Traffic	V3 CPDLC		
Timeline		2006 → (t)				

id	Identification
pos	Position
veh	Vehicle
ROP	Runway Occupancy Planning
AMM	Airport Moving Map
HUD	Head-Up Display
S1	Surveillance Service for ATCOs step 1
C1	Control Service for ATCOs step 1
G1	Ground guidance means Service for ATCOs step 1
R1	Routing Service for ATCOs step 1
A1	Onboard Service for Flight Crews step 1
V1	Onboard Service for Vehicle Drivers step 1

Clustering A-SMGCS service steps to implementation packages must take into account the airport specific characteristics. The ICAO Manual for A-SMGCS proposes three main criteria to classify aerodromes [1]: 1) complexity of the airport layout (basic, simple, complex), 2) the visibility conditions (VIS1 through VIS4), and 3) the average traffic density (low, medium, heavy). A full combination provides 3x4x3 = 36 aerodrome types, which is a large number and thus of little practical use.

To make this huge number of combinations more tangible, EMMA project partners decided to focus on *complex* airports with either *medium* or *heavy* traffic density and four categories of the prevailing visibility conditiontable II), which was deemed much more practical.

TABLE II
IMPLEMENTATION PACKAGE RAW-MATRIX

Layout	Traffic density	Visibility			
		Vis 1	Vis 2	Vis 3	Vis 4
COMPLEX	Medium	Implementation Package (IP) 1	IP2	IP3	IP4
COMPLEX	Heavy	IP5	IP6	IP7	IP8

Each of these eight cells provides a different airport environment and different requirements to fulfil the users' needs. But what are the general users' needs? In principle all potential A-SMGCS users aim to operate safer and more efficient on the ground. This is also reflected in the A-SMGCS definition: *A system providing routing, guidance and surveillance for the control of aircraft and vehicles in order to maintain the declared surface movement rate under all weather conditions within the aerodrome visibility operational level (AVOL) while maintaining the required level of safety* [1]. That means, the two main questions when defining implementation packages at each cell must be: *Which A-SMGCS service is needed to enable safe and efficient ground operations?*

But before this question can be answered satisfactorily, well designed operating procedures have to be developed to enable a service to provide its operational benefit. For instance, when ATCOs are assisted by an A-SMGCS *surveillance and monitoring/alerting* services but they are not allowed to use such services as a primary source of information through relevant operational procedures, safety and efficiency aspects would not fully exploited.

The same applies to planning or onboard guidance: Introduction of these services apart from the right procedures to influence the ground traffic would not lead to throughput benefits. The equipment, on the other hand, is more seen as a catalyst or as a prerequisite to evoke a potential service. But procedures are always the core to enable a service that meets the users' needs.

Initial procedures developed in EMMA are outlined in the D135 EMMA Operational Requirement Document [2]. These initial procedures for higher A-SMGCS services still lack sufficient maturity but were used as a starting point to form EMMA implementation packages. In successor projects (e.g. EMMA2) these procedures will be tested in simulation exercises and updated by the gained results.

IV. IMPLEMENTATION PACKAGES

The following sections describe each of the eight IPs and explain the reasons why certain services are recommended to be implemented. In addition to that, *options* for additional service implementations are proposed for each IP to further increase safety and efficiency.

A. Implementation Package 1

Concerning table II this IP is recommended for a complex airport (more than one RWY) with medium traffic density (20 < movements/h < 35) operating under visibility 1 (VIS1). That is, the ATCO can control the traffic by visual reference over the complete aerodrome, at all times, and the traffic is not as heavy as the ATCOs or pilots reach their mental capacity limits.
A-SMGCS could help here to provide surveillance (position and identification = S1) of aircraft and vehicles on the airport manoeuvring area to enhance ATCO's situation awareness, to complete ATCOs situation assessment (e.g. who is who in a taxi queue far away from the control tower or to allow them to go without pilot position reports). Furthermore, a runway safety net (C1) helps to prevent runway incursions. All this would contribute to safety and efficiency.

Optional: Since an Airport Moving Map (AMM) with its basic service (showing the position on the ground) is independent on ground equipage, Airlines and Airports could equip their aircraft and vehicle fleets with an AMM (A1 + V1) to increase the pilot's and driver's situation awareness what would increase safety again. Automatic routing or ROP (R3/R4[5]) could be initiated when the surface movements are identified as too inefficient compared to runway or gate capacity. When the route is known to the system it can be transferred onboard via data link provided that aircraft support a CPDLC service (A2) and provided that an input device for the ATCO and proper procedures are available.

B. Implementation Package 2[6]

The side conditions with IP2 are the same as with IP1 except that VIS2 is predominating now, i.e. the ATCO cannot see the traffic outside. Therefore, the ATCO[7] should be provided with a surveillance that covers the complete movement area with position and identification information of all aircraft (S2). Since pilots and vehicle drivers can still see and avoid each other (visibility 2), conflict alerting service is concentrated on the runways (C1) where providing separation is the most difficult and safety-critical part.

Optional: A Ground Traffic display showing the surrounding traffic by receiving TIS-B information (A2 + V2) would be an option to increase safety. As surveillance covers the whole airport also conflict alerting extended to the taxiway could be implemented (C2). However, detection of conflicts on taxiways by automation is a very complex task because much information has to be known to the control function, e.g. the cleared taxi route. That is why, it is assumed that *see and be seen* seems to be rather efficient with VIS2. When the route is known to A-SMGCS, CPDLC (A2) can be implemented as well to increase safety and efficiency.

[5] Manual (R1) and semi-automatic (R2) routing are identified as implementation steps and are certainly needed to evolve automatic routing (R3) at an airport. However, working with R1 and R2 are rated as too inefficient to support the ATCOs – therefore these first implementation steps are not considered with A-SMGCS implementation packages.

[6] IP2 nearly complies with EUROCONTROL implementation levels I&II.

[7] At some airports there is a separate Apron or Ramp Control that are not ANSP. However, within this context the ATCO term is also used for non-ANSP control units.

C. Implementation Package 3

Visibility decreases further so that pilots are not able to see and avoid each other anymore (visibility 3). Conditional taxi clearances (e.g. "follow CSA123") that base on the pilots ability to see and avoid the other traffic cannot be applied anymore. Currently ATCOs take into account these new limitations and give taxi instruction following procedural control operations (one aircraft only within a segmented area). Those procedures for low visibility operations (LVO) (PANS ATM, doc4444, §7.10, [8]) maintain safety (as the topmost objective) but on the expense of throughput.

A solution to maintain safety <u>and</u> throughput would be that aircraft are still able to see and avoid each other by providing them a step 2 onboard service (A2). A2 enables them to see the surrounding traffic by receiving surveillance information from ground stations (Ground traffic display enabled by ADS-B-in and TIS-B). The main issue with this solution is the transition phase: It would be needed that all movements are equipped with this service because non-equipped movements cannot avoid other ones and they cannot be controlled in a mixed mode environment. Even when A2 would be the best solution it cannot be assumed that all aircraft are equipped from one day to the next. That is why this solution cannot be preferred for the near future.

A first interim solution would be to assist the <u>ATCO</u> to provide the control service for all movements on the movement area. The runway safety net (C1) would be extended to the taxiways and aprons (C2 + C4) including a route deviation alerting service (C3). To make the route known to the system, automatic routing should be available (R3). Surveillance would be extended to step 3 (S3), what would enable the ATCO to provide traffic information to aircraft and vehicles on the apron area as well. But as control of the whole apron area would be hard to achieve on the basis of its surveillance display, the ATCO would only be responsible for designated areas of the apron area (taxi lanes, active stands, passive stands). Only authorised movements (vehicles must be equipped and must ask for permission to enter) would be permitted to use such areas. Other movements would be restricted to parts of those areas (ICAO doc 9830, §3.5.16.3 [1]).

A second and perhaps more likely interim solution would be to equip vehicles, which have to move on these designated areas, with a ground traffic display (V2). This would enable them to avoid conflicts with moving aircraft and they could move uncontrolled - without regard of the ATCO. Which solution will eventually be applied is highly dependent on the airport layout, equipment reliability and standards, local procedures, and decisions met by the local stakeholders.

Optional: Since S2 is available (includes TIS-B) a ground traffic display (A2) could be used by the airlines to increase situation awareness and efficiency of taxi movements. Routing can be extended to a *runway occupancy planning* (ROP) (R4) when cost/benefit data support this implementation.

D. Implementation Package 4

Visibility is now insufficient to taxi by visual reference. Onboard service has to be extended to step 3 (A3) that includes a head-up display (HUD) that enhances the pilot's local situation awareness by a HUD that provides scene-linked and conformal symbology to enhance the outside view. Step 2 surveillance (S2) and step 2 control (C2) assist the ATCOs and provide them the required situation awareness. Vehicles are equipped with ground traffic displays (V2) whereby they can move without additional traffic information from ATCO.

Optional:
Service to flight crews can be extended by an auto steering function (A4), which keeps the aircraft's nose wheel on the yellow taxi line automatically. Additionally, alerting can be extended to the apron area (C4) and automatic routing (R3) and ROP (R4) can be implemented if shortages with safety or efficiency are found.

E. Implementation Package 5 through 8

IP5 through IP8 are designed for the operational needs of complex airports with heavy traffic density, greater than 35 movements per hour. Since the traffic density is very high and thus the human operators often reach their capacity limits, surveillance should always be step 2 (S2) and control should always be step 3 (C3). These service steps would provide the ATCO with a complete surveillance and a safety net of the overall movement area. This increases mainly safety. To increase or maintain throughput automatic routing including a ROP (R4) should be implemented to support the users by optimised and negotiated times and taxi routes (on a CDM basis).

With VIS3 (IP7) it is insufficient for pilots to avoid collisions with other traffic by visual reference. As mentioned above with IP3 the ATCO should be provided with an additional safety net that detects conflicts not only on the runways and on the taxiways but also on the apron areas (C4). Vehicles moving on the designated apron areas (where they can conflict with aircraft) should be equipped with a ground traffic display (V2) to enable the vehicle drivers to see the surrounding traffic and to avoid it.

With VIS4 (IP8) it is insufficient for pilots to taxi by visual guidance only. As with IP4, the onboard service has to be extended to step 3 (A3) that includes a head-up display (HUD) with enhanced symbology to improve the pilot's local situation awareness.

Optional:
Optional but very beneficial with all IPs with heavy traffic would be the step 2 onboard service to flight crew and vehicle drivers (A2 and V2). With this service pilots and vehicle drivers are always able to see where they are, where they have to go, and where the surrounding traffic is. Particularly with dense traffic, this would contribute to safety, but also to faster taxiing what is an efficiency aspect. Vehicles can be equipped further on with vehicles service step 3 (V3) what would allow them to receive a taxi route, or the exact location of an accident, or other information via data link. This would be particularly beneficial with VIS3 and VIS4 when they cannot see the destination just by looking outside their windows. Table III gives an overview of all eight implementation packages:

TABLE III
RECOMMENDED IMPLEMENTATION PACKAGES

Layout	Traffic density		Visibility			
			Vis 1	Vis 2	Vis 3	Vis 4
COMPLEX	Medium	Recommended	IP1 S1+C1	IP2 S2+C1	IP3 S2+C3/4 +V2+R3	IP4 S2+C2+ A3 + V2
		optional	A1+V1 R3/R4+A2+ V1	A2+V2 C2+R3/R4+A2 +V1	R4+A2	C4+ A4+R3/R4
	Heavy	Recommended	IP5 S2+C3+ R4	IP6 S2+C3+ R4	IP7 S2+C4+ V2+R4	IP8 S2+C3+ A3+V2+ R4
		optional	A2 + V2	A2 + V2	A2 + V3	A4 + V3

V. CONCLUSION

The present paper outlines an improved concept to support future A-SMGCS research and implementations. The concept bases on gained results and experiences from 15 years of A-SMGCS research, e.g. BETA project results ([9], [10], and [11]). This paper does not aim to provide results or even a CBA. It is more focused on giving a theoretical basis for future A-SMGCS research in order to derive reasonable hypotheses and significant results.

ACKNOWLEDGEMENT

The project EMMA was funded by the European Commission DG-TREN. I would like to thank the DG-TREN for their support. Thanks as well to the project partners of this project.

REFERENCES

[1] D1.3.1u EMMA Air-Ground Operational Service and Environmental Description www.dlr.de/emma/
[2] D1.3.5u EMMA Operational Requirement Definition, www.dlr.de/emma/
[3] EUROCONTROL, Operational Concept & Requirements for A-SMGCS Implementation Level I, 2003
[4] EUROCONTROL, Operational Concept & Requirements for A-SMGCS Implementation Level II, 2004
[5] ICAO, Advanced Surface Movement Guidance and Control Systems (A-SMGCS) Manual, Doc 9830 AN/452, First Edition 2004.
[6] ICAO, Manual of All Weather Operations, doc 9365, 2nd edition 1991
[7] ICAO, Manual of Surface Movement Control and Guidance Systems (SMGCS), Doc 9476-AN/927 First Edition 1986
[8] ICAO Procedures for Air Navigations Services. Air Traffic Management, Doc 4444 ATM/501. Fourteenth Edition 2001
[9] Jakobi, J.; Lorenz, B.; Biella, M. (2004): Evaluation of an Onboard Taxi Guidance System. In: Dennis A. Vincenzi; Ph.D.; Mustapha Mouloua; Ph.D.; Peter A. Hancock [Ed.]: Human Performance, Situation Awareness and Automation: Current Research and Trends, HPSAA, I, Lawrence Erlbaum Associates, Inc., Publishers, S. 143 - 149, HPSAA II, Daytona Beach, 22.-25.03.2004, ISBN 0-8058-5341-3
[10] Klein, K. & Jakobi, J. (2003). BETA Recommendations Report. EC Growth Project BETA, Document 1999-RD.10804, D26 / 2003-05-07.
[11] Meier, Christoph; Jakobi, Jörn, "Verification and Validation Results from the Operational A-SMGCS Field Trials of the Project BETA", FAA-EUROCONTROL ATM-R&D Seminar, Baltimore, USA, 2005

Jörn Jakobi received his diploma in psychology from the University of Göttingen. Since 2000 he has been a human factors expert with DLR Institute of Flight Guidance. He was editor of the BETA operational concept and has coordinated operational A-SMGCS on-site trials at Hamburg, Prague, and Zurich. Currently he is managing the sub-project 'Concept' of the FP6 integrated project EMMA.

Functional Relationship of Elements of Apron and Terminal Building at the Airport

Igor Poltoracki

Abstract— **This paper deals with the problem of determination of the required number of certain elements of airport passenger terminal complex (aircraft parking positions, check-in counters, gate lounges, and baggage claim devices); as well as with determination of the functional relationship between these elements. For each of the abovementioned airport passenger terminal complex elements, an adequate simulation model for determination of the required number of these elements is developed, and several experiments - numerical examples, based on different input data (traffic scenarios), are carried out, showing that models reflect well the ideas they are based on. A set of analyses of results from numerical examples for particular scenarios, as well as comparative analysis of results obtained for different traffic structures in different models of airport passenger terminal complex elements are subsequently performed.**

***Index Terms*—Airport, Apron, Gate, Passenger Terminal, Capacity, Modelling**

I. INTRODUCTION

There are many papers addressing the problem of the capacity of existing or planned facilities at the airport. One particularly important and interesting problem related to modern airports is determination of the capacity of certain elements of a passenger terminal complex (e.g. aircraft parking stands). Namely, most of the large airports are faced with such a demand during peak hours that it exceeds the available number of aircraft stands (positions) adjacent to the terminal building.

However, this problem could be viewed from another perspective, in the sense that the scope is to determine required number of certain elements of airport passenger terminal complex. The motivation for this research comes from the fact that there are very few studies related to this problem, even though it is a very significant aspect in airport passenger terminal complex planning and management.

Within this research, only aircraft parking (gate) positions, and those elements (facilities) of an airport passenger terminal that are by its function connected to aircraft stands, and have direct impact on them, such as check-in counters, gate lounges, and baggage claim devices, were taken into consideration. The other elements (facilities) of an airport passenger terminal that may or may not appear in an arrival or departure flow, such as passport control, customs, security checks, etc., were disregarded.

I. Poltoracki is with the Air Transport Department, Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade, Serbia and Montenegro (e-mail: i.poltoracki@sf.bg.ac.yu)

The basic parameters that have influence on required number of previously mentioned elements of airport passenger terminal complex, are comprised in, or influenced by traffic schedule, e.g. aircraft arrival rate at the apron, fleet structure, gate (stand) occupancy time, passenger arrival rate at the airport, number of passengers and baggage per flight (both arriving and departing), etc.

The aim of this research was to develop several models for determination of the required number of previously mentioned elements of airport passenger terminal complex.

II. RELATIONSHIP OF OPERATION OF GATE POSITION AND OTHER FACILITIES

The relationship of operation of single gate position and other facilities (resources) that are related to it, and obligatory in departure and arrival flow (check-in counters, gate lounge, and baggage claim devices), is depicted in Fig. 1.

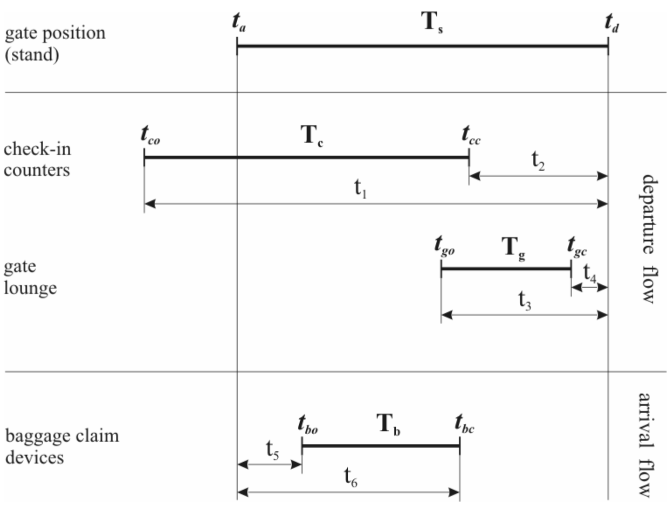


Fig. 1. Relationship of operation of single gate position and other facilities related to it

Where:
T_s - gate position (stand) occupancy time,
T_c - check-in counters occupancy time,
T_g - gate lounge occupancy time,
T_b - baggage claim devices occupancy time,

III. Model Definition

A. Gate (Parking) Positions

For the purpose of determination of gate (parking) position requirements, a couple of simple models were developed: one analytical, and one computer-based.

Several assumptions were introduced which were valid for both models:

- all aircraft are grouped in three classes: class 1 (e.g. ATR 72), class 2 (e.g. A320), class 3 (e.g. B767),
- aircraft of a certain size can only use gates that are specifically designed for these aircraft. However, a gate for a large aircraft can be used by all smaller-size aircraft,
- all aircraft will be served without delay or rejection.

1) Analytical Model

The analytical model is based on Horonjeff's [2] deterministic model for computing the required number of gate positions (G), based on volume of arrivals in aircraft per hour (A), and mean gate occupancy time in hours (T):

$$G = AT$$

This formulation does not account for the time separation required for manoeuvring aircraft between departure from a gate position and the next arrival, thus it underestimates the gate position requirement.

This lower-bound estimation of gate number can be increased either by introducing a "utilization" factor (U), as suggested by Horonjeff [2]:

$$G = AT/U,$$

or by adding a time period that represents the aircraft separation (buffer) time at a gate (S) to the gate occupancy time, as suggested by Bandara and Wirasinghe [3]:

$$G = A\ (T + S).$$

This separation time consists of push-out or power-out time, the time required by departing aircraft to clear the apron area, and the time required by arriving aircraft to move in from the apron entrance to the gate position.

The previous formulation was used in the analytical model for computing gate position requirements for certain size (class) of aircraft (i), as:

$$G_i = A_i\ (T_i + S),$$

as well as for total gate position requirements:

$$G = \Sigma G_i = \Sigma A_i\ (T_i + S),$$

where:

T_i is standard (average) gate position occupancy time for certain size (i) of aircraft. Adopted values were 40, 50, and 80 minutes, for aircraft class 1, 2, and 3, respectively,

A_i is the volume of arrivals of certain size aircraft (i) during the design hour, and $A = \Sigma A_i$. It should be emphasized that design hour doesn't have to be the same for different types (size) of aircraft, e.g. it is possible that there are several peak periods during the day with different distributions by type (size) of aircraft, so that the design hour is derived as some sort of "envelope" of those peak periods.

The adopted value for gate separation time (S) in the model was 10 minutes, regardless of the size of aircraft.

2) Computer-based (Simulation) Model

As was mentioned earlier, the previous analytical model is deterministic. In order to allow (introduce) stochastic, i.e. variation in time of arrivals at the gate position, as well as in gate (stand) position occupancy time, a simple computer-based (simulation) model was developed.

This model was made in MS Excel. Input for this model could be any traffic data (e.g. schedule) for one day (24h). Namely, model regards period between 00:00 and 23:59, whereas the changes of state are tracked down with an increment of one minute.

Input data (inserted manually) are times of aircraft arrival and departure at gate positions, i.e. appropriate schedule, modified so that it takes into account separation time at a gate (S), as in [3]. As in the analytical model, the adopted value for (S) was 10 minutes, regardless of the size of aircraft.

The logic of the model is actually based on a graphical method of superimposing the number of certain resources (facilities) that are occupied at the observed moment.

Besides determination of the total gate position requirement, which is determined based on peak traffic period (as in analytical model), this model provides information about the change in number of simultaneously occupied gate positions during the day. Since this model was made in Excel, it was very easy to produce appropriate charts and graphically represent output results.

B. Check-in Counters

In the case of check-in counters, only the computer-based (simulation) model was developed. As for the previous model, it was made in MS Excel, and the logic is based on the graphical method of superimposing the number of certain resources (facilities) that are occupied at the observed moment.

The following assumptions were introduced for this model:

- check-in counters are operated as dedicated, i.e. per-flight check-in, whereas the number of counters that are opened for a particular flight depends on the size of aircraft (1, 2, and 3 counters for aircraft class 1, 2, and 3, respectively),
- opening and closing times of the counters for certain flight are linked to aircraft departure time from the gate, regardless of the size of the aircraft. The adopted values were 90 and 20 minutes prior to aircraft departure time, for opening and closure of the counters, respectively.

The outputs from the model are also appropriate charts, which graphically represent the change in number of simultaneously operated check-in counters during the day (24h), with an increment of one minute.

C. Gate Lounges

The following assumptions were introduced for the gate lounge model:

- there are only gate (parking) positions adjacent to the terminal building (no "open" positions on distant apron),
- each flight uses a separate gate lounge, but it is possible (allowed) that in certain situations a single gate lounge, in a short time period, "handles" sequentially several aircraft parked on different stands (positions) adjacent to the terminal building. This is possible due to passengers being able to come from the gate lounge to certain aircraft in two different ways:
  a) directly through the air-bridge (if the aircraft is parked on a stand right in front of the observed gate lounge), or
  b) walking across the apron (in the case of the aircraft being parked at a nearby position adjacent to the terminal building). This case is feasible only if the aircraft is of the same or smaller size (class) than the size (class) of the observed gate lounge.
- opening and closing times of a gate lounge for certain flight are linked to aircraft departure time from the gate, whereas the period it is in use depends on the size of the aircraft. The adopted values for gate lounge usage time per flight were 15, 25, and 30 minutes for aircraft class 1, 2, and 3, respectively, whereas the closure time of the lounges was set to 5 minutes prior to aircraft departure time.

The model for gate lounges had also been developed in MS Excel, using the logic of the graphical method of superimposing the number of certain resources (facilities) that are occupied at the observed moment.

The outputs of this model are also appropriate charts, which graphically represent the change in the number of simultaneously used (occupied) gate lounges during the day (24h), with an increment of one minute.

D. Baggage Claim Devices

The following assumptions were introduced for the baggage claim devices model:

- all baggage claim devices are of the same size (capacity) and are sufficient for the largest aircraft (class 3),
- based on the previous assumption, the baggage from each flight performed by aircraft class 2 or 3 are placed on a separate baggage claim device; whereas it is possible, if there is a need, to use one baggage claim device to simultaneously deliver baggage from two flights performed by smallest (class 1) size aircraft (based on assumed size of such aircraft, compared with class 3 aircraft),
- beginning and final times for using a baggage claim device for certain flight are linked to the aircraft arrival time at the gate, whereas the period it is in use depends on the size of aircraft. The adopted values for baggage claim device usage time per flight were 15, 20, and 30 minutes for aircraft class 1, 2, and 3, respectively, whereas the starting time for usage of these devices was set at 10 minutes after the aircraft arrival time.

The model for baggage claim devices is also computer-based, and developed in Excel, using the logic of the graphical method of superimposing the number of certain resources (facilities) that are occupied at the observed moment.

As in previous models, the outputs are appropriate charts, which graphically represent the change in number of simultaneously used baggage claim devices during the day (24h), with an increment of one minute.

IV. Numerical Examples - Scenarios

For the purpose of validating the logic of the abovementioned models, several numerical examples - scenarios were performed.

Traffic data (scheduled and realized traffic for three peak days) obtained from Tivat Airport (a small seaside airport in Montenegro with a high seasonal peak in the summer) had served as a basis for these numerical examples. Fig. 2 depicts the apron at Tivat Airport (source: AIP Serbia and Montenegro). Beside traffic data, the data about usage of check-in counters for the same days was available.

In order to be implemented in the models, the above mentioned traffic data for three peak days had to be modified in a sense that cancelled, special and general aviation flights were excluded from the sample.

From those three peak days, the one with the largest total number of aircraft was chosen to be introduced into the models. After the previously cited modifications were made, a total of 24 aircraft during the above mentioned day remained, whereas the distribution by classes was: 4, 17, and 3 aircraft of class 1, 2, and 3, respectively.

Four different scenarios were considered within these numerical examples, out of which the first three scenarios are deterministic and fourth one is stochastic, where:

1) Scenario 1 represents scheduled traffic for a selected day, modified in a sense that average gate occupancy times for different size (class) of aircraft (calculated based on traffic data for all three days) are used. The adopted (average) values were 40, 50, and 80 minutes, for aircraft of class 1, 2, and 3, respectively,
2) Scenario 2 (basic scenario) represents original scheduled traffic for the same day, and is shown in Fig. 3,
3) Scenario 3 represents realized traffic for the same day as in Scenarios 1 and 2,
4) Scenario 4 represents a random variation of Scenario 2 in respect of aircraft arrival times at the gate positions, and gate occupancy times. These variations were generated using the Monte-Carlo simulation, based on appropriate distributions and cumulative frequencies of aircraft arrival lateness at the gate position, and deviation of realized gate occupancy time compared to the one planned by schedule, respectively.

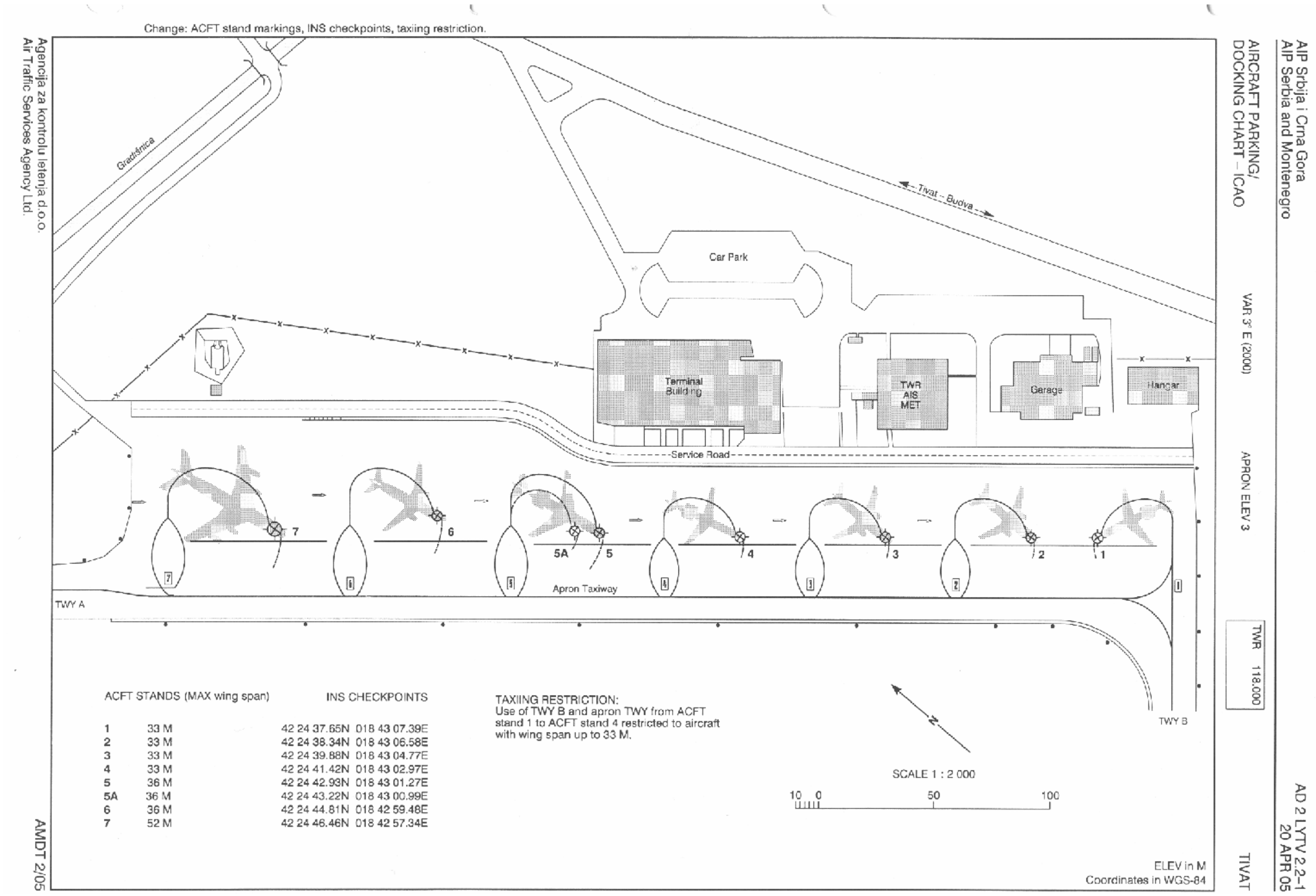


Fig. 2. Apron at Tivat Airport with the gate (parking) positions disposition (source: AIP Serbia and Montenegro)

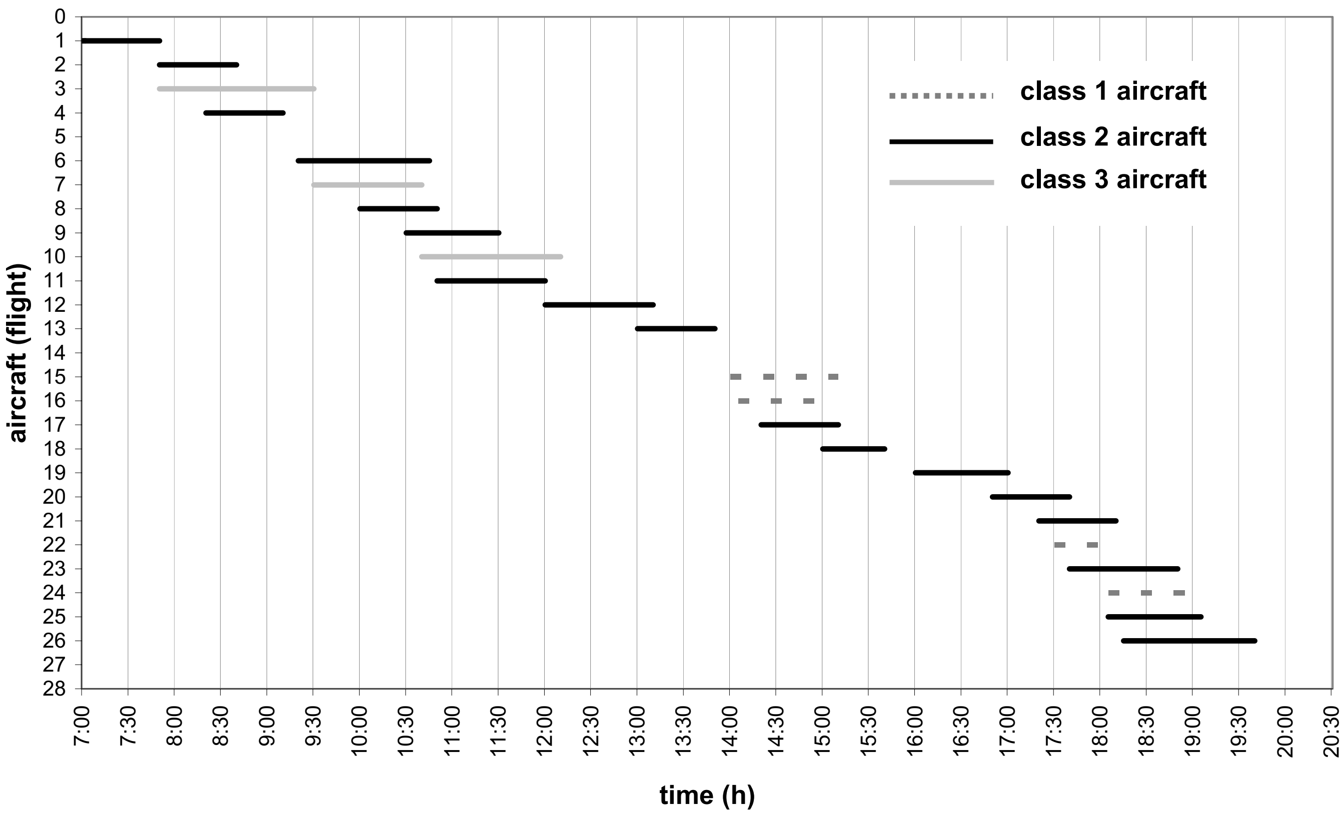


Fig. 3. Time distribution of requests for the gate (parking) positions (traffic data for Scenario 2)

Both of these distributions and appropriate cumulative frequencies were determined from traffic data for all three days. The distributions of arrival lateness and deviation from gate occupancy time, used for generating the variations in Scenario 4 are shown in Fig. 4, and Fig. 5, respectively. Within this scenario, 10 simulation runs (i.e. 10 iterations) have been performed for each of four defined simulation models of elements of airport passenger terminal complex.

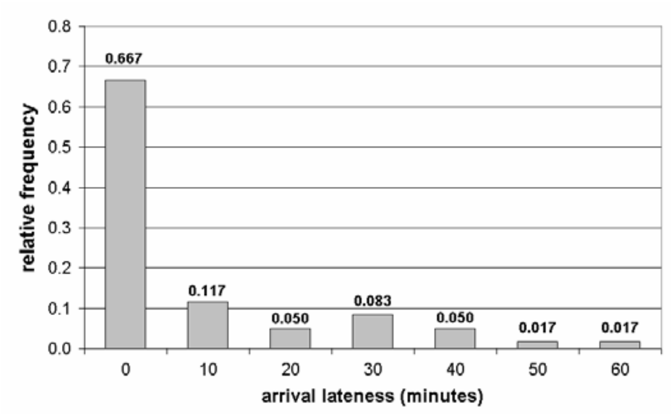


Fig. 4. Distribution of arrival lateness at the gate (parking) position

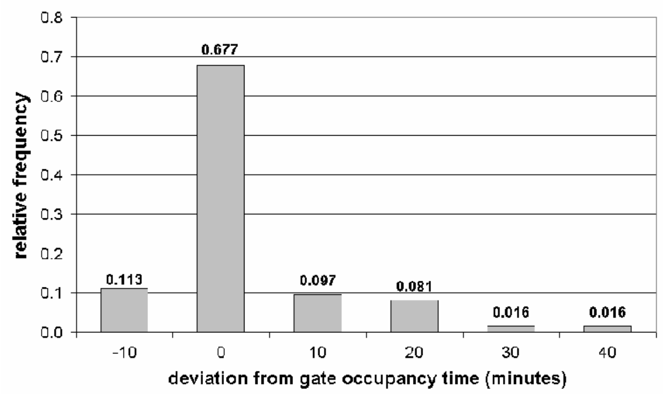


Fig. 5. Distribution of deviation from gate occupancy time

Within each of aforementioned scenarios, output results for each of the four simulation models (gate positions, check-in counters, gate lounges, and baggage claim devices) were obtained and analysed.

V. Output Results

For the purpose of illustration, the results obtained for the first three (deterministic) scenarios, for each of the models, are presented combined; whereas the results of the fourth (stochastic) scenario are shown separately, compared with the basic (second) scenario.

A. Gate (Parking) Positions

As it was mentioned in Section III-A, for the purpose of determining the gate (parking) position requirements, a couple of models were developed: one analytical, and one computer-based. With the goal of comparing and validating the logic of computer-based model, both models were first tested using Scenario 1 traffic.

The input values for the analytical model were two peak periods during a selected day with a different aircraft mix, as shown in Table I. This shows that design hour is not the same for aircraft of different sizes (as it was emphasized in Section III-A-1).

TABLE I
PEAK PERIODS USED AS INPUT DATA FOR ANALYTICAL MODEL

peak period	total number of arrivals	number of arrivals by the size of aircraft		
		A_1	A_2	A_3
10:00 - 11:00	4	0	3	1
17:20 - 18:20	6	2	4	0

The output results of the analytical model, rounded up to a first larger whole number are shown in Table II. It can be seen that results differ between the two observed periods, hence the ultimate values were obtained as a sort of envelope of values for those two periods, having in mind the adopted principle that smaller-size aircraft can park at gate positions for larger aircraft, whereas the opposite is not permitted.

TABLE II
ADOPTED VALUES OF OUTPUT RESULTS FOR ANALYTICAL MODEL

peak period	required number of gate positions by classes			total
	G_1	G_2	G_3	ΣG_i
10:00 - 11:00	0	3	2	5
17:20 - 18:20	2	4	0	6
envelope	**1**	**3**	**2**	**6**

The output results for the computer-based (simulation) model are shown in Fig. 6. From this figure it is easy to perceive two peak periods used for calculation in the analytical model. But beside that, the information about the change in number of simultaneously occupied gate positions during the day (total and by classes) is provided.

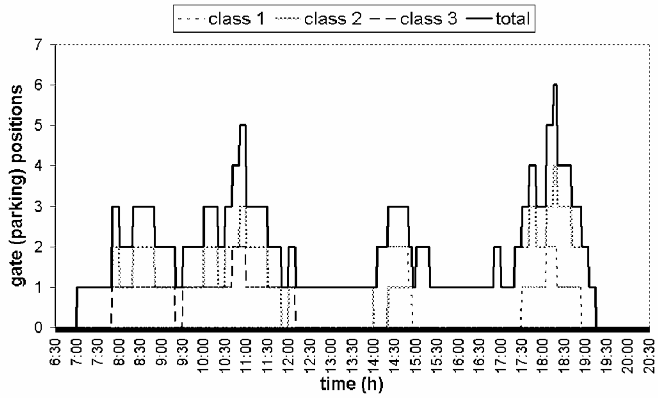


Fig. 6. Output results for simulation model of gate positions for Scenario 1

It can be seen that the values (both total and by classes) obtained using the simulation model for two previously mentioned peak periods are identical to the ones from the analytical model. This verifies that simulation model could be used for calculation of required number of gate positions. Having in mind the above mentioned advantages of this model (e.g. more information that it provides), it was decided that only the simulation model will be used for subsequent scenarios.

The above mentioned advantages of the simulation model for gate positions provided, as well, argument for usage of other simulation models (for check-in counters, gate lounges, and baggage claim devices), since the logic used in them is very similar.

As indicated at the beginning of this Section, the output results for the first three (deterministic) scenarios (total values only) are presented as combined results in Fig. 7, for better illustration and comparison.

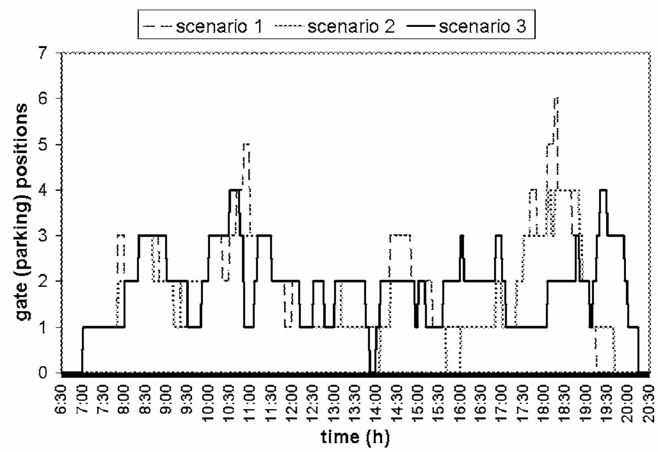


Fig. 7. Comparison of output results for model of gate (parking) positions for Scenarios 1, 2, and 3

Fig. 7 shows that only in Scenario 1 the required number of gate positions exceeds 4 (in two aforementioned peak periods).

The previous graph can be presented in another way, so that the total duration of states with a certain number of simultaneously used gate positions is shown, as in Fig. 8.

This type of graph is very useful for comparison when there are several sets of similar data that change over time.

From Fig. 8 it can more easily be seen that situations (states) when there are more than 4 simultaneously used gate (parking) positions have a very short total duration, which leads to very low utilization of some of the gate positions. This poses a question as to whether the adopted assumptions should be changed so that it is allowed that some aircraft will have to wait for the gate position to become free and experience some delay; whereas in turn, costs for the airport (both infrastructure and operational – stuff and equipment-vise) will be reduced.

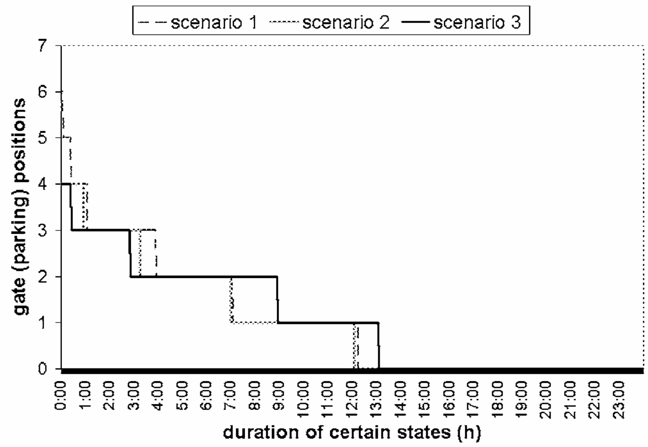


Fig. 8. Comparison of duration of states with certain number of simultaneously used gate (parking) positions for Scenarios 1, 2, and 3

The output results for Scenario 4 (stochastic one) are shown in Fig. 9. This graph depicts results from each of 10 iterations (total values only), combined with appropriate results from the basic (second) scenario, for better illustration and comparison.

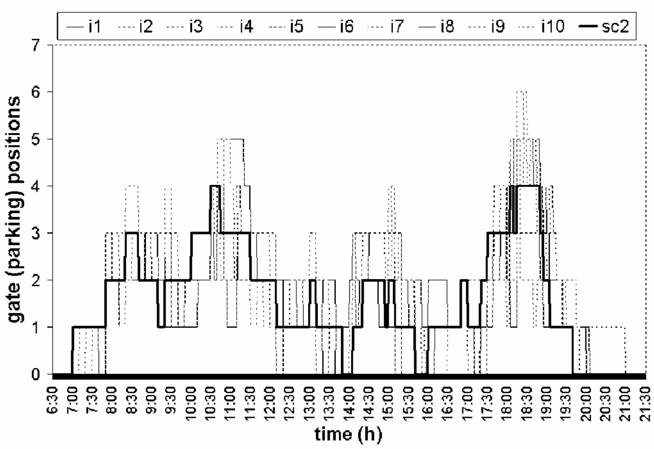


Fig. 9. Comparison of output results for model of gate (parking) positions for Scenario 4 (all 10 iterations) and Scenario 2 (basic scenario)

This figure shows that the results from 10 iterations are spread around the values from the basic (second) scenario, and follow the trend in change of those, which was expected. It can also be observed that for couple of iterations, the number of simultaneously used gate positions reaches the value of 6 during peak periods, whereas in the basic scenario the maximum value is 4.

The previous graph can also be presented in another way, so that the total duration of states with certain number of simultaneously used gate positions for each of 10 iterations from Scenario 4 are shown, compared with corresponding values from Scenario 2, in Fig. 10.

From this graph, it can be seen that in most iterations duration of certain states are longer, and that there are some iterations in which the maximum value of 4 simultaneously used gate positions from the basic scenario is exceeded. This is due to the fact that the share of positive values of distributions used for generating the variations in Scenario 4 is by far greater than of those with negative values (see Section IV, Fig. 4 and Fig. 5).

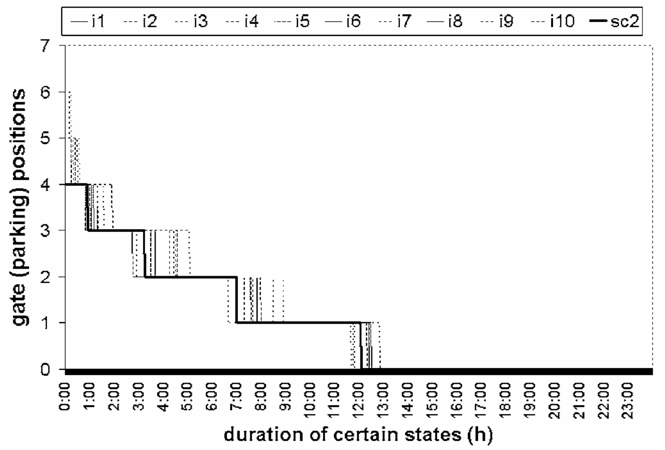


Fig. 10. Comparison of duration of states with certain number of simultaneously used gate (parking) positions for Scenario 4 (all 10 iterations) and Scenario 2 (basic scenario)

B. Check-in Counters

The output results for the simulation model of check-in counters for the first three (deterministic) scenarios (total values only) are presented combined in Fig. 11, for better illustration and comparison.

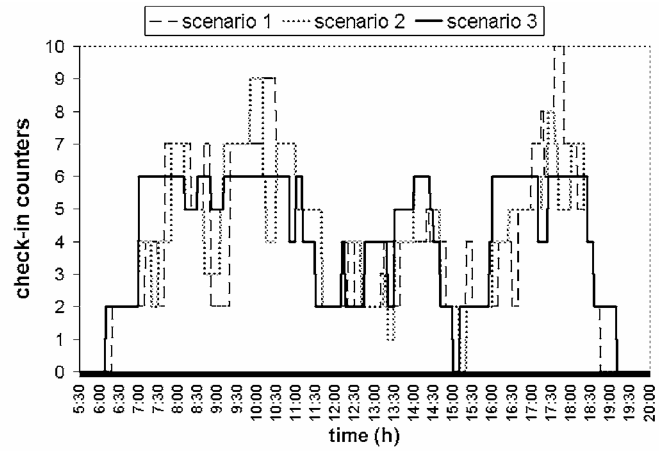


Fig. 11. Comparison of output results for model of check-in counters for Scenarios 1, 2, and 3

It can be seen that the results for scenarios 1 and 2 are almost matching, and show two peak periods which are linked to (precede) the previously indicated peak periods in the gate position requirements. The only greater difference between the results for the first two scenarios appears during the second peak period, where in Scenario 1 there is a request for 10 simultaneously opened counters, compared to request for 8 counters in Scenario 2. The peak periods in Scenario 3 are matching to a great extent the ones in the first two scenarios. The difference in the maximum number of simultaneously opened counters comes from the fact that input data for Scenario 3 (realisation) represents real data about usage of check-in counters obtained from Tivat Airport, and that there are only 6 counters installed and used at that airport. The tactic which is used to meet the requirements at peak periods is that not all of the check-in counters dedicated to a certain flight are used for the same period, i.e. some of them are opened later or closed earlier in order to switch between the overlapping requirements of different flights at peak periods.

As in the previous model, the data from the previous graph can be presented in another way, so that the total duration of states with a certain number of simultaneously opened check-in counters is shown, as in Fig. 12.

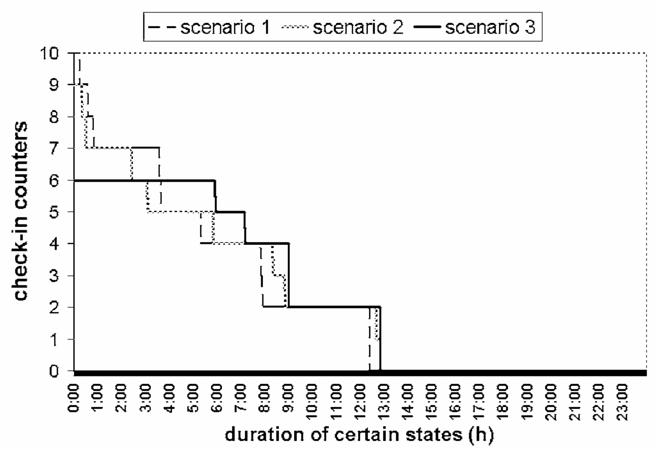


Fig. 12. Comparison of duration of states with certain number of simultaneously opened check-in counters for Scenarios 1, 2, and 3

From this graph it can be noticed even more easily that the previously explained corrections in length of period that some counters are dedicated to certain flights, can contribute to the reduction of investment in procurement of check-in counters, and avoid low utilization of some of the counters that appear in scenarios 1 and 2. Naturally, due to these corrections, the total duration of situations (states) with 5 or 6 simultaneously opened check-in counters in Scenario 3 is quite longer than in scenarios 1 and 2.

The model of check-in counters was not used with Scenario 4 (stochastic one). Namely, there was no point in varying opening and closing times of the counters, since the nature of the check-in process is such that the counters have to be opened during the scheduled period (regardless of eventual lateness in arrival or departure of a flight), as the passengers will come to the airport to perform their check-in according to the flight schedule.

C. Gate Lounges

Fig. 13 shows the combined output results for the simulation model of gate lounges for the first three (deterministic) scenarios (total values only). It can be seen that maximum required number of simultaneously used gate lounges is 3 for all three deterministic scenarios, whereas these maximum requirements appear in different periods during the day.

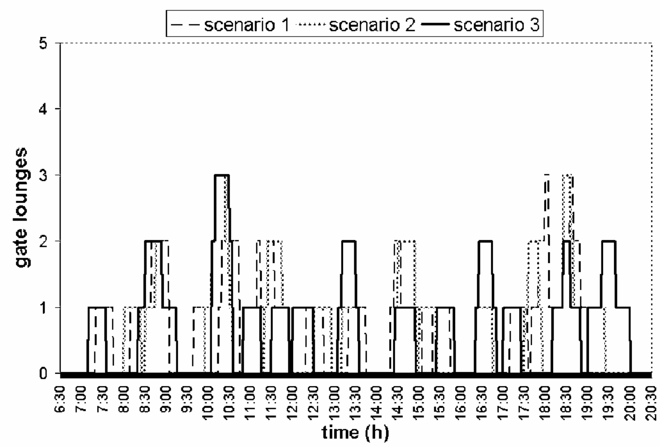


Fig. 13. Comparison of output results for model of gate lounges for Scenarios 1, 2, and 3

The data from the previous graph can, similarly to preceding models, be presented in another way. Therefore, a comparison of the total duration of states with a certain number of simultaneously used gate lounges for the first three (deterministic) scenarios is shown in Fig. 14. This figure shows that the maximum required number of 3 simultaneously used gate lounges appears in a very short period during the day in all three deterministic scenarios. Hence, it would be wise to perform small corrections in the planed length of usage of gate lounges for certain flights, and in turn reduce infrastructure or operational costs for the airport (depending on the planning level), and increase utilization of the remaining gate lounges at the same time.

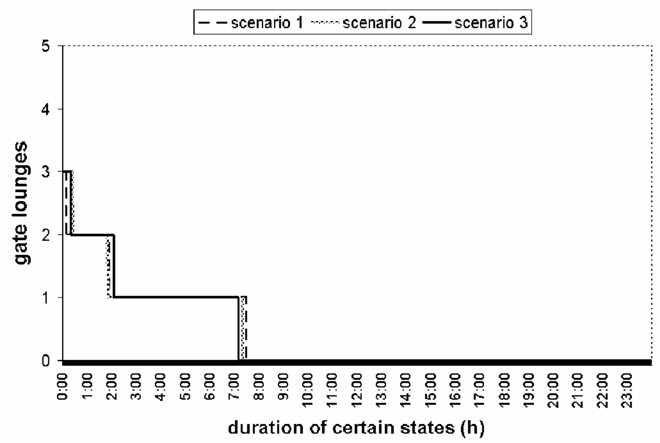


Fig. 14. Comparison of duration of states with certain number of simultaneously used gate lounges for Scenarios 1, 2, and 3

Since the opening and closing times of gate lounge for certain flight are linked to aircraft departure time from the gate, the variations of arrival time and gate position occupancy duration in Scenario 4 resulted only in shifting the planned period of usage of gate lounges for those flights that experienced delay in departure. This means that the passengers will, in case their flight is delayed in departure, stay longer in the central hall (or lounge) after check-in, waiting for a call and the gate to be opened.

The output results for Scenario 4 are shown in Fig. 15. This graph depicts results from each of 10 iterations (total values only), combined with appropriate results from the basic (second) scenario, for better illustration and comparison.

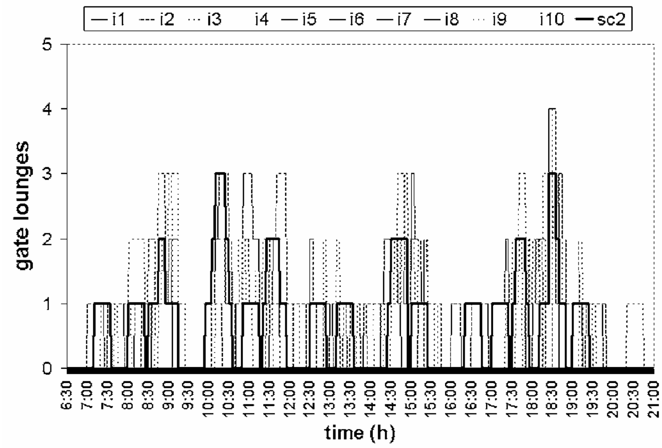


Fig. 15. Comparison of output results for model of gate lounges for Scenario 4 (all 10 iterations) and Scenario 2 (basic scenario)

Fig. 15 shows that results obtained in iterations are very similar to those in the basic scenario. The discrepancies are mostly up to one gate lounge, which was expected since the duration of usage of gate lounges for the flights with departure delay had remained unchanged.

It can be seen that maximum value of simultaneously used gate lounges in a couple of iterations reaches 4, whereas in basic scenario maximum value is 3.

For the sake of easier comparison, the data from the previous graph is presented in another way so that in Fig. 16 the total duration of states with a certain number of simultaneously used gate lounges for each of 10 iterations from Scenario 4 are shown, compared with corresponding values from Scenario 2.

This graph shows even more clearly that the corresponding values in iterations are, due to above mentioned remark, very similar to those in the basic scenario.

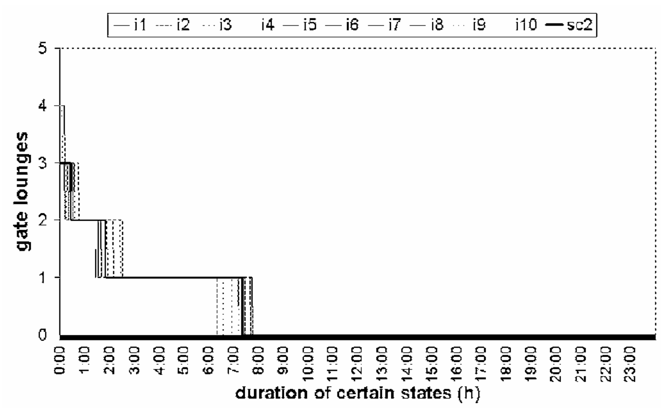


Fig. 16. Comparison of duration of states with certain number of simultaneously used gate lounges for Scenario 4 (all 10 iterations) and Scenario 2 (basic scenario)

D. Baggage Claim Devices

The output results for the simulation model of baggage claim devices for the first three (deterministic) scenarios (total values only) is shown as combined in Fig. 17. It should be emphasized that the results for scenarios 1 and 2 are identical (they are overlapping in the graph).

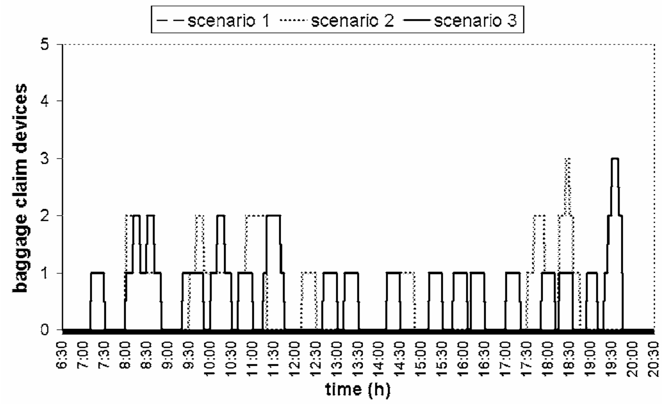


Fig. 17. Comparison of output results for model of baggage claim devices for Scenarios 1, 2, and 3

It can be observed from Fig. 17 that the maximum required number of simultaneously used gate lounges is 3 for all three deterministic scenarios, whereas these maximum requirements appear only once in each of the scenarios, but at different times during the day.

As in previous cases, the data shown in Fig. 17 can be presented in another way, for better comparison between scenarios (Fig. 18). The previous remark about the results from scenarios 1 and 2 being identical is also valid here.

It is noticeable that the values from Scenario 3 are very similar to those from the first two scenarios, and that the maximum required number of 3 simultaneously used baggage claim devices appears in a very short period during the day in all three deterministic scenarios. So, it would be reasonable to make small corrections in the planned length of usage of baggage claim devices for certain flights, in order to avoid engagement of a third device with its very low utilization.

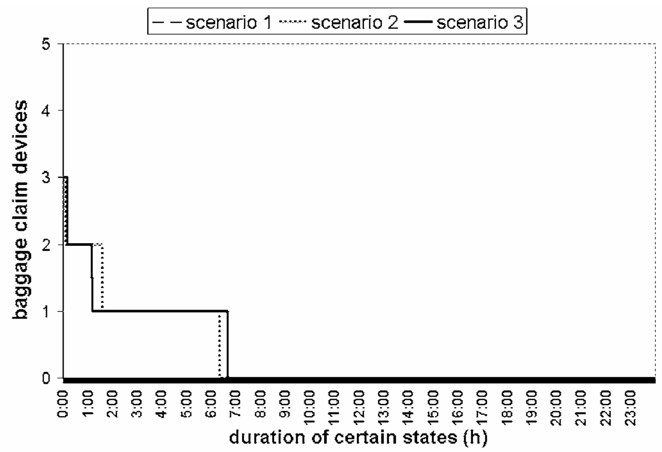


Fig. 18. Comparison of duration of states with certain number of simultaneously used baggage claim devices for Scenarios 1, 2, and 3

As the beginning and final times for usage of baggage claim device for certain flight are linked to aircraft arrival time at the gate, the variations of arrival time and gate position occupancy duration in Scenario 4, resulted only in shifting the planned period of usage of baggage claim devices for those flights that experienced arrival delay; whereas the duration of their usage remained unchanged.

The output results for Scenario 4 are shown in Fig. 19. This graph depicts results from each of 10 iterations (total values only), combined with appropriate results from the basic (second) scenario, for better illustration and comparison.

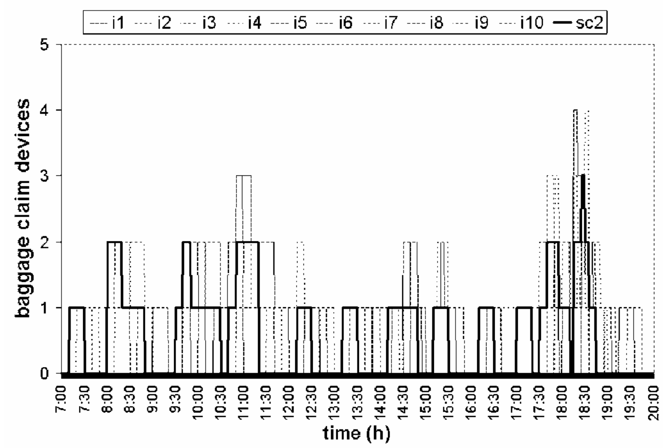


Fig. 19. Comparison of output results for model of baggage claim devices for Scenario 4 (all 10 iterations) and Scenario 2 (basic scenario)

The previous figure shows that results obtained in iterations are very similar to those in the basic scenario, and that differences are mostly up to one baggage claim device, which was expected in accordance with the remark about unchanged duration of usage of gate lounges for the flights with an arrival delay.

It can be seen that the maximum value of simultaneously used baggage claim devices in a couple of iterations reaches 4, whereas in the basic scenario the maximum value is 3.

For the sake of easier comparison, the data from the previous graph is presented in another way so that Fig. 20 depicts the total duration of states with a certain number of simultaneously used baggage claim devices for each of 10 iterations from Scenario 4, compared with corresponding values from Scenario 2.

This graph shows in an even clearer way that the corresponding values in iterations are very similar to those in the basic scenario, due to the above mentioned remark about unchanged duration of usage of baggage claim devices for flights with an arrival delay.

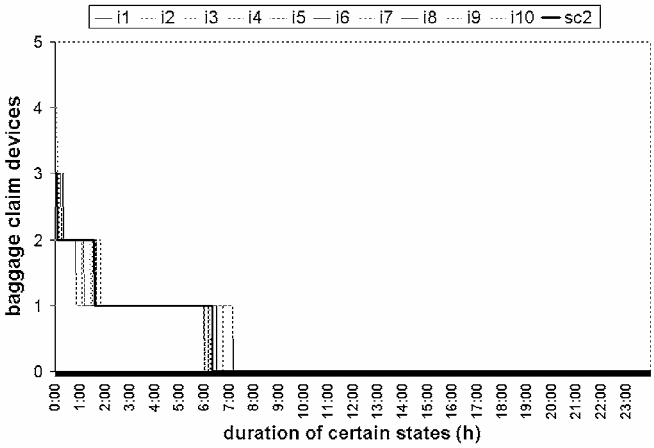


Fig. 20. Comparison of duration of states with certain number of simultaneously used baggage claim devices for Scenario 4 (all 10 iterations) and Scenario 2 (basic scenario)

VI. Conclusion

The research presented here considers the problem of determination of the required number of certain elements of an airport passenger terminal complex (aircraft parking positions, check-in counters, gates, and baggage claim devices)

For the purpose of determination of gate (parking) position requirements, a couple of models were developed: one analytical, and one computer-based. With the goal of comparison and validation of the computer-based (simulation) model, both models were tested using Scenario 1 traffic, showing that the logic of the simulation model is correct. Having in mind the advantages of this simulation model (e.g. more information that it provides), it was decided that only this model will be used for subsequent scenarios.

These advantages of the simulation model for gate positions provided, as well, argument for development and usage of other simulation models (for check-in counters, gate lounges, and baggage claim devices), since the logic used in them is very similar.

Several experiments - numerical examples, based on different input data (traffic scenarios), were performed, using each of previously mentioned simulation models. Namely, four different scenarios were considered within the cited numerical examples, out of which the first three scenarios were deterministic, whereas the fourth one was stochastic

A set of analysis of results from numerical examples for particular scenarios, as well as comparative analysis of results obtained for different traffic structures in different models of airport passenger terminal complex elements were performed afterwards.

The methodology developed and presented here could be used in the process of planning (dimensioning) the required resources (capacities) of airport passenger terminal complex (on tactical and strategic level).

Possible further research could be to introduce costs (take them into consideration), as well as to determine the cancellation probability.

References

[1] I. Poltoracki, "*Functional Relationship of Elements of Apron and Terminal Building at the Airport*," Master thesis, Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade, Serbia and Montenegro, (expected in June 2006).

[2] R. Horonjeff, *Planning and Design of Airports*, 2nd ed., McGraw-Hill, New York, 1975, pp. 150–153 and 265-268.

[3] S. Bandara, and S. C. Wirasinghe "Airport Gate Position Estimation Under Uncertainty," *Transportation Research Record 1199*, TRB, National Research Council, Washington, D.C., 1988, pp. 41-48.

[4] S. G. Hamzawi, "Management and Planning of Airport Gate Capacity: A Microcomputer-Based Gate Assignment Simulation Model," *Transportation Planning and Technology*, Vol. 11, 1986, pp. 189-202.

[5] M. I. Hassounah, and G. N. Steuart, "Demand for Aircraft Gates," Transportation *Research Record 1423,* TRB, National Research Council, Washington, D.C., 1993, pp. 26-33.

[6] G. N. Steuart, "Gate Position Requirements at Metropolitan Airports," *Transportation Science*, Vol. 8, 1974, pp. 169-189.

[7] S. C. Wirasinghe, and M. Shehata, "Departure Lounge Sizing and Optimal Seating Capacity for a Given Aircraft/Fleet Mix - (i) Single Gate. (ii) Several Gates," *Transportation Planning and Technology*, Vol. 13, 1988, pp. 57-71.

[8] V. Tošić, O. Babić, and M. Janić, "Airport Passenger Terminal Subsystems Simulation Models," *Proceedings SYM-OP-IS '81*, Herceg Novi, 1981, pp. 353-369.

[9] V. Tošić, "A Review of Airport Passenger Terminal Operations Analysis and Modelling," *Transportation Research,* Vol. 26A, No. 1, 1992, pp. 3-26.

[10] V. Tošić, and O. Babić, "Modelling Passenger Terminal Operations at Medium-sized Airports," *Proceedings - Air Transport in Central Europe Conference*, Žiline, 1993.

[11] M. Janić, *Air Transport System Analysis and Modelling (Capacity, Quality of Services and Economics)*. Gordon and Breach Science Publishers, 2000, pp. 45-48.

[12] M. Mirković, "*Air Passenger Check-in Simulation*," Graduation thesis, Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade, Serbia and Montenegro, 1999.

[13] M. Stojković, and V. Tošić, "A Model for Calculation of Passenger Terminal Elements Load," *Proceedings SYM-OP-IS '86*, Herceg Novi, 1986, pp. 537-544.

[14] *AIP Serbia and Montenegro*, SMATSA, Belgrade, 2006.

[15] *Passenger Terminal Complex (Chapter 3)*, IATA - Airport Terminal Reference Manual, 7th Edition, Geneva, 1989.

[16] *ATR 72 Airplane Characteristics for Airport Planning*, ATR Product Support, Technical Publications Department, Toulouse, 1988.

[17] *A320 Airplane Characteristics for Airport Planning*, Airbus Industrie Product Support, Technical Publications Department, Blagnac Cedex, 1985.

[18] *B767 - 200/300 Airplane Characteristics for Airport Planning*, Boeing Commercial Airplanes, Seattle, 1989.

Igor D. Poltoracki was born in Belgrade, Serbia and Montenegro in 1973. He has graduated at the Faculty of Transport and Traffic Engineering (FTTE), University of Belgrade at Department of Air Transport in 1999. He is currently postgraduate student and is finishing his Master Thesis (expected in June 2006). He is working at the Department of Air Transport, FTTE, since 1999. Besides work on lecturing and research, he has been involved in several aviation related projects and studies. His main field of interest are: Airports, ATM, Modelling, Simulation and Transport and Traffic Analysis.

Session 3

Analytical Modeling 1

The Effect of the Planning Horizon and the Freezing Time on Take-off Sequencing

Jason A. D. Atkin and Edmund K. Burke
School of Computer Science and I.T.
University Of Nottingham, Jubilee Campus
Nottingham, NG8 1BB, England
Email: jaa/ekb@cs.nott.ac.uk

John Greenwood
NATS CTC
4000 Parkway, Whiteley, Fareham
Hampshire, PO15 7FLT, England

Dale Reeson
National Air Traffic Services
Heathrow Airport, Hounslow
Middlesex, TW6 1JJ, England

Abstract*— In this paper we consider the problem faced by runway controllers at London Heathrow Airport as they attempt to determine the best order for aircraft to take off. The order in which aircraft take off can have a large effect upon the throughput of the runway and the consequent delay for aircraft. Although two runways are available for use at Heathrow, in order to control noise for residents on the flight paths, only one is used for departures at any time. The other is used for arrivals. As this is one of the busiest international airports in the world, and the runway is the bottleneck for the departure system, it is important to attain high runway throughput. A runway controller re-arranges the aircraft for take-off within holding points at the ends of the runways, currently performing this task manually. The decision about the take-off order has to be made with very limited decision time, precipitating investigations into the development of an on-line decision support system to aid in this task. We have developed a model for such a decision support system and a simulation of the departure system, and discuss these in this paper. Our experiments predict that our system can provide suggestions fast enough to be of use in practice, while also being able to consider more aircraft than a real runway controller can. Thus it can obtain consequent benefits from highlighting potential problems before they occur. In order to maximise the potential benefits of a decision support system, it is important to understand how the decisions are affected by the planning horizon, how far ahead the system considers aircraft. From this information, we can better understand the inputs that the system would need to be able to fulfil its role. The position of an aircraft in the take-off order has to be frozen at some point prior to take-off. Our investigations revealed a trade-off between the time at which the take-off order was frozen and the planning horizon necessary to get the best results from the system. This paper considers this trade-off, evaluates where the planning horizon needs to be, and shows the relationship with the point at which the take-off schedule is frozen. We present results which show that there is a substantial delay benefit from including taxiing aircraft. Furthermore, we show that, with a schedule frozen for two minutes before take-off, the vast majority of the benefit can be gained from a knowledge of only those aircraft which will arrive at the holding point within the next eight minutes, and that aircraft can be ignored until they push-back from their stands. Finally, the results also show how the planning horizon must increase if the time for which the schedule is frozen increases.

I. Introduction

A schematic diagram of Heathrow can be seen in figure 1. The airport currently has two runways, north and south, and both can be used in either direction. Each runway is named according to the current direction of use and whether it is on the left or right from that direction. There are currently four terminals, labelled T1 to T4 in figure 1, although a fifth is being built. Near the ends of each runway, the taxiways form holding points, labelled HP in the diagram, within which a runway controller manually performs the complex task of ordering the aircraft for take-off. Heathrow's proximity to housing means that only one runway can be used for departures at any time, with the other being utilised by arrivals. This makes it vital to achieve a high throughput on the one available departure runway.

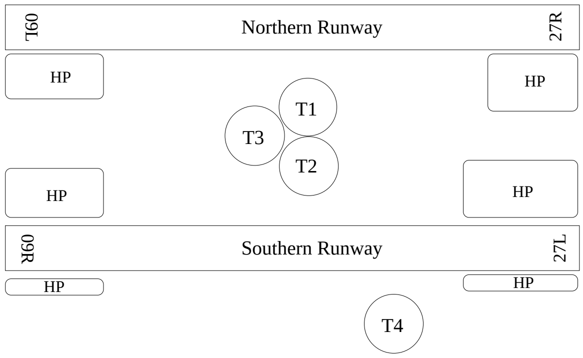


Fig. 1. The layout of London Heathrow Airport

We have designed a decision support system for the runway controller, to aid in the difficult task of ordering the aircraft for take-off. In order to make the best use of this, it is vital to understand the effect of other parts of the departure system upon it. In this paper we consider the advantages of moving the planning horizon, giving the system knowledge of some or all of the aircraft on the taxiways, allowing these to be considered in the schedule, before they actually come under the control of the runway controller. We also consider the effect of changing the time at which the schedule is frozen, effectively changing the other end of the planning window, and consider the relationship between this and the planning horizon.

In this paper, we first present the take-off scheduling problem and discuss the planning horizon and schedule freezing issues. We then present our decision support system and departure system simulation, which we use to investigate the effects of varying the planning horizon and the time at which the schedule is frozen. We end the paper by presenting the results of our experiments and by drawing conclusions about the implications.

II. The take-off scheduling problem

The take-off scheduling problem can be summarised as finding a take-off order with low delay, low workload for pilots and controllers and maximal compliance to the various temporal and physical constraints upon the aircraft in the schedule. It is important, as the delay in the schedule has obvious implications for airlines, passengers and the airport.

When aircraft take off, a minimum separation time is required between them. Unnecessarily large separations delay the following aircraft from taking off, causing unnecessary fuel burn, with the associated financial and environmental costs. The runway controller's aim is to order the aircraft so that the number of larger separations that are required is reduced, consequently reducing the delay for the aircraft awaiting take-off.

Mandatory separation times are required between aircraft at take-off, to ensure that the wake vortices created by the previous aircraft to take off have had time to dissipate. This *'wake vortex separation'* depends upon the weight classes of the current and preceding aircraft. Every time the take-off schedule has a lighter class aircraft following a heavier class aircraft, a larger separation will be required, adversely affecting runway throughput and delaying later take-offs.

Aircraft take off along fixed routes, called *'Standard Instrument Departure'* (SID) routes. A further mandatory separation time is required to ensure that in-flight separation distances will be attained. This *'SID separation'* is based upon the relative SID routes on which the aircraft are departing. Aircraft departing on the same or similar SIDs may need a larger separation to be applied at take-off to ensure that the in-flight separation is attained. This separation is further modified according to the speed groups of the aircraft, to allow for the fact that a faster or slower following aircraft will decrease or increase the separation distance. A good schedule will, therefore, often ensure that aircraft with similar SIDs are not in adjacent positions in the take-off order.

Some aircraft have a *'Calculated Time Of Take-off'* (CTOT) assigned to them. This designates a target take-off time and is assigned to avoid congestion, en-route or at destination airports, by staggering the arrivals to the congested position. As aircraft are permitted to take off up to five minutes before or ten minutes after the CTOT time, it effectively designates a fifteen minute take-off window. CTOTs are assigned without regard to the source airport, so can be difficult, or impossible, for the runway controllers to achieve at busy periods. To allow for this, a limited number of five-minute extensions can be used, but as few as possible should be employed.

Take-off scheduling has been considered in a number of research papers in the past. Idris et al. considered the departure system at Boston Logon airport in [1] and [2], concluding that the runway represents the bottleneck. Anagnostakis et al. developed a two-stage departure planner, which they presented in [3] and [4]. A constraint satisfaction based model was applied to the take-off scheduling problem by van Leeuwen et al in [5]. Trivizas used a maximum position shift approach in [6], with dynamic programming.

There are similarities between arrival scheduling and departure scheduling, but also important differences. Abela et al, [7], Beasley et al, [8] and [9], and Ernst et al, [10], have all considered the arrival scheduling problem and applied different methods to solving the problem. However, although there are similarities in the wake vortex separation rules, the downstream constraints and the constraints inflicted by the holding point structure do not apply. Thus, the direct application of such methods to our problem is not possible.

The departure problem was considered by Bianco et al. as a special case of the cumulative asymmetric travelling salesman problem (ATSP) with release dates, in [11]. However, the equivalency does not hold for Heathrow, as the SID separation rules do not obey the triangle inequality, so it is not always sufficient to only ensure adequate separations between adjacent take-offs.

At Heathrow, aircraft are directed around the taxiways by Ground Movement Controllers (GMC) to holding points near the end of the runway, within which a runway controller aims to reorder them for take-off. Due to the complexity of the roles of the controllers involved, and the time constraints upon them, there is often little co-ordination between them for individual aircraft. The GMC usually determines to which of a number of holding point entrances each aircraft should be delivered, considering issues such as ease of reaching the entrance, congestion in the holding area and on the taxiways, and any requests the runway controller may have made.

Performing the overtaking within the holding points, in the last few minutes before take-off, ensures that the uncertainties in the take-off schedule are minimised. The controller does not need to consider the variability of taxi-times nor the contention with arriving aircraft for stands that would occur if the scheduling was performed at the stands. However, the structure of the holding points restricts how much overtaking can be performed. To be overtaken, there must be a position in which an aircraft can wait and a clear path for the overtaking aircraft to go past.

As, at Heathrow, the overtaking to achieve the take-off schedule is performed within the holding points, the constraints they inflict upon what overtaking is possible must be considered in the solution method chosen. The applicability of the aforementioned research to the take-off scheduling at Heathrow is, therefore, limited. In [12], Craig et al. considered the effect of a simplified holding point structure at Heathrow, applying a dynamic programming approach to solve it. The position of the aircraft have to be included as they have a great effect upon the feasibility of further reordering. The holding points are actually much more complicated than in [12], however, so the feasibility of the dynamic programming approach quickly breaks down as the number of possible positions for aircraft increases. We presented an alternative approach in [13], further developed the approach in [14] and apply it in this paper to investigate the planning horizon and take-off freezing time.

III. Planning Horizon and Take-off Freezing

Departing aircraft can be considered to pass through the following states:

1) At the stand.
2) On the taxiway.
3) At the holding point.
4) At the holding point in a frozen take-off order.
5) Taken off.

A runway controller will usually only consider the aircraft that are already within the holding point when determining the take-off order. The first aspect that we consider in this paper is the effect of increasing the planning horizon beyond the holding point arrival time, so that at least some of the aircraft on the taxiways, and possibly also some that are still at their stands, are considered in the take-off schedule.

The earliest take-off time is a vital piece of information for any take-off scheduling system. In practice, however, the variability in the taxi times means that there is a degree of uncertainty in the earliest time at which an aircraft could reach the runway. This uncertainty decreases as the aircraft gets closer to the holding points and the runway.

It is useful to understand the effects of the planning horizon as it is easier to obtain shorter term predictions of taxi times than longer term ones. The task of a taxi time prediction system (providing these to a decision support system) would, consequently, be simplified. Understanding the effects can avoid the need for an unnecessarily large planning horizon. Additionally, larger planning horizons mean incorporating more aircraft in the search, so the search space is much larger, making the job of the decision support system much harder. Finally, it is useful to know whether aircraft need to be considered before they have pushed-back from their stands, as there will always be much more uncertainty involved in the holding point arrival time until that point.

The second aspect we wish to consider is the effect of take-off freezing. At some point before take-off the aircraft will have been given instructions for take-off or actually be lining up for take-off. At this point it is impractical to change the position of that aircraft in the schedule, except in extreme circumstances such as an inability to take-off for some reason.

The runway controller will usually give some conditional clearances to pilots, telling them to line up for take-off following another specified aircraft. Conditional clearances effectively freeze that part of the take-off schedule. They are used as they have workload advantages for a runway controller, as well as giving the pilots more visibility of their planned take-off time. The take-off order could still be changed after conditional clearances have been given, but doing so will involve more work than before they are given.

Additionally, a controller will normally have a planned take-off order in mind for the aircraft in the holding points, probably with gaps later in the schedule that future arrivals could fill, if possible. Reconsidering the order for these aircraft requires time and effort, so is not always practical due to the extremely busy workload of the controller. Therefore, often, more of the schedule, beyond that where conditional clearances have been given, can be considered as frozen, for all practical purposes, as it may be frozen from the point of view of the controller making the decisions. Of course, if necessary, changing this part of the schedule is less costly than changing the schedule where conditional clearances have been given, it will just not necessarily be considered.

We can consider a conditional clearance to move an aircraft from state 3 to state 4 in the list above. Conditional clearances given by controllers are not actually as constraining as freezing a part of the take-off schedule is. It is often possible, with only limited extra work, for a controller to fit an extra take-off into an existing large separation in the take-off schedule, especially if the take-off times for the other aircraft are unaffected, assuming, of course, that the new aircraft can perform the required overtaking within the holding points.

To investigate the effects of early freezing of the schedule, and of the planning horizon, we used a decision support system we had designed, together with an abstract simulation of the departure system of the airport. To ensure that the experiments were realistic, we used real, historic data provided by National Air Traffic Services. This data included temporal information such as the times at which the aircraft pushed back from their stands, arrived at the holding points and took off, as well as the details of the aircraft involved, such as the weight classes, SID routes and speed groups. In the following sections we explain the decision support system, simulation and experiments.

IV. Decision Support System

A decision support system for a runway controller must make decisions about the desirable take-off order using only the information available at the time. The responsiveness of an on-line decision support system to a changing situation is determined by the search time required to make each decision. Our system has a search time of around a second on a 2.4GHz pentium 4, making it fast enough for use as a real-time system.

A decision support system will have to solve a sequence of problems over time, each of which consists of only a snapshot of the daily schedule. A system would be running constantly, re-deciding what to do as the situation evolves, responding to changing circumstances such as new aircraft pushing back from stands or the effects of previous decisions. The aim, however, is to obtain the best overall schedule for the day, rather than necessarily the solution of highest throughput at any particular instant in time. The suggested schedule that the system returns must, therefore, not cause problems for the scheduling of later take-offs.

Furthermore, as the overtaking required to achieve a take-off scheduling takes place within the holding points at the end of the runway, our decision support system not only considers whether a take-off order is desirable, but also whether it is achievable within the holding points.

Our decision support system has three main parts. The first part uses a tabu search to investigate the possible take-off orders and to seek a high quality one. The second part is a system to verify whether the overtaking required to achieve a desired take-off schedule is possible. The third part determines how good a take-off schedule is by predicting take-off times for aircraft and evaluating a consequent cost for the schedule. Each of these parts is described below.

A. Tabu search

Tabu search was first introduced by Glover in [15] and has been applied to many different kinds of optimisation problems. Further information about Tabu Search can be found in [16]. The flow of our tabu search is illustrated in figure 2. Starting at an initial, achievable, take-off schedule (called a solution) the search makes progressive changes, seeking better and better schedules.

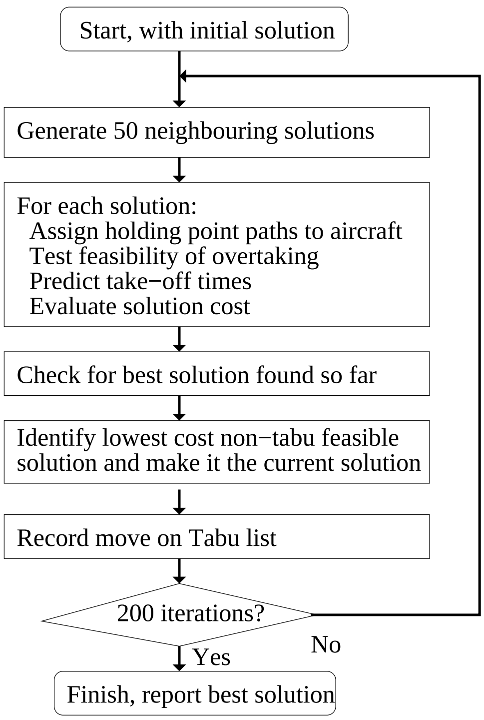


Fig. 2. The tabu search

Each neighbouring solution is created by applying a random move. Available moves include exchanging the positions of two aircraft in the schedule or moving up to five aircraft forward or backward in the schedule. In this way, fifty different schedules are created at a time. The search then selects the best of these and adopts it as the new current solution.

In order to inhibit the search from cycling between a few good schedules, a tabu list is applied. Whenever a move is made, the details of the old positions of the aircraft that were moved in the schedule are stored. The next ten moves after that one are explicitly prohibited from returning all of these aircraft to the positions from which they came. Any move which would do so is declared to be tabu and will not be accepted as the new current schedule. This avoids the search returning to schedules that have been recently evaluated, aiding it in escaping some of the local optima. Even if a solution is declared to be tabu, however, it is still recorded as the best if it has a lower cost than the best solution found so far.

A check is performed for each solution to verify that the required overtaking is achievable within the holding point. If it is not, the solution is declared as infeasible and will not be adopted. During the search the best solution found so far is always maintained. This is returned as the suggested schedule at the end of the search.

B. Overtaking within the holding points

It is important to verify that the required overtaking can be performed within the holding points. To ensure this we have a two-stage process. First sensible paths are assigned to aircraft, then the holding point model is used to verify that the overtaking can all take place.

We use a directed graph model of the current holding point structure for this verification. An example graph, for the 27R holding point, is given in figure 3. The graph has a node for each valid waiting position for an aircraft and arcs for transitions aircraft can make between nodes as they move through the holding point.

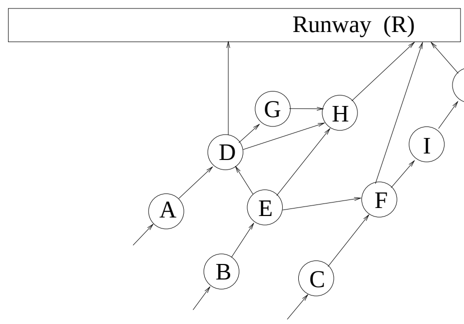


Fig. 3. The 27R holding point network

A path assignment heuristic is used to allocate paths through the holding point to aircraft. Paths are assigned based upon the overtaking that is required. Each path can be uniquely identified by the set of nodes it passes. Fast paths are assigned to aircraft which overtake and slower ones to those which are overtaken. For example, with the holding point given in figure 3, if two aircraft that arrive at entrance A must reverse their order the overtaking one will be assigned the path ADHR and the overtaken one ADGHR. The heuristic method ensures that the path allocation is sensible, from the point of view of the runway controller. This is important as, otherwise, a solution would be immediately rejected, and the decision support system would be worthless.

Once paths have been allocated, a feasibility check is performed to verify whether the required overtaking is possible. The holding point graph for the current holding point is used to determine this.

Initially, each aircraft that is already in the holding point is placed in the node related to its current position. If the aircraft is between nodes it is placed at the node it will next enter. Aircraft that are currently on the taxiways are placed in queues, in predicted arrival order, at the holding point entrance at which they are predicted to arrive. If any entrance node is empty then the first aircraft from that entrance queue is placed in the node and removed from the queue.

The overtaking test is then performed by moving aircraft, one node at a time, through the holding point graph towards the runway. Each aircraft can only enter the next node on its path if the node is empty, and entering it will not block another aircraft from reaching the runway on time. A fast method has been developed for verifying that the latter condition has been met. This works by building partial take-off orders for each node and tracking the number of free nodes that an aircraft could use to move out of the way of an aircraft that should take off before it.

If all aircraft can enter the runway node in the desired take-off order then the order is achievable within the holding point. If not, then the schedule is discarded as being unachievable, or infeasible.

C. Schedule cost

The tabu search requires an objective measure of the cost of a schedule, its aim being to find a low cost solution. To measure the cost of a schedule, take-off times are predicted for all aircraft, then the total delay is measured and the number of missed CTOTs are counted.

Predicting take-off times requires that the earliest time of take-off is calculated for each aircraft, allowing sufficient time for the aircraft to reach the runway and ensuring that all required separations are maintained.

The time at which an aircraft will be able to reach the runway is calculated by adding an expected traversal time of the holding point to the predicted holding point arrival time. The traversal times used are pessimistic to ensure they are achievable.

The earliest time an aircraft can take off while obeying separation rules can be calculated by considering each previous take-off in turn. Adding the higher of the required wake vortex or SID separations to the take-off time of the earlier take-off gives an earliest take-off time for this aircraft.

If the aircraft has a CTOT then it is not permitted to take off more than 5 minutes before the CTOT time. This forms an additional constraint for the earliest take-off time.

As there is never an advantage in delaying the take-off, aircraft are assumed to take off as early as they can, so the predicted take-off time is assumed to be the earliest time which allows the aircraft to physically reach the runway, obey any CTOT and fulfil all required separation rules.

Once the take-off times have been predicted, a cost for the schedule can be calculated. The cost takes account of the number of CTOT extensions required, the total time for which the aircraft are in the holding points and the deviation of the schedule from the first-come-first-served schedule. The cost for the schedule is a combination of these three factors, weighted so that reducing the CTOT extensions is the primary criterion, reducing delay is secondary and reducing deviation is tertiary. If any schedule involves an aircraft which is scheduled too late even for a CTOT extension it is given a prohibitively high cost, preventing the search from adopting these schedules.

We refer to the time in the holding points as the delay for aircraft. We use a *delay-based* rather than *throughput-based* measure for the cost of the solutions to the problems. Both delay-based and throughput-based measures share the characteristic of penalising larger separations. However, a delay-based measure penalises these separations more if they are earlier in the schedule, as they delay more aircraft. It, therefore, favours schedules where the large separations are later on in the schedule and are more likely to be utilised by later aircraft. This often aids later iterations of the search to perform better.

V. Simulation

Our departure system simulation allows us to investigate the consequences of adopting the take-off sequence recommended by the decision support system. The simulation maintains information about the expected arrival time at the holding point for aircraft on the taxiways and the position within the holding point of any aircraft currently there.

The simulation uses a one minute timestep, presenting the decision support system with a sequence of problems consisting of only the information an on-line decision support system would have. At each step, the decision support system is given the opportunity to order the aircraft for take-off, providing the simulation with information about the desired take-off order and how it will be achieved. The simulation is responsible for using this information to update the aircraft positions before presenting the next problem to the system to solve.

As mentioned in section III, some or all of the aircraft on the taxiways may be included in the problem the simulation presents to the system. For these aircraft, the system is given a predicted arrival time at the holding point. In our tests we assumed no uncertainty in the arrival time, as we wished to evaluate the effects of varying the planning horizon without having to account for the effects of the uncertainty. The prediction is, therefore, always accurate in these tests.

The decision support system also needs to know which holding point entrance (for example, A, B or C in figure 3) the aircraft is due to arrive at, in order to test the feasibility of performing the required overtaking. For these experiments we assumed the Ground Movement Controller (GMC) delivered each aircraft to the nearest entrance to its stand. Experiments have shown that better results can be obtained by delivering the aircraft more evenly across entrances, allowing more flexibility in rescheduling. However, assigning aircraft to the nearest entrance is the only allocation method that is guaranteed not to involve prohibitive work from the GMC under any circumstances.

At the end of each experiment, all aircraft will have been scheduled for take-off and the final schedule can be evaluated to determine the CTOT compliance and delay for aircraft.

TABLE I

DATASET DETAILS AND MANUAL RESULTS

Dataset	Number of aircraft				Manual results	
	Heavy	Medium	Light	CTOTs	CTOTs	Delay (s)
1	90	239	1	100	5	366
2	100	244	1	98	5	312
3	64	193	2	42	4	372

VI. RESULTS

We performed a series of experiments to evaluate the effects of varying the planning horizon and varying the schedule freezing time. Each experiment was performed ten times, for each of the same three datasets, because of the stochastic nature of the tabu search. The mean results are shown in the tables. In fact, in every case, the results for each of the ten executions were very close to each other, the CTOT compliance being the same in all cases and the delay varying only slightly. The number of aircraft of each weight class and the number with CTOTs are shown in table I, along with the performance of the real controllers in terms of the number of CTOTs missed and the average delay per aircraft.

In our first experiment we evaluated the effect of varying the planning horizon by altering the amount of knowledge the system has about taxiing aircraft. Aircraft are included in the simulation if they will both arrive at the holding point within a given time window and have already left their stand. For example, if this taxi knowledge time is 0, then aircraft are only considered once they have reached the holding point. Table II presents details of the performance of the system as the taxi knowledge time is changed. The number of CTOTs missed and the total delay for the aircraft in the system is given for each of the three datasets. In this experiment, the decision support system froze the take-off order for aircraft within the holding point two minutes before take-off and penalised changes made within three minutes of take-off.

We can clearly see from table II that the delay is significantly improved by giving the decision support system some knowledge of the aircraft currently on the taxiways, and that, in general, the system performed better with more knowledge. In some cases, the CTOT compliance was also improved. In all cases, however, the delay seems to plateau at around eight minutes taxi knowledge.

We note, however, that aircraft were only added once they had left their stands, so there is a limit to how early the simulation will include each aircraft, regardless of how far ahead the system is permitted to know about aircraft. To determine whether this was the reason for the plateau we performed a second experiment, where the aircraft were added to the simulation a given number of minutes before arrival at the holding point, even if they had not left their stands at that time.

Table III shows the results of this second experiment. The results are very similar to those in table II. We conclude that the plateau in performance is not entirely due to the push-back times. We will return to the reason for this plateau later in this section.

We also wished to consider the effect of varying the point at which the take-off order is frozen. When freezing the take-off order we usually assume that the order can only be frozen for aircraft actually at the holding point. This reflects the fact that controllers can only give conditional clearances for take-off based on aircraft actually in the holding point. It also avoids a possible feasibility problem, as the holding point arrival order is not fixed for aircraft still on the taxiway, as aircraft could push back from stands in front of them. As the arrival order is not fixed, the required overtaking to achieve a schedule may

TABLE II

CTOT COMPLIANCE AND DELAY FOR VARIABLE TAXI KNOWLEDGE

Taxi knowledge(s)	Dataset 1		Dataset 2		Dataset 3	
	CTOT	Delay	CTOT	Delay	CTOT	Delay
0	5	307	2	281	4	312
30	4	294	1	255	4	307
60	4	274	2	241	4	274
90	3	266	2	237	4	262
120	3	262	2	233	4	264
180	3	259	1	223	4	257
240	3	248	1	216	4	251
300	3	249	1	217	4	252
360	3	246	1	215	4	248
420	3	246	1	213	4	244
480	3	242	1	213	4	244
540	3	242	1	213	4	244
600	3	242	1	213	4	244
660	3	242	1	213	4	244
720	3	242	1	212	4	244
780	3	242	1	212	4	244

TABLE III

CTOT COMPLIANCE AND DELAY FOR VARIABLE TAXI KNOWLEDGE, IGNORING STAND LEAVING TIME

Taxi knowledge(s)	Dataset 1		Dataset 2		Dataset 3	
	CTOT	Delay	CTOT	Delay	CTOT	Delay
0	5	307	2	281	4	312
30	4	294	1	255	4	307
60	4	274	2	241	4	274
90	3	266	2	237	4	262
120	3	262	2	232	4	264
180	3	259	1	223	4	257
240	3	248	1	216	4	251
300	3	249	1	217	4	252
360	3	246	1	214	4	248
420	3	246	1	213	4	244
480	3	241	1	213	4	244
540	3	241	1	213	4	244
600	3	241	1	213	4	244
660	3	241	1	213	4	244
720	3	240	1	211	4	243
780	3	240	1	208	4	243

cease being possible, which is a problem if that part of the schedule has been frozen.

We performed experiments that varied the length of time for which the take-off order was frozen. In these experiments we assumed that the system was aware of all aircraft on the taxiways at that time. However, the results were remarkably similar to the results when the take-off order was not frozen, so are not documented here. Investigation revealed the reason for this. The schedules produced at any stage often had aircraft that were still on the taxiways placed in early positions in the schedule, for instance to fit a taxiing aircraft between two at the holding point that need a large separation. This significantly limited how much of the schedule was actually frozen. Combined with this, as the next experiment will show, the system was often scheduling the aircraft before they reached the holding point, so fixing the schedule for aircraft at the holding point made not difference.

In the absence of taxi time uncertainty, there is an obvious relationship between the effects of varying the planning horizon and varying the amount of time the schedule is frozen for, as both vary the number of aircraft that are available for scheduling. To examine the relationship further, and to better understand the results for the previous experiments, we performed a series of experiments where we varied both the planning horizon and the time for which the take-off order was frozen. To change the planning horizon we changed the time before arrival at the holding point at which each aircraft was considered by the search. The position of an aircraft in the take-off schedule was frozen a given number of minutes before take-off, regardless of whether it was at the holding point at that time. In order to be able to freeze the schedule without introducing the aforementioned feasibility problem, aircraft were added to the simulation the given number of minutes before holding point arrival, regardless of whether they had pushed-back or not.

The results are presented for dataset 1 in tables IV, V and VI. Table IV shows the number of CTOTs missed for taxi knowledge of up to 5 minutes. Beyond 5 minutes the number of CTOTs missed is 3 in all but one case (for frozen time 540 seconds and knowledge 360 seconds it is 4) so the results are not shown here. Tables V and VI show how the total delay in the schedule changes as the taxi knowledge and frozen time vary. The results for the other datasets show similar trends.

The diagonal pattern in the tables clearly shows the expected relationship between freezing the take-off order earlier and knowing earlier about taxiing aircraft that will soon arrive at the holding point. These affect opposite ends of the planning window and, as expected, have similar effects. Exact symmetry would not be expected, even if using an exact solution method to find the best schedule, as, even though the planning window may be a similar size the problems presented to the decision support system at each iteration will involve different aircraft, and sometimes even different numbers of aircraft.

For a larger taxi time knowledge, freezing the schedule earlier does not change the delay for the aircraft. This implies that the take-off order for these aircraft had already been determined by that point. Examination of the last time at which the take-off order was changed for aircraft supports this interpretation. This explains the earlier results, where varying the frozen time for which the schedule made little effect when the system had full knowledge of the taxiing aircraft.

Additionally, the fact that aircraft are scheduled earlier also tells us that sufficient knowledge was available to be able to do this much earlier than was actually necessary. The results obtained for very large taxi knowledge match the value at which tables II and III were plateauing. This, finally, explains the plateau in tables II and III. With around eight minutes warning about aircraft that will later arrive at the holding point, the system had enough knowledge to perform the scheduling, so additional information helped very little.

So, in summary, there is a benefit from accurate knowledge of taxiing aircraft, up to around eight minutes before holding point arrival, beyond which there is little benefit to delay. But, as tables II and III are similar, there is little gain from knowing about aircraft before they are ready for push-back. Additionally, tables IV, V and VI show that fixing the schedule for longer before take-off requires more knowledge of taxiing aircraft if similar performance is to be obtained from the system.

TABLE IV

FROZEN TIME VS TAXI KNOWLEDGE - CTOTS

Frozen time(s)	Taxi knowledge for : 0s	60s	120s	180s	240s	300s
0	3	3	3	3	3	3
60	3	3	3	3	3	3
120	4	3	3	3	3	3
180	5	4	3	3	3	3
240	6	4	3	3	3	3
300	6	6	4	3	3	3
360	6	6	6	4	3	3
420	8	6	6	6	4	3
480	11	7	6	6	6	4
540	14	9	8	6	6	6

TABLE V

FROZEN TIME VS TAXI KNOWLEDGE - DELAY

Frozen time(s)	Taxi knowledge for : 0s	60s	120s	180s	240s	300s
0	284	257	255	251	247	245
60	285	262	255	251	247	246
120	287	261	260	252	246	247
180	321	278	262	259	248	249
240	345	306	273	259	252	249
300	391	337	305	276	252	252
360	409	374	332	302	272	253
420	489	421	373	332	302	272
480	520	463	412	373	331	302
540	574	513	473	410	373	331

TABLE VI
FROZEN TIME VS TAXI KNOWLEDGE - DELAY

Frozen	Taxi knowledge for :					
time(s)	360s	420s	480s	540s	600s	660s
0	244	244	241	241	241	241
60	244	244	241	241	241	241
120	246	244	241	241	241	241
180	246	246	241	241	241	241
240	247	247	243	241	241	241
300	249	247	244	243	241	241
360	252	249	250	244	243	241
420	253	257	247	250	244	243
480	270	258	254	247	250	244
540	302	270	255	254	247	250

VII. CONCLUSION

Take-off scheduling within the holding points at Heathrow is a very complicated problem. Due to the amount of time runway controllers have to spend communicating with pilots and monitoring the airport, there is limited time available for them to consider the take-off scheduling task. It is common to have little knowledge of the aircraft on the taxiways, usually only considering them if large separations are seen within the take-off schedule for aircraft already at the holding points. The result of this is that problematic sequences of aircraft on the taxiways are not always foreseen in time to do something about them.

In this paper we have evaluated the effect of changing the planning horizon, including knowledge of taxiing aircraft that will soon arrive at the holding points. To do this, we used a simulation of the departure system, with a real-time decision support system that we designed taking the place of a runway controller. We have shown that there are definite delay and CTOT compliance benefits from having knowledge of taxiing aircraft, and we feel that this is where a decision support system will be of most benefit to the controllers.

We have shown that, for a two-minute frozen take-off order, as long as the system can accurately predict the holding point arrival times of aircraft that will arrive within the next eight minutes the vast majority of the benefit can be gained. There is little need for making longer term predictions, or for predicting the arrival of aircraft that have not yet left their stands.

However, the frozen parts of the discovered schedules often contained aircraft that were still on the taxiways. These schedules would not be possible if the take-off schedule was frozen considering only aircraft in the holding points at that time, as has to be done for conditional clearances. There is, therefore, a consequent danger of missing better schedules if too many conditional clearances are given.

Finally, we have seen that there is a correlation between the time at which the take-off order is frozen and the required planning horizon. If the schedule is frozen for longer, the planning horizon will need to be moved, requiring a longer term taxi time prediction.

ACKNOWLEDGMENT

This work was supported by EPSRC (The Engineering and Physical Sciences Research Council) and NATS (National Air Traffic Services) Ltd. from a grant awarded via the Smith Institute for Industrial Mathematics and Systems Engineering.

REFERENCES

[1] H. R. Idris, B. Delcaire, I. Anagnostakis, W. D. Hall, N. Pujet, E. Feron, R. J. Hansman, J.-P. Clarke, and A. Odoni, “Identification of flow constraint and control points in departure operations at airport systems,” *AIAA Guidance, Navigation and Control conference, Boston, MA, August*, 1998.

[2] H. R. Idris, I. Anagnostakis, B. Delcaire, R. J. Hansman, J.-P. Clarke, E. Feron, and A. R. Odoni, “Observations of departure processes at logan airport to support the development of departure planning tools,” *Air Traffic Control Quarterly*, vol. 7, pp. 229–257, 1999.

[3] I. Anagnostakis and J.-P. Clarke, “Runway operations planning, a two-stage heuristic algorithm,” *AIAA Aircraft, technology, Integration and Operations Forum, Los Angeles, CA, October 1st-3rd*, 2002, available at: http://icat-server.mit.edu/Library/Download/167_paper0024.pdf [24 February 2006].

[4] I. Anagnostakis and J.-P. Clarke, “Runway operations planning, a two-stage methodology,” *36th Hawaii International Conference on System Sciences (HICSS-36), Hawaii, January 6-9, 2003*, 2003.

[5] P. van Leeuwen, H. Hesselink, and J. Rohling, “Scheduling aircraft using constraint satisfaction,” *Electronic Notes in Theoretical Computer Science*, no. 76, 2002.

[6] D. A. Trivizas, “Optimal scheduling with maximum position shift (mps) constraints: A runway scheduling application,” *Journal of Navigation*, vol. 51, pp. 250–266, 1998.

[7] J. Abela, D. Abramson, M. Krishnamoothy, A. de Silva, and G. Mills, “Computing optimal schedules for landing aircraft,” *The 12th National Conference of the Australian Society for Operations Research, Adelaide, July 7-9*, 1993, pp. 71–90, available at: http://www.csse.monash.edu.au/~davida/papers/asorpaper.pdf [24 February 2006].

[8] J. E. Beasley, M. Krishnamoorthy, Y. M. Sharaiha, and D. Abramson, “Scheduling aircraft landings - the static case,” *Transportation Science*, vol. 34, no. 2, pp. 180–197, May 2000.

[9] J. E. Beasley, J. Sonander, and P. Havelock, “Scheduling aircraft landings at london heathrow using a population heuristic,” *Journal of the Operational Research Society*, vol. 52, pp. 483–493, 2001.

[10] A. T. Ernst, M. Krishnamoorthy, and R. H. Storer, “Heuristic and exact algorithms for scheduling aircraft landings,” *Networks*, vol. 34, no. 3, pp. 229–241, 1999.

[11] L. Bianco, P. Dell’Olma, and S. Giordani, “Minimizing total completion time subject to release dates and sequence-dependent processing times,” *Annals of Operations Research*, vol. 86, pp. 393–416, 1999.

[12] A. Craig, R. Ketzscer, R. A. Leese, S. D. Noble, K. Parrott, J. Preater, R. E. Wilson, and D. A. Wood, “The sequencing of aircraft departures,” *40th European Study Group with Industry, Keele*, 2001, available at: http://www.smithinst.ac.uk/Projects/ESGI40/ESGI40-NATS/Report/AircraftSequencing.pdf [24 February 2006].

[13] J. A. D. Atkin, E. K. Burke, J. S. Greenwood, and D. Reeson, “A meta-heuristic approach to departure scheduling at London Heathrow airport,” *The 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT), San Diego, USA, August 9-11*, 2004. Available at: http://fugazi.engr.arizona.edu/caspt/atkin.pdf

[14] J. A. D. Atkin, E. K. Burke, J. S. Greenwood, and D. Reeson, “Hybrid meta-heuristics to aid runway scheduling at London Heathrow airport,” accepted for publication in *Transportation Science*, to appear 2006.

[15] F. Glover, “Future paths for integer programming and links to artificial intelligence,” *Computers and Operations Research*, vol. 13, pp. 533–549, 1986.

[16] F. Glover and M. Laguna, *Tabu Search*. Kluwer Academic Publishers, 1998.

Equitable Airport Arrival Scheduling using Airline Costs

Maarten Soomer*† and Geert Jan Franx†
*National Aerospace Laboratory NLR, Amsterdam, the Netherlands
†Vrije Universiteit, Amsterdam, the Netherlands

Abstract— **Arrival runways are a critical resource in the air traffic system. Arrival delays have a great impact on airline operations and cost. Therefore, tactical arrival planning is becoming increasingly important. It also plays a key-role in the currently researched free-flight concept.**

In this paper we consider the tactical single runway arrival problem. The current focus on collaborative decision making is reflected by giving airlines the possibility to provide cost functions related to arrival delays for their flights. A scaling method for these cost is introduced to ensure equity. Our formulation will assign landing times to the flights, while taking the costs into account. A problem-specific local search heuristic has been implemented to obtain reasonable solutions within acceptable computation times.

A large number of instances created using schedule data from a major European hub have been tested. These experiments show large cost savings for the airlines compared to current practice. All airlines achieve improvements, which is important for the acceptance of the method, and is a validation for our scaling method.

The heuristic is able to solve instances with over 100 flights in a few minutes. This makes the method very suitable to apply dynamically in a practical setting.

I. Introduction

Arrival runways are a critical resource in the air traffic system. Airport arrival capacity is subject to large changes, mainly because of weather and visibility conditions. The demand however, is largely predictable because it mainly consists of scheduled flights. During peak hours, this demand is close to the maximum capacity. This is especially the case at hub airports.

Air Traffic Control is responsible for guiding aircraft in a safe, equitable and efficient manner. Flights approaching the airport are under the guidance of the approach controller from approximately 30 minutes before the actual landing, when they enter the so called radar range of the airport. From this moment on, the controller must create a correctly separated flow of aircraft towards the runway. To maintain safety, a minimum separation between landing aircraft is required. This separation depends on the weight categories of the aircraft. There are usually 3 categories (Light, Medium and Heavy) and a heavier aircraft flying behind a lighter aircraft needs less separation, than the reverse order. Therefore, the categories of the considered flights play an important role in determining an efficient landing sequence. The minimum separation distance becomes larger in low visibility conditions (regardless of the weight categories).

Efficiency cannot be considered as the only objective. Punctuality is an important issue for airlines and their passengers. A sequence that gives the maximum runway throughput might be unbeneficial with respect to punctuality. Further, flights with a lot of (transfer-) passengers might be considered more important by the airline than other flights. Other considerations can also play a role in the valuation of an arrival delay for a specific flight by an airline, such as the transfer times of the passengers, fuel cost and the propagation of the delay. The latter can be caused by consecutive flights using the same aircraft or crew.

This leads to the idea to consider airline preferences for individual flights when determining the arrival schedule. This fits in the current focus on collaborative decision making (CDM, [1], [2]).

Arrival planning plays also an important role in the currently researched free flight concept. According to Andreatta et al. [3], the free flight concept was born among the air carriers of the United States, who were asking the air traffic control to provide them only with an arrival time, leaving the airlines the freedom of selecting, for each flight its departure time, route and speed, as long as they are able to arrive at the assigned time. Currently, hardly any tactical arrival planning (between 24 and 1 hours before landing) is done. Consequently, the flow of aircraft entering the airport radar range is not very orderly. The landing sequence of the flights will be similar to that entering the radar range, although controllers have the ability to make some small changes for reasons of safety and efficiency. In the free flight concept, the use of tactical arrival scheduling will create a more orderly arrival flow. This will make it easier for the controller to land the flights efficiently and according to the arrival schedule.

Another advantage of tactical arrival planning is, when large delays are assigned to certain flights early, these delays can be absorbed on the ground before departure rather than en-route. This is usually much cheaper for the airlines.

This paper proposes a model in which airlines provide a cost function for each arriving flight, which relates the arrival time of the flight to cost. These cost functions can be used in an ATC decision support tool in order to generate arrival schedules with minimal cost. However, it is also the role of air traffic control to assure that air traffic proceeds in an equitable manner. Thus, ATC decision support tools that take into consideration the airlines' operating costs must do so in a manner that does not show favour to some airlines at the expense of others. In Section II-A, we propose a scaling mechanism that addresses airline equity issues that stem from the modelling of delay costs for individual flights. Subsequently we will extend the mixed-integer zero-one optimization model proposed by Beasley et. al. [4] by using these scaled cost functions as objective. This model formulation, is given in section II-B. A local-search heuristic to efficiently obtain reasonable schedules is introduced in section III.

A large number of problem instances, created from real-life data concerning arrivals during a week at a major European hub, are used as input for our method and the results are shown in section IV. These results include an analysis of the costs and fairness of the obtained schedules as well as the performance of the heuristic (both the quality of the solutions and computation times).

Our formulation is deterministic, meaning all parameter values are assumed to be known, before running the algorithm. In practice this is not true. Departure delays and unexpected weather changes, will influence possible landing intervals and required separation times. However, our formulation and heuristic allow for a dynamic use of the model by recalculating the schedule, every time the circumstances change.

II. Model

The model can be used to determine an arrival sequence and arrival times for a given set of flights on a single runway, complying to the separation rules. This means that we assume flights are already assigned to a runway, if more runways are currently used for arrivals. We also assume that runways are independent, meaning that the flight paths of all runways in use, are separated sufficiently, such that the runways can be treated separately.

It is assumed that during the planning period the runway considered will only be used for arriving flights. The latter is not a real restriction, because, due to safety and cost considerations, landing operations take precedence over take-offs in mixed mode. Further, we assume that there is a closed time interval for each flight in which it will land. For flights en route, this interval is determined by the distance left, maximum and minimum speed and amount of fuel left. For not yet departed flights, the minimum flight time and maximum departure delay determine the interval. For reasons of equity it is advisable to use the same maximum departure delay for all flights.

Next, the representation of airline costs and the scaling mechanism is introduced. The mixed-integer zero-one optimization formulation is given in section II-B.

A. Airline Costs and Equity

The cost incurred by an airline for an arriving flight depends on its landing time. Every flight might have different characteristics, such as the number of transfer passengers and their transfer times and the departure time of a subsequent flight with the same aircraft or crew. Therefore the airline is allowed to provide a different cost function for every individual flight. Each function relates the arrival time of the flight to cost.

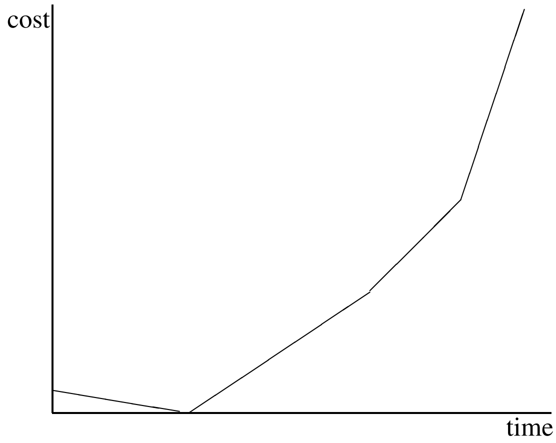


Fig. 1. Example of a convex piecewise linear cost function

We want to allow the airlines as much flexibility as possible in determining the cost functions. At the same time the functions need to be suitable to determine an equitable arrival schedule. We propose a restriction on the shape of the functions and the application of a scaling mechanism to obtain this.

The functions are required to be convex and piecewise linear and to have a minimal cost of zero at a time within the possible landing interval. Convex functions in general seem to be a realistic description of the costs: A larger delay means a larger increase in costs per time unit. Any convex function can be accurately approximated by a (convex) piecewise linear function. These are convenient to use in a mathematical programming formulation, as given in section section II-B. An example of such a function is depicted in figure 1. The airlines still have a lot of flexibility in determining the number and locations of the breakpoints and the slopes of the lines. The breakpoints should represent times at which cost will make a big step, or after which cost will increase much faster than before. These breakpoints will behave as thresholds.

The considered costs can vary a lot between different airlines. Now, minimising the sum of all cost functions, will favour airlines that define relatively large cost for their flight delays. At first sight this seems reasonable, since this will lead to minimal total cost. However, we have to bear in mind that all airlines are allowed to define their own cost functions. Therefore, they will be able to obtain higher priorities for their flights, by (falsely) representing very large cost for delays. This leaves room for manipulation. This troublesome aspect can be overcome by rescaling all cost functions, such that the average assigned cost per time unit are equal. However, the cost-ratio between flights of a single airline should be preserved, to reflect the economic trade-off for this airline. Therefore, the same scaling factor will apply to all flights of the same airline. The individual cost functions of the flights are multiplied by the scaling factor of their airline. An optimal schedule is not defined as the schedule with minimal total cost, but as the schedule with minimal total *scaled* cost.

Let us make this more precise. Consider airline j with N_j arriving flights, with convex piecewise linear cost functions $\kappa_1(t), \dots, \kappa_{N_j}(t)$. Let E_i and L_i be the earliest and latest possible landing times of aircraft i, respectively. So, $\kappa_i(t)$ is defined on the interval $[E_i, L_i]$.

To obtain equity, these cost functions will be scaled to new cost functions $f_i(t) = \alpha_j \kappa_i(t)$ $(i = 1, \dots, N_j)$. The scaling factors α_j are determined per airline. This ensures that the ratio between costs of their own flights are preserved in the scaled objective functions. α_j is defined such that:

$$\frac{1}{N_j} \sum_{i=1}^{N_j} \frac{\int_{E_i}^{L_i} \alpha_j \kappa_i(t) dt}{(L_i - E_i)^p} = 1.$$

So,

$$\alpha_j = N_j \left(\sum_{i=1}^{N_j} \frac{\int_{E_i}^{L_i} \kappa_i(t) dt}{(L_i - E_i)^p} \right)^{-1},$$

where p is a parameter to minimise the effect of differences in the length of the landing intervals. It is preferable to choose $p \geq 2$.

Let us explain this. Consider two airlines with only one flight and identical cost functions $\kappa_1(t) = \kappa_2(t) = t$. The possible landing interval of flight 1 and 2 are $[0, T_1]$ and $[0, T_2]$ respectively, with $T_2 > T_1$. Now

$$\begin{aligned} \alpha_1 &= \quad 2T_1^{(p-2)} \\ \alpha_2 &= \quad 2T_2^{(p-2)} \end{aligned}$$

and the scaled objective functions are:

$$\begin{aligned} f_1(t) &= \quad 2T_1^{(p-2)} t \\ f_2(t) &= \quad 2T_2^{(p-2)} t. \end{aligned}$$

If $p < 2$ then $f_1(t) \geq f_2(t)$ for $0 \leq t \leq T_1$. This cannot be considered fair: Flight 2 has a larger interval and therefore more possible landing times. If it would be impossible to land both aircraft before time T_1, aircraft 2 is able to land between time T_1 and T_2, while aircraft 1 is not. When choosing $p < 2$ it is always cheaper to land aircraft 1 before aircraft 2 (even if both aircraft can be scheduled before time T_1) and airline 2 costs ($\kappa_2(t_2^*)$) will be larger than airline 1 costs ($\kappa_1(t_1^*)$). So when $p < 2$, airline 1 is better off, while airline 2 provides more flexibility.

If $p = 2$ then $f_1(t) = f_2(t)$ for $0 \leq t \leq T_1$, which can be considered fair. Another choice is to reward aircraft 2 for providing more flexibility (by a larger interval) by choosing $p > 2$.

At most airports, there are peak periods during the day. If the α_j's were determined for a day or longer, an airline with flights scheduled in peak and non-peak periods has an advantage over an airline with flights only scheduled in peak periods. The first airline could assign large cost to delays for their flights during peak periods, compared to their flights outside peak periods. The average scaled delay cost per time unit will be much larger for the first airline compared to the second for their respective flights in the peak periods. Therefore it is cheaper to assign less delay to flights from airline 1 in peak periods. Delays are much more likely to happen in peak periods, because in these periods the demand is close to (or even temporarily exceeds) capacity. Outside peak periods, delays do not occur frequently, so flights from airline 1 are expected to receive little delay, again. While for airline 2 it is relatively cheap to assign delays to in the peak periods, where all their flights are. More fairness can be achieved by determining new scaling factors for separate time periods (such as peak and non-peak periods).

Our scaling approach also offers the possibilities of trading of preference levels between airlines. Consider airlines j and k, having scaling factors α_j and α_k, respectively. Now airline j could pay airline k such that the new scaling factor of airline k would be $0.9\alpha_k$ and airline j would receive a scaling factor $(1 + 0.1\frac{N_k}{N_j})\alpha_j$.

B. MIP Formulation

In this section a Mixed Integer Programming (MIP) formulation of the model is given. The basic notation and constraints are similar to those of Beasley et al. [4]. The objective function represents the total scaled costs.

Let

N : Number of arriving flights to schedule

E_i : Earliest possible landing time for flight i $\quad i = 1, \dots, N$

L_i : Latest possible landing time for flight i $\quad i = 1, \dots, N$

S_{ij} : Required separation time when flight i lands before flight j $\quad i, j = 1, \dots, N; i \neq j$

and decision variables

t_i : landing time for flight i $\quad i = 1, \dots, N$

$$\delta_{ij} = \begin{cases} 1 & \text{if flight } i \text{ lands earlier than flight } j \\ 0 & \text{otherwise} \end{cases} \qquad i, j = 1, \dots N; i \neq j$$

We will now introduce the basic constraints, ensuring a feasible and safe schedule. A more extensive description of those can be found in Beasley et al. [4].

$$E_i \leq t_i \leq L_i \qquad i = 1, \dots, N \tag{1}$$

This defines the possible landing time.

Considering pairs of flights (i, j):

$$\delta_{ij} + \delta_{ji} = 1 \qquad i = 1, \dots, N-1; j > i \tag{2}$$

This set of constraints states that either flight i must land before flight j ($\delta_{ij} = 1$) or flight j must land before flight i ($\delta_{ji} = 1$).

Using the separation times and (overlap of) possible landing time intervals we can define three sets of pairs of flights:

U : the set of pairs (i, j) of flights for which it is uncertain whether flight i lands before flight j

V : the set of pairs (i, j) of flights for which flight i definitely lands before flight j, but for which the separation is not automatically satisfied

W : the set of pairs (i, j) of flights for which aircraft i definitely lands before flight j, and the separation is automatically satisfied

For elements in the set V and W the order of the flights is known. So trivially the following constraints hold:

$$\delta_{ij} = 1 \qquad \forall (i, j) \in W \cup V \tag{3}$$

Note that for flights in set V we still need to ensure the proper separation:

$$t_j \geq t_i + S_{ij} \qquad \forall (i, j) \in V \tag{4}$$

Finally we also need to ensure separation if the order of the flights is not known (pairs of flights in set U):

$$t_j \geq t_i + S_{ij}\delta_{ij} - (L_i - E_j)\delta_{ji} \qquad \forall (i, j) \in U \tag{5}$$

Either, flight j cannot land earlier than the minimum separation time between i and j after the landing time of flight i:

$$t_j \geq t_i + S_{ij},$$

in case flight i lands before flight j ($\delta_{ij} = 1$) or this constraint is superfluous, since

$$t_j \geq t_i - (L_i - E_j)$$

is always satisfied (because $t_i \leq L_i$ and $t_j \geq E_j$), in the case flight j lands before flight i ($\delta_{ij} = 0$). Note that the separation in the latter case is ensured by the constraint for the pair (j, i).

As explained in section II-A each airline provides a convex piecewise linear cost function for each of its flights, that is then scaled to obtain equity among airlines. The scaled convex piecewise linear function $f_i(x)$ for flight i can be written as a set of linear functions on a number of connected intervals:

$$f_i(x) = \begin{cases} A_{i0}x + B_{i0} & 0 \leq x \leq X_{i1} \\ A_{i1}x + B_{i1} & X_{i1} \leq x \leq X_{i2} \\ \vdots & \vdots \\ A_{iK_i}x + B_{iK_i} & X_{iK_i} \leq x \end{cases}$$

where

K_i : Number of breakpoints of $f_i(x)$ $\quad i = 1, \ldots, N$

Now:

$$f_i(x) = \max_{k=0,\ldots,K_i} \{A_{ik}x + B_{ik}\}$$

This can be seen using that $f_i(x)$ is continuous, meaning

$$f_i(X_{i,k}) = A_{i,k-1}X_{i,k} + B_{i,k-1} = A_{i,k}X_{i,k} + B_{i,k} \qquad k = 1, \ldots, K_i$$

and $f_i(x)$ is convex, meaning

$$A_{i0} < A_{i1} < \ldots < A_{iK_i}.$$

The objective of the MIP will be to minimise the sum of these cost functions. However, these function are not linear in the current decision variables t_i, and therefore new decision variables c_i are introduced:

c_i : cost for landing flight i $\quad i = 1, \ldots, N$

The c_i will represent the cost function $f_i(t_i)$. To ensure this the following constraints are introduced:

$$c_i \geq A_{ik}t_i + B_{ik} \qquad i = 1, \ldots, N; k = 0, \ldots, K_i \tag{6}$$

These ensure for flight i that

$$c_i \geq \max_{k=0,\ldots,K_i} \{A_{ik}t_i + B_{ik}\} = f_i(t_i).$$

Next we will introduce our objective function:

$$z = \min \sum_{i=1}^{N} c_i \tag{7}$$

Equations 6 and 7 ensure that $c_i = f_i(t_i)$:

Suppose $t_1^*, t_2^*, \ldots, t_n^*, c_1^*, c_2^*, \ldots, c_n^*$, is an optimal solution of our MIP, where: $c_i^* > f_i(t_i^*)$ for some i. Let $c_i' = f_i(t_i^*)$. Replacing c_i^* with c_i' will decrease the objective by $c_i^* - f_i(t_i^*)$ without violating any of the constraints 6:

$$c_i' = f_i(t_i^*) = \max_{k'=0,\ldots,K_i} \{A_{ik'}t_i^* + B_{ik'}\} \geq A_{ik}t_i^* + B_{ik}$$

for $k = 0, \ldots, K_i$. So $c_i^* > f_i(t_i^*)$ cannot be optimal (and $c_i' = f_i(t_i^*)$ is).

The use of non-convex (piecewise linear) cost functions is possible, but will introduce additional decision variables to this formulation.

III. Local Search Heuristic

Beasley et. al. [4] show that the considered decision problem is NP-complete. This is done by reduction from a job shop scheduling problem with sequence dependent-setup times and zero processing times (or zero setup times and sequence dependent processing times). This means no efficient (polynomial) solution algorithm is known. Solving the MIP-formulation will provide an optimal solution, but is not efficient.

Therefore, we want a heuristic that provides reasonable solutions very fast. The idea behind our heuristic is to repeatedly find an improved sequence for the flights and determine the optimal landing times given this sequence.

This idea stems from the following observation: If the landing sequence of the flights is given, the MIP-formulation becomes a LP formulation (since the values of all the binary variables are known). This formulation consists of (1), (4), (6) and (7). The solution of this LP provides the optimal landing times, given the sequence.

To find an improved sequence of the flights, local search can be used. Local search uses a neighbourhood of the current solution to find a new (improved) solution. The neighbourhood is defined in such a way that new sequences will be “close” to the current sequence, meaning they are very similar. This means the corresponding LP formulations will also be. LP solvers are able to solve such a formulation very efficiently by using the previous solution.

The general local search algorithm is given below.

LOCAL SEARCH()
$S =$ initial feasible solution
while there is a neighbour of S of better quality
  do $S =$ neighbour of S of better quality

Next we will specify how to find an initial feasible solution, the definition of the neighbourhood and the selection procedure for a neighbour of better quality. There are standard techniques available to do this. However, it is beneficial to use problem specific features in these procedures.

A. Initial Feasible Solution

An initial sequence is found by sorting the flights according to their preferred landing time (time with minimum cost). Solving the LP gives optimal feasible landing times for this sequence.

It is, however, possible that the LP is not feasible, meaning there are no feasible landing times given this sequence. In that case new sequences are repeatedly generated, by swapping two adjacent flights for which the earlier one has a larger latest landing time, until the LP of such a sequence is feasible. If this still not gives a feasible solution, new sequences are repeatedly generated, by swapping two adjacent flights, for which the total sequence require less separation.

B. Neighbourhoods

We used two neighbourhoods.

The first is a problem-specific extension of a swap-neighbourhood. The swap neighbourhood consists of all sequences that are equal except that two flights have swapped positions.

The second is a problem-specific extension of a shift-neighbourhood. The shift neighbourhood consists of all sequences that are equal except that one flight is removed from its original position and inserted at a new position.

The problem-specific extensions to these neighbourhoods convert an infeasible sequence into a feasible one. This can be done if the infeasibility is caused by non-overlapping landing time intervals of a flight that is swapped or shifted and some other flights. By changing the positions of these other flights, a feasible sequence can be obtained.

C. Selection of a neighbour

Since the evaluation of a neighbour involves solving a LP, which takes (relatively) long, it is preferable to find a selection method that finds an improvement by evaluating as little neighbours as possible.

Therefore we will evaluate promising sequences first. For all neighbours of the current solution, an estimated gain in objective is calculated. This estimation uses the landing times in the current solution, to estimate the landing times and involved scaled cost in the new solution, without solving the LP. The neighbour with the largest estimated gain is evaluated first. If this neighbour indeed gives a better solution, it is selected. Otherwise the neighbour with the second largest estimated gain is evaluated, etc.

IV. Computational Results

In this section, we asses the performance of the local search heuristic. A large number of instances, created using schedule data from a major European hub, were tested. This data contains all arrivals from a week in September 2004. The data included, airline, flight number, aircraft type, arrival runway and scheduled and actual arrival times.

The following assumptions were made about the cost functions and possible landing times of the flights.

It is assumed that the cost function of every flight has a minimum of zero cost at the scheduled time of arrival of the flight, according to the timetable.

The structure of the cost functions for the home carrier and its partners, was determined in cooperation with specialists from this airline. Its perceived delay costs are strongly related to the number of missed transfers. This is quite natural, since this airline uses the airport as hub, and consequently has a lot of passengers transferring at the airport. Exact passenger flows and related costs were not provided by the airline for confidentially reasons. Instead the number of missed transfers per 15 minutes of delay, were generated randomly. These were translated into convex piecewise linear cost functions by using the cumulative number of missed transfers up to time t as the slope of the cost function at time t.

The main objective for the other airlines is punctuality. The latest landing time before other flights from the airline are affected by the delay of the flight considered, is also taken into account. This is represented in the cost function by choosing a positive slope between the scheduled time and this time and a steeper slope hereafter. This point in time and the exact slopes are generated randomly for each flight.

Our planning horizon is several hours before the flights will land. That means that flights departing from Europe have not departed when planning. These flights are assumed to have a maximum departure delay and a possible landing interval of 3 hours. Intercontinental flights are en-route and are able to arrive between 25 minutes before and 30 minutes after schedule.

TABLE I
WAKE VORTEX SEPARATION IN NAUTICAL MILES FOR DIFFERENT WEIGHT CATEGORIES

		following aircraft		
		Light	Medium	Heavy
leading	Light	3	3	3
aircraft	Medium	5	3	3
	Heavy	6	5	4

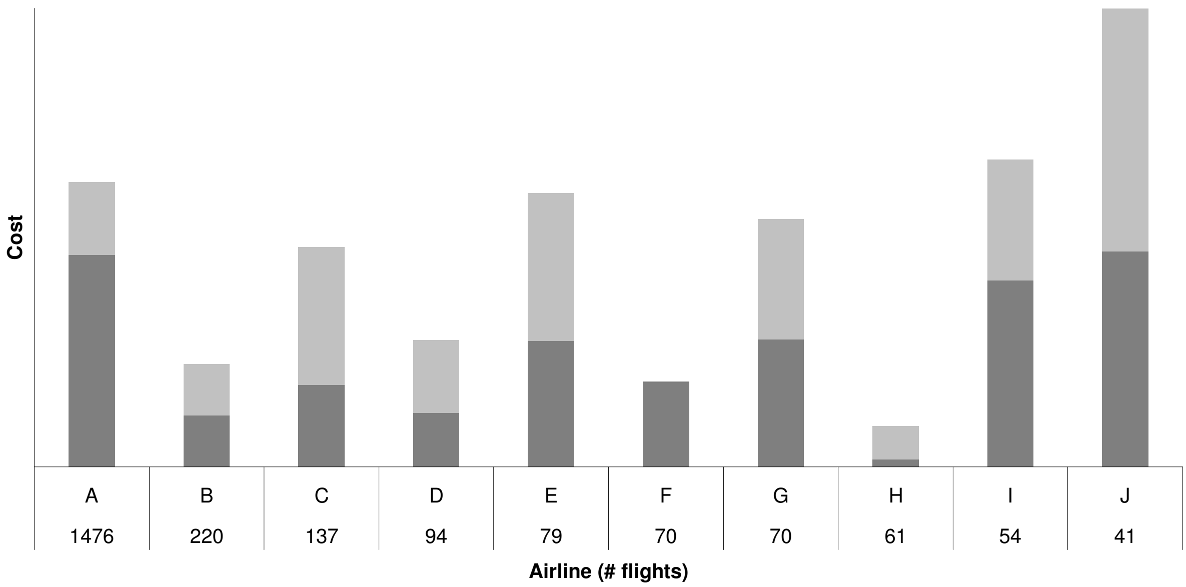


Fig. 2. Average cost per flight in FCFS schedule (total column) and the local search schedule (dark gray)

In Table I the arrival separation distances under good visibility conditions, according to international regulations, are listed.

In low-visibility conditions, the separation distances must be larger. The required separation distance under low-visibility conditions is the maximum of the required wake-vortex and low-visibility separations. The actual required separation *time* between two flights is calculated using the weight categories and approach speeds of the aircraft. The aircraft types used for the flights are available from the data.

In the experiments, low-visibility conditions requiring 6 nautical miles separation were used. This separation causes a decreased arrival capacity which will often result in large delays. It is interesting to asses the (scaled) costs and their spread among the airlines, resulting from those delays. All optimised results are compared to the initial feasible solution used in the local search, which resembles the first come, first served (FCFS) sequence currently used in practice.

96 instances were created by dividing the 4064 arrivals by runway and time. A runways is only used continuous for at most a few hours, depending on demand and weather conditions. The flights landing in such a period on a runway, are considered as a single instance. These instances contained between 1 and 501 flights. All the instances were solved using the local search heuristic with the adapted swap neighbourhood. The heuristic was implemented in C++. The LP problems are solved using ILOG CPLEX 7.5.

The schedules obtained by the local search algorithm, yields tremendous savings compared to the FCFS costs. The overall average savings per flight were 33% of FCFS costs (47% in terms of scaled costs).

To evaluate the equity of the schedule, the average costs per flight for the airlines are examined. Because the costs for an individual flight can fluctuate a lot, it is only useful to compare airlines with a substantial number of flights. In figure 2 the average cost per flight for the 10 airlines with the largest number of flights, in the week considered, are shown. The names of the airlines are omitted for reasons of confidentiality. For these airlines, the savings are 47% of the FCFS costs. The savings in terms of scaled costs are on average 52% per airline. All these airlines achieve improvements, which is important for the acceptance of the method, and is a validation for our scaling method.

The (FCFS) costs of the airlines are partly determined by the original timetable. Flights in peak periods are likely to receive a delay and the corresponding costs. This indicates it might be useful to consider airline-costs issues also when designing the timetable. This is, however, beyond the scope of this paper.

Another experiment has been done to asses the quality of the solutions obtained by the local search heuristic. They were compared to the optimal solutions, obtained by CPLEX's MIP-solver. This was done for 24 instances, that could be solved by the MIP solver in reasonable time. One of these instances contained 117 flights, the others between 15 and 51 flights each. 19 instances were solved, assuming good visibility conditions. 5 instances were solved under the low-visibility conditions used in the previous experiment.

The total optimal scaled costs were 19% of the total scaled FCFS costs (81 % savings obtained) for these 24 instances. The heuristic using the swap and shift neighbourhood, were able to achieve 94% and 98% of the total possible savings and gave the optimal solution in 5 and 13 instances, respectively. In the 5 instances with low-visibility conditions the swap and shift neighbourhood achieved 90% and 85% of the total possible savings. This shows the heuristic provides solutions, which obtain a large amount of the maximum possible savings. Note that the heuristic always provides a solution that is

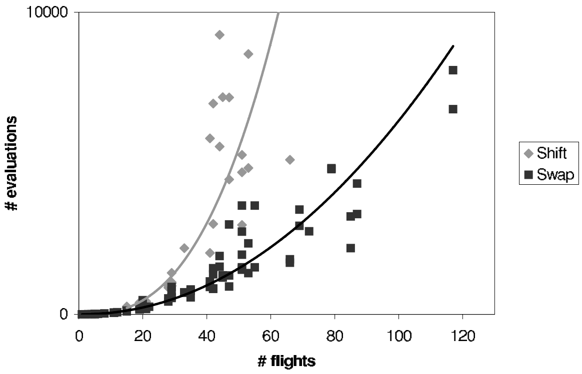


Fig. 3. Number of neighbourhood evaluations required using the shift and swap neighbourhoods

as least as good as the FCFS solution, since this is the initial solution used. The shift neighbourhood seems to give better solutions under good visibility conditions and the swap neighbourhood under low-visibility conditions.

The computation time of the heuristic, is determined by the number of neighbour LP evaluations. In each iteration the neighbours of the current solution are evaluated until a better solution is found or all neighbours are evaluated. In the latter case the heuristic terminates. In figure 3 the relation between the number of evaluations and the number of flights in the instance is shown. It is clear that the swap neighbourhood requires less evaluations. This can be explained by the number of neighbours for a arbitrary sequence of N flights. The swap and shift neighbourhood contain $N(N-1)$ and $\frac{N(N-1)}{2}$ neighbours respectively. The shift neighbourhood requires also more iterations on average before the solution which has no improving neighbours, is found. An evaluation takes approximately the same time for both neighbourhoods, because a similar LP-formulation of the same size has to be solved. The time required to solve this LP increases somewhat with the number of flights in the instance.

The total running time of the heuristic, when using the swap neighbourhood, is within a minute for instances with up to 40 flights and this increases to about ten minutes for instances with around 80 flights. All computations were performed on a Compaq computer with an Intel Pentium III processor (866 MHz), 256 MB physical memory and a Linux operating system.

In a practical setting a schedule might be needed fast. When running the heuristic, the best found solution so far is available at any time.

The computation times can be reduced by evaluating only neighbours that are promising (considered by their estimated gain). Another possibility is to limit the size of the neighbourhoods by allowing only swaps of flights that are less than a certain number of positions apart in the current sequence and shift a flight less than a certain number of positions. These changes might however affect the quality of the solutions.

V. Conclusion and Further Research

In this paper we considered the single runway arrival problem. A model to determine an arrival schedule, that takes airlines cost into account, was presented. A scaling method for these cost was introduced to ensure equity. A local search heuristic was implemented to obtain reasonable solutions fast.

A large number of instances created using schedule data from a major European hub were tested. These experiments show large cost savings for the airlines compared to a schedule that resembles current practice. All airlines achieve improvements, which is important for the acceptance of the method, and is a validation for our scaling method.

The heuristic is able to solve instances with over 100 flights in a few minutes. This makes the method very suitable to apply dynamically in a practical setting. In a dynamic setting, available information regarding visibility conditions and possible landing intervals keeps changing over time. The heuristic is then used several times, for the same flights, with the updated information.

The model can be extended to a multiple runway model, including runway assignment. Integral scheduling of arriving and departing flights is also interesting, because their arrival and departure times and cost are related.

Acknowledgment

We thank the airline specialists for the fruitful discussions. Also we would like to thank Ger Koole from the Vrije Universiteit, Amsterdam and Rene Verbeek, Michel van Eenige and Ronny Groothuizen from the National Aerospace Laboratory NLR for their input.

References

[1] Eurocontrol, "European cdm portal," `http://www.euro-cdm.org`. [Online]. Available: http://www.euro-cdm.org

[2] E. Gilbo and K. Howard, "Collaborative optimization of airport arrival and departure traffic flow management strategies for cdm," in *3rd USA/EUROPE Air Traffic Management R&D Seminar*, 2000. [Online]. Available: http://atm-seminar-2000.eurocontrol.fr/acceptedpapers/pdf/paper18.pdf

[3] G. Andreatta, L. Brunetta, and G. Guastalla, "From ground holding to free flight: An exact approach," *Transportation Science*, vol. 34, no. 4, pp. 394–401, 2000.

[4] J. Beasley, M. Krishnamoorthy, Y. Sharaiha, and D. Abramson, "Scheduling aircraft landings - the static case," *Transportation Science*, vol. 34, no. 2, pp. 180–197, 2000.

Statistics of the Approach Process at Detroit Metropolitan Wayne County Airport

Babak G. Jeddi, John F. Shortle, Lance Sherry

Abstract—**Dealing with uncertainty is a necessary and difficult aspect in operations analysis of complex systems such as the air transportation system. We represent these uncertainties in the approach process by probability distributions. This paper provides statistical observations of the approach and landing process at Detroit Metropolitan airport (DTW) for one week of February 2003. These probability distributions are based on aircraft track record data at DTW which are collected by a multilateration surveillance system. After explaining characteristics of the database and its short comings, we present a methodology to extract necessary statistical samples. From this, we obtain appropriate probability distributions for landing time interval (LTI), inter arrival distance (IAD), and runway occupancy time (ROT) presented under instrument flight rules (IFR) and peak traffic periods.**

Index Terms—**Aircraft approach, probability, risk, safety, stochastic processes**

I. INTRODUCTION

UNDERSTANDING the stochastic behavior of the approach and landing process is critical to analyze runway separation risk and runway capacity. Statistical analysis is a method for this purpose. In recent years, multilateration systems have been installed in some airports, including Detroit Metropolitan Wayne County airport (DTW). These systems provide reasonably accurate time-position estimates of all transponder-equipped aircraft (a/c) operating in the airport vicinity in all weather conditions. These data can be used to obtain samples of landing process variables, such as the Landing Time Interval (LTI) between successive aircraft to the runway threshold, the Inter-Arrival Distance (IAD) between two successive aircraft at the moment that the lead aircraft crosses the runway threshold, and Runway Occupancy Times (ROT). ROT is the length of time required for an arriving aircraft to proceed from over the runway threshold to a point clear of the runway. This paper considers *LTI*, *IAD*, and *ROT* as (random) variables.

Manuscript received February 26, 2006. This work is supported in part by Wayne Bryant, Wake Program Manager, NASA Langley Research Center.

Babak G. Jeddi is a Ph.D. candidate in the Department of Systems Engineering and Operations Research at George Mason University, Fairfax, VA 22030 US (Corresponding author, Phone: +1 (703) 505-5994; fax: +1 (703) 993-1521; e-mail: BGhalebs@gmu.edu).

John F. Shortle is an assistant professor in the Department of Systems Engineering and Operations Research at George Mason University, Fairfax, VA 22030 US. (e-mail: JShortle@gmu.edu).

Lance Sherry is an associate research professor in the Department of Systems Engineering and Operations Research at George Mason University, Fairfax, VA 22030 US. (e-mail: LSherry@gmu.edu).

Levy *et al*. [1] use multilateration data of Memphis International airport (MEM) to obtain probability distributions of *LTI* and average landing speed conditioned on the type of follow-lead aircraft in visual meteorological condition (VMC).

Probability distributions for *LTI* and *ROT* are also estimated by Haynie [2] for Atlanta International airport (ATL) using his field observations from this airport. References [3]-[5] provide distribution fits for Haynie's observations, as well as field observations from LaGuardia airport (LGA). However, the sample sizes are small, and the results are not conditioned on aircraft weight class type, or heavy traffic times. Also, they have not obtained samples of *IAD*, which we provide here. Reference [6] provides analysis of *IAD*, *LTI*, and runway utilization in peak periods at Dallas/Fort Worth International Airport (DFW). They include both instrument meteorological condition (IMC) and VMC times in their study; however, they do not fit known specific probability distributions to the observations. They use radar data, which normally do not extend to the runway threshold, so they must extrapolate aircraft flight paths to the most likely runway threshold.

Vandevenne and Lippert [7] develop a model to represent *LTI* and provide a probability distribution fit. This model is the convolution of exponential and normal distributions. Andrews and Robinson [8] extend the capabilities used in [6]. They fit probability distribution functions for *LTI* using the Vandevenne and Lippert model [7]. Rakas and Yin [9] use Performance Data Analysis and Reporting System (PDARS) database to estimate probability distribution of *LTI* under VMC in Los Angeles International Airport (LAX). For this purpose, they develop a PDF which they name it *double-normal* distribution.

The Center for Air Transportation Systems Research (CATSR) at George Mason University (GMU) has access to multilateration surveillance system data of DTW via Volpe National Transportation Systems Center, an organization within the US Department of Transportation. The original multilateration data are de-identified by Sensis Corporation, and the filtered data are used in this study. However, as discussed later, there are still some outliers, noise, and missing data present in the database.

This paper introduces the characteristics and organization of available data and the algorithms that we have initiated to extract recorded data of landing times and position over the runway thresholds, runway exit times, and the position of the following aircraft when its lead crosses the runway threshold. Then this algorithm is used to investigate DTW data from Feb2, 2003 to Feb8, 2003 (in Greenwich Mean Time) in order to provide probability distributions for *LTI*, *IAD*, at runway thresholds, and *ROT*. Fig. 1 is a simplified diagram of this airport.

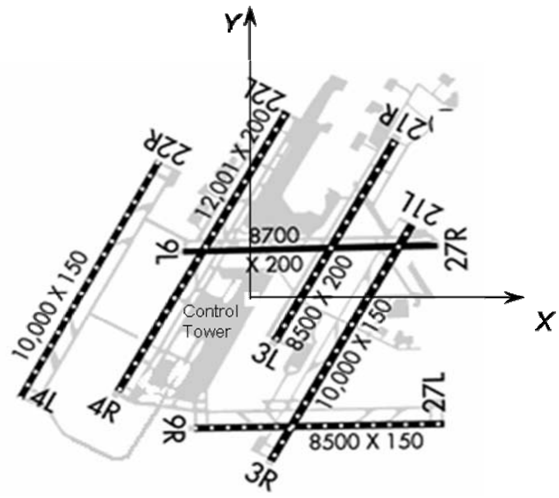


Fig. 1. Simplified DTW airport diagram (http://www.airnav.com/airport/KDTW)

We have organized this paper as follows. Section 2 provides details about the database structure, noise and outliers, and data preparation necessary to extract the required landing samples. Section 3 presents statistical findings and probability distribution fits for "peak traffic period landing variables" *LTI*, *IAD*, and *ROT*, under IMC. (In this paper, IMC is defined by the IMC / VMC flag in the ASPM database, which provides conditions at the airport.) Section 4 presents conclusions of the study and some topics for future research.

II. Database Structure and Sample Extraction Procedure

Multilateration data must be processed to perform probabilistic analysis of the operations. There are two categories of short comings with the data. First, the data contain noise, outliers, and missing records, and second, the data provide the aircraft time-position tracks but do not specify when aircraft cross certain positions. In this section, we discuss these problems and explain our strategy to extract necessary samples for statistical analysis of the approach process.

We make use of five fields from the multilateration data (out of a possible eighteen): aircraft mode-s, time (t in seconds), longitude (X in meters), latitude (Y in meters), and mode-c. The mode-s field is a number of an attached transponder that uniquely identifies an aircraft. The transponder is generally attached somewhere close to the center of the aircraft. The mode-c field is a barometer-based value that can be converted to altitude (in feet) by multiplying it by 25 and adding 10,000 to the result. However, the obtained value is not very reliable for this purpose due to pressure change and barometer errors under for different weather conditions. Time and position of aircraft are recorded every second. For the week Feb 2, 2003 to Feb 8, 2003, the database includes 33,030,878 records, requiring 1GB of disk space.

A. Data preparation

The database is in Oracle format and we use SQL+ to obtain queries. Necessary manipulations and sample extractions are done in MATLAB.

To start, we sort the Oracle data by mode-s and then by time. We also change the time stamp to the format "dd/mm/yy hh:mi:ss." Basic queries demonstrated that the mode-s is missing for some records. In some cases, the mode-s of an entire aircraft track is missing. In other cases, we are missing the mode-s of only a few points along a track. We eliminate all of these data points. In the latter case, we retain the basic track path, since we can linearly interpolate the path of the aircraft from the other points with mode-s. In the former case, we discard the entire track. This may result in some inter-arrival times that are too long. However, because of the available data, loosing some possible landing records does not significantly influence the study.

In the database, the origin (X=0,Y=0) of the Euclidian coordinate is the FAA control tower located between runways 21R and 22L, as shown in Fig. 1, and the Y axis indicates the true north. Runway 21L, and all other runways parallel to it, have a Magnetic angle of 214.8^o. True North and Magnetic North have an angle of 6.1^o W, as indicated in airport diagram [10]. Thus, the true angle of runway 21L is 214.8 – 6.1 = 208.7^o, or equivalently 61.3^o from the X-axis. Since data are collected in the true coordinates, we observe the same results by tracking the aircraft course on the runways [11]. In the same manner, we calculate the true angle of runways 27L/09R and 27R/09L as 1.3^o from the X-axis.

To simplify working with the database, we rotate coordinates to make the runways parallel to the X-axis. To find the aircraft position in the rotated coordinates we multiply the observed (X,Y) position by the rotation matrix R as

$$R = \begin{pmatrix} Cos\,(a) & Sin\,(a) \\ -\,Sin\,(a) & Cos\,(a) \end{pmatrix}, \tag{1}$$

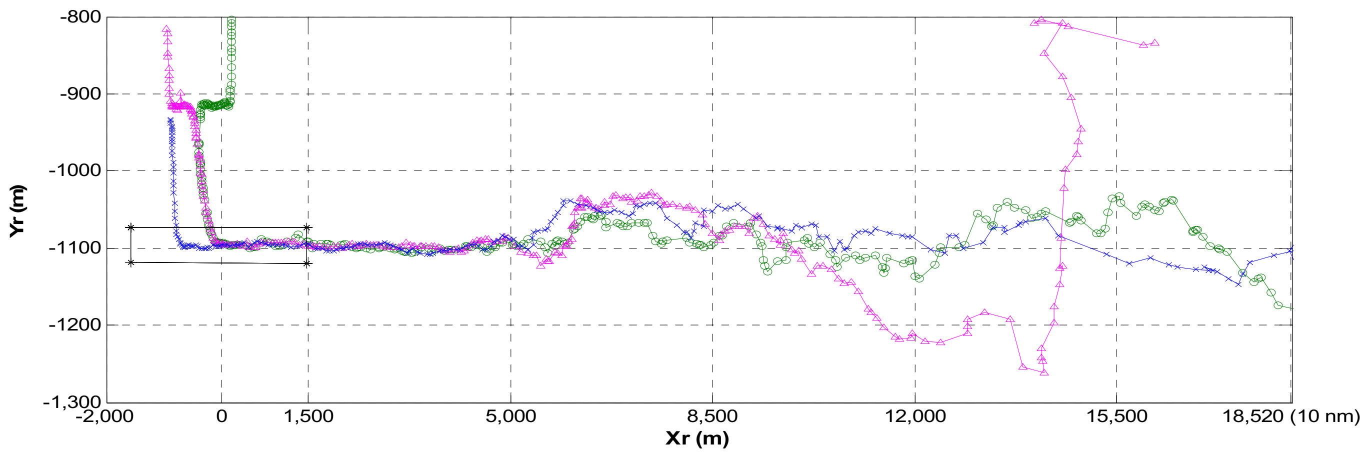
where a is the rotation angle which is 61.3^o for runways 21L/03R, 21R/03L, 22L/04R and 22R/04L, and 1.3^o for runways 27L/09R and 27R/09L as described before. That is, the aircraft position in the rotated Euclidian coordinates is

$$\begin{pmatrix} X_r \\ Y_r \end{pmatrix} = R * \begin{pmatrix} X \\ Y \end{pmatrix}. \tag{2}$$

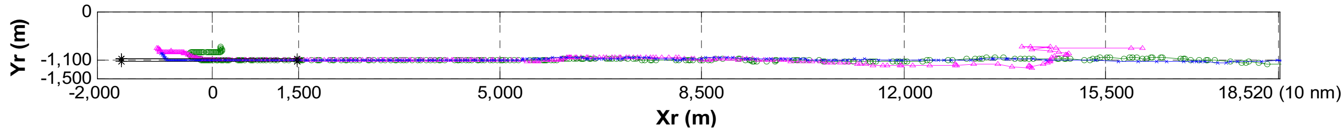
Using the rotation formula (2), we also transform the runway coordinates to the new coordinates.

Preliminary queries and plots demonstrated some noise in the data. Fig. 2 is the ground projection (bird's eye view) of the track plot of sample aircraft landings on runway 21L. Fig. 2a (the lower figure) is drawn to scale, whereas Fig. 2b is expanded in the Y_r-axis. In the figure, two aircraft exit the runway from the high-speed exit located after the middle of the runway. Based on visual investigation, the noise of X-Y positions is assumed to be in an acceptable range, for a given time, as demonstrated in Fig. 2b. In frequent cases, there are two or more records of a given aircraft at the same second. We average the records in such cases.

Since landings are the subject of study, it is sufficient to



a) Exaggerated in Yr axis



a) Drawn to scale

Fig. 2 (a) Landing and exit track of three large a/c on/from runway 21L, (b) exaggerated version of figure (a) in Yr axis

consider data in a rectangle, which we call the "query box." The sides of the query box are parallel to the sides of the runway rectangle, and the box includes the runway and the common landing path extended about 10 nm from the runway threshold. For runway 21L, for example, we consider the rectangle -1350m<X_r<18500m, and -3000m<Y_r<-800m. Fig. 2 illustrates that data beyond this rectangle are dropped from the query. We obtained queries for every runway for the entire week. We transform the time stamp of each of these outputs to second-format with respect to a time reference. We consider 12am on January 1, 2003 as time zero.

Position is recorded at a second rate; however, there are time gaps when position is not recorded. For such cases, if the gap is at most 10 seconds, we linearly interpolate the time-position of the aircraft between two boundaries of the time gap for every second. We do not apply this interpolation for the time gaps of more than 10 seconds. This procedure is implemented in MATLAB.

We also need to attach wake vortex weight classes, and weather conditions (Instrumental Meteorological Condition IMC or Visual Meteorological Condition VMC) information to data records. In our one week sample, there are totally 1496 distinct mode-s values. For these aircraft, we managed to obtain wake vortex weight class of 93% of them, of which 67% is provided by Sensis Corporation and the rest is obtained by matching and search of tables of the FAA aircraft registration database, including MASTER, ACFTREF, and Aircraft Information tables. The weather condition for every quarter hour is reported in Aviation System Performance Metrics (ASPM) database in local time. Considering the time column of the data, we add a new column to records to indicate IMC and VMC weather condition.

After data preparation in the aforementioned manner, we now discuss how to extract samples of random variables of the landing process, and compute desired landing statistics. Recorded data of a given aircraft might include many landings, departures, or fly-overs, but these operations are not differentiated in the database. We now introduce an algorithm to distinguish landings from other operations, and to calculate samples of *LTI*, *IAD*, and *ROT* samples.

B. Algorithm to Extract Samples

The procedure should recognize landings then extract necessary records through the following steps:

1. For each mode-s, divide all records of a single aircraft into separate operations (landings, departures, etc). We suppose that a new operation begins whenever there is a time gap of more than 15 minutes between any two records of that aircraft. Any of these operations might be a landing, departure, fly over, or a ground operation.
2. Check if a given operation is a landing on a given runway, 21L for example, by checking if it passes the following tests:
   - Let t_{min} and t_{max} be the first and last times for which the aircraft is in the "query box."
   - If $X(t_{min}) - X(t_{max}) > 5{,}000$ m, then the aircraft proceeds from right to left, and has been long enough in the runway direction to be a candidate for a landing on runways 21L, 21R, 22L, 22R, 27L, or 27R. Similarly, if $X(t_{min}) - X(t_{max}) < -5{,}000$ m, then it is a candidate for a landing on runways 03R, 03L, 04R, 04L, 09R, or 09L.
   - Check if the aircraft ever crosses the threshold of the specific runway and is observed over the runway
3. Repeat step two for all operations and aircraft, and record their threshold time and location. Record the time and location of aircraft when it is first observed outside of the runway rectangle after landing, i.e. taxi-in time and

TABLE I
NUMBER OF PEAK TIME LANDINGS OBSERVED FROM FEB2, 2003 TO FEB8, 2003

	Runway													
a/c Type	03L	03R	04L	04R	09L	09R	21L	21R	22L	22R	27L	27R	Total	%
Not Available	-	1	3	-	-	-	11	0	0	7	1	2	26	1.4
Small	-	19	26	-	-	-	98	0	3	101	18	17	280	15.1
Large	-	96	158	-	-	-	445	1	18	483	107	111	1418	76.2
B757	-	8	15	-	-	-	39	0	0	51	5	11	129	6.9
Heavy	-	0	4	-	-	-	1	0	1	1	0	0	7	0.4
Total	0	124	206	0	0	0	594	1	22	643	131	141	1862	100

location. If the aircraft track disappears over the runway, then exit from runway is not recorded, record zero or blank for the exit time.

4. Sort landings in ascending manner, to recognize follow-lead aircraft. Record the location of any follow aircraft at the moment its lead crosses the runway threshold.
5. Calculate *ROT* for any aircraft, and *LTI*, and *IAD* for any pair of lead-follow aircraft. ■

Depending on the objective of a study, observations shall be classified based on weather condition, weight class of follow-lead aircraft, arrival rate, etc.

III. LANDING STATISTICS

We define a peak period for a given runway to be a quarter-hour with at least seven landings on that runway. For the week Feb 2, 2003 to Feb 8, 2003 we observed 1862 peak period landings out of 4313 landings observed for the entire week on all twelve runways. Peak period landings are distributed among runways and aircraft types as shown in Table I. The majority of these landings occur on runways 21L and 22R. Only 1.4% of wake vortex weight classes of peak period landings could not be recognized.

Fig. 3 shows arrival rates per quarter hour for runway 21L. The horizontal axis is in local time. Observations start at 7:00pm Feb 1, 2005. Shaded periods over the time axis indicate IMC periods for the airport. The arrival pattern for runway 22R is similar to this one since the arrival traffic is equally directed to these two parallel runways whenever these runways are in the landing configuration.

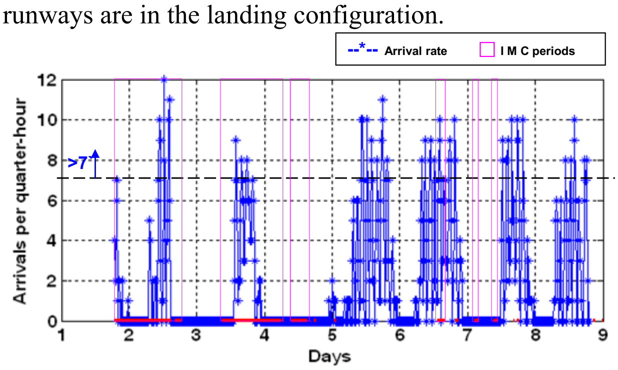


Fig. 3. Arrival rate to runway 21L from late Feb 1 to 8, 2003 local time

To double check completeness of observations in the multilateration database and to validate our data preparation and sample extraction algorithm, we compared the number of landings reported in ASPM database with the results from our study. The comparison plot is given in Fig. 4. Overall for this week, ASPM reports 160 more landings than ours. This corresponds to a small proportion of 3.6% (=100*160/4473) of ASPM records. Average and standard deviation of "Observed minus ASPM" rates are 0.24 and 1.7 arrivals per quarter-hour, respectively. This difference can be the result of missing mode-s and unrecorded landings that might have happened because of off transponders or non-transponder aircraft.

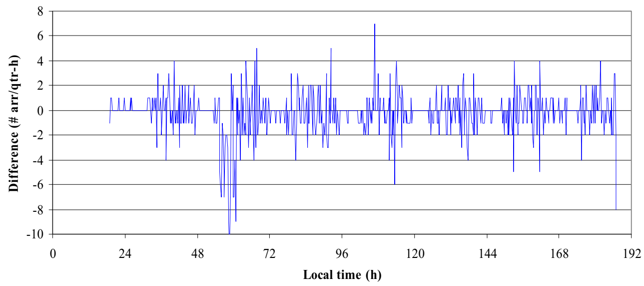


Fig. 4. (Observed – ASPM) per qtr-h arrival rate. The mean and standard deviation are 0.24 and 1.7 arr/qtr-h, respectively.

To analyze system operations it is also important to know the proportion of follower-leader aircraft pairs. Table II shows this proportion for our data (peak times only), which is also called a transition matrix. About 59% of the landings are L-L aircraft. In 77.1% and 77.3% of the times a large aircraft was the lead and the follow aircraft, respectively.

TABLE II
FOLLOW-LEAD AIRCRAFT TRANSITION MATRIX (% OUT OF 1805 PAIRS) IN PEAK PERIODS

Follow \ Lead	Small	Large	B757	Heavy	Sum
Small	1.7	12.5	1.2	0.1	15.5
Large	12.8	58.8	5.4	0.3	77.3
B757	0.9	5.4	0.6	0.0	6.9
Heavy	0.1	0.3	0.0	0.0	0.4
Sum	15.5	77.1	7.1	0.3	100

In 59% of landings a large aircraft follows another large one. In 77% of the time a large aircraft is the following (the leading) one.

A. Peak time ILS Landing Probability Distributions

In risk and capacity analysis, the pattern of the approach process behavior in peak periods is of interest. For this reason, we focus on periods during which there are seven or more landings per quarter hour. Also, the approach process under IMC is the subject of sampling and distribution estimation in this paper.

Table III is the default standard for the "approach in-trail threshold separation minima" under Instrument Flight Rule (IFR) put forth by Federal Aviation Administration. We are interested to know what the probability distributions of *LTI* and *IAD* are for class of follow-lead aircraft with the 3 nmi separation spacing minima indicated in Table III, i.e. pairs S-S, L-S, B757-S, H-S, L-L, B757-L, and H-L. In specific situations, 3 nmi spacing standard may be reduced to 2.5 nmi [12]-[13]. However, differentiating these situations is not the subject of this paper.

TABLE III
IFR APPROACH IN-TRAIL THRESHOLD SEPARATION MINIMA (NMI)

Follow a/c	Lead a/c			
	Small	Large	B757	Heavy
Small	3	4	5	6
Large	3	3	4	5
B757	3	3	4	5
Heavy	3	3	4	4

We have obtained 511 samples of *IAD* and 523 samples of *LTI* for the class of pairs of interest. Independence of samples is examined by "one-lag scatter plot" in Fig. 5 for *IAD*; for more information on statistical concepts discussed in this paper see, e.g., [14]-[16], for example. The plot does not demonstrate a specific pattern of dependency among the samples and one-lag correlation coefficient is 0.25. Higher degrees of lags have lower correlation coefficients Thus independence of *IAD* samples, which is required for distribution fitting purposes, is accepted. In the same manner, we conclude independence of *LTI* samples by examining the one-lag scatter plot with related correlation coefficient of 0.25.

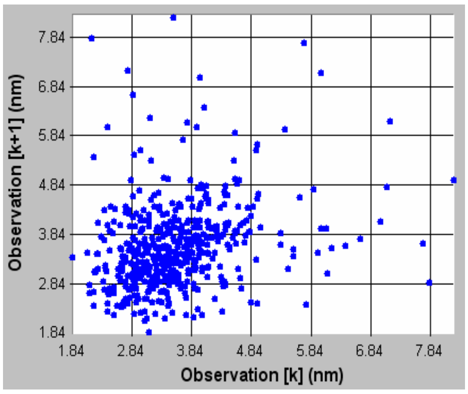


Fig. 5. The one-lag scatter plot of peak-IMC period *IAD* of pairs with 3 nmi separation standard (511 samples). The one-lag correlation coefficient is 0.25. Correlation coefficients for higher degrees of lags are smaller.

We presented histograms and probability distribution function (PDF) fits for *IAD* and *LTI* in Fig. 6 and Fig. 7 with increments of 0.5 nmi and 15 s, respectively. For practicality, in fitting a distribution, we limit *IAD* to a minimum of 1.5 nmi and estimate its distribution by Erlang(1.5;0.35,6) where the values represent location (shift), scale, and shape parameters, respectively. The mean of the Erlang distribution is [(location par.)+ (shape par.)*(scale par.)], and the variance is [(shape par.)*(scale par.)^2]. We use the Maximum Likelihood Estimation (MLE) method for this estimation and for estimations of *LTI* and *ROT* probability distributions. The fit passes Kolmogorov-Smirnov test (KS-test) for significance levels less than 0.10. The Log-Logistic(1.5;1.9,4.5) distribution provides a slightly better fit where values represent location, scale, and shape parameters, respectively.

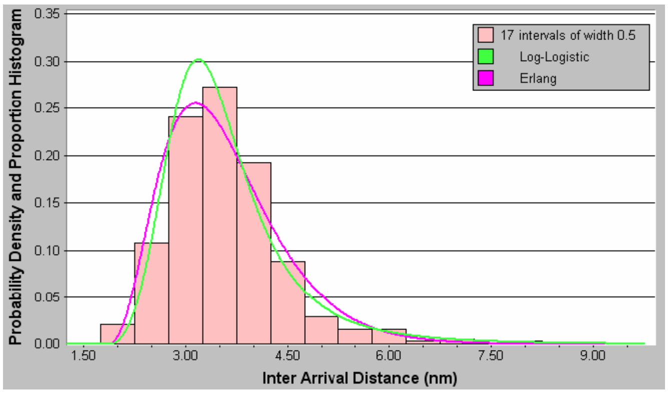


Fig. 6. *IAD* histogram and distribution fits for 511 samples. Sample mean is 3.6 and standard deviation is 0.88 nmi. Erlang(1.5;0.35,6) fit has the mean 3.6 and standard deviation 0.86 nmi.

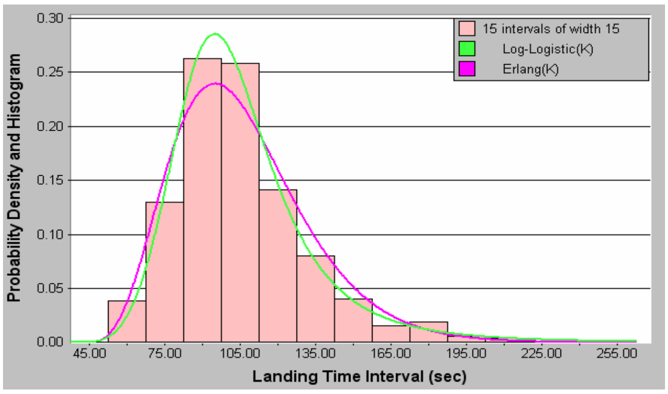


Fig. 7. *LTI* histogram and distribution fits for 523 samples. Erlang(40;11,6) has the mean 106 and standard deviation 27 seconds.

We estimated probability distribution of *LTI* by Erlang(40;11,6) when we enforce minimum of 40 seconds. Similar to the *IAD* case, the Log-Logistic(40;61,4.4) distribution provides a slightly better fit than Erlang distribution, which is a specific case of the gamma distribution. The Erlang fit is accepted by KS-test for significance levels of 0.05 or smaller.

We have obtained 669 samples of *ROT* in peak IMC periods. We conclude that they are independent because the *N*-lag correlation coefficient, *N*=1, 2,… , is less than or equal to 0.08; also the one-lag scatter plot does not show any specific pattern of relationship. The histogram and two distribution fits for *ROT* samples are shown in Fig. 8. We estimate the distribution of *ROT* using three different distributions – gamma, beta, and normal. Using the MLE method, the best fits are Gamma(25;2.8,8.5) in the enforced range of (25,∞) s, Beta(25,110;6.1,15.4) in the enforced range of (25,110) s, where 3rd and 4th values represent shape parameters, and N(49,8.1^2) in the open range of (-∞,∞). Gamma is the best fit among these three with the maximum likelihood criterion; however, the beta distribution might be preferred because, as in real situations for *ROT*, it has lower and upper bounds (for example it can not be negative). The normal distribution, which is used in [3]-[5], is rejected for *ROT* samples in the 0.1 significance level. Mean and variance

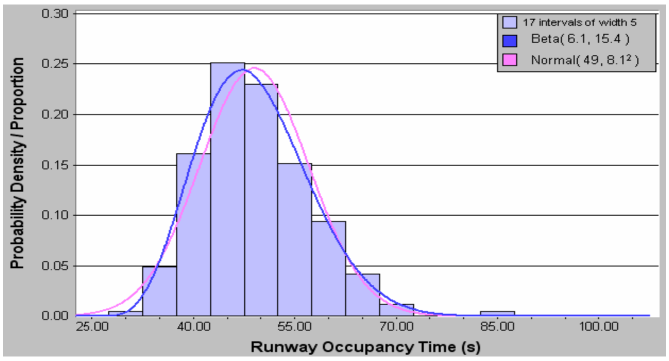


Fig. 8. Peak-IMC periods *ROT* histogram and distribution fits for 669 samples for all aircraft types. Sample mean and standard deviation are 49.1, and 8.1 s. Beta(25,110;6.1,15.4) fit has the mean 49.1, and standard deviation 8.1 s.

of Beta($L, U; \alpha, \beta$) are

$$\begin{cases} \mu = \mathrm{L} + \dfrac{1}{\mathrm{U - L}} \cdot \dfrac{\alpha}{\alpha + \beta} \\ \sigma^2 = \left(\dfrac{1}{U - L}\right)^2 \cdot \dfrac{\alpha \cdot \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)} \end{cases}.$$

We also want to know if *ROT* is different under IMC and VMC weather conditions. Fig. 9 shows histograms of *ROT* under VMC and IMC for the runways with similar taxiway configurations, i.e. 21L/03R and 22R/04L. The structure of runways 27L and 27R seems to be different from 21L/03R and 22R/04L, as also seen in Table IV.

Visual inspection of the figure does not suggest any

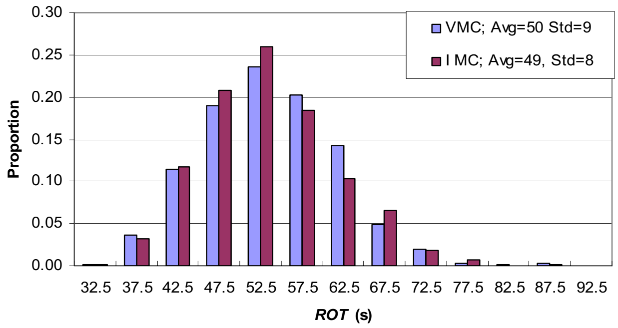


Fig. 9. Histogram of *ROT* under VMC (895 samples) vs. IMC (590 samples) for runways 21L/03R and 22R/04L. We can not observe significant difference between IMC and VMC samples.

significant difference between IMC and VMC *ROT* for this week of data. (Here, IMC / VMC is only distinguished by the corresponding flag in the ASPM database. We have not conditioned on other complementary variables, like surface visibility, which might also affect ROT. Thus, data from a different week in which surface visibility is reduced might show a distinction between IMC / VMC.) From the data, the average and standard deviation of *ROT* in VMC are 50 s, and 9 s, respectively. In IMC, the average and standard deviation are 49 s, and 8 s.

Also of interest is the probability (or frequency) that the *LTI* between two consecutive aircraft is less than the *ROT* of the leading aircraft. We represent this probability by P{$LTI_{k,k+1} < ROT_k$}, k=1, 2, …, and name it "*runway-related approach risk*." Fig. 10 shows pairs of observations ($LTI_{k+1,k}$, ROT_k) observations which is the *ROT* of the lead aircraft k versus the

TABLE IV

ROT IN PEAK PERIODS FOR LANDING RUNWAYS

Runway	Statistic	VMC	IMC
03R	*N*	60	30
	Range (s)	[33,68]	[32,64]
	Avg (s)	48	47
	Std (s)	7	8
04L	*N*	63	91
	Range (s)	[40,60]	[39,68]
	Avg (s)	48	49
	Std (s)	6	6
21L	*N*	271	148
	Range (s)	[31,70]	[29,72]
	Avg (s)	45	48
	Std (s)	8	10
22L	*N*	22	-
	Range (s)	[40,79]	-
	Avg (s)	55	-
	Std (s)	9	-
22R	*N*	283	171
	Range (s)	[26,72]	[39,70]
	Avg (s)	53	50
	Std (s)	6	6
27L	*N*	60	38
	Range (s)	[38,58]	[39,60]
	Avg (s)	48	48
	Std (s)	5	5
27R	*N*	72	26
	Range (s)	[38,105]	[36,84]
	Avg (s)	57	54
	Std (s)	12	11
Total	*N*	828	504
	Range (s)	[26,105]	[29,84]
	Avg (s)	49	49
	Std (s)	8	8

Overall, from this table, we can not observe significant difference between *ROT* under IMC and VMC. Based on the sample, runway 27R has higher mean and variability than other runways.

LTI between aircraft k and k+1 for peak period landings. We have limited *LTI* in the figure to 200 seconds for the purpose of clarity. In this figure, there are two observations having *ROT* of 105 s which correspond to landings on runway 27R in VMC. They are exceptional cases since they are far from other sample population and we consider them as outliers. They are 19 s bigger than the second largest sample 86 s, for example.

Fig. 10 also demonstrates independence of $LTI_{k,k+1}$ and ROT_k., for all k. The Kendall "sample-correlation statistic," which measures dependency in non-parametric statistics, is 0.085 and supports independence of these random variables; for more discussion on this parameter see [16]. The sample correlation coefficient is 0.15 and also confirms the independence hypothesis.

Now we provide an empirical and a theoretical "point estimation" for P{$LTI_{k,k+1} < ROT_k$}, k=1, 2, …, in peak periods

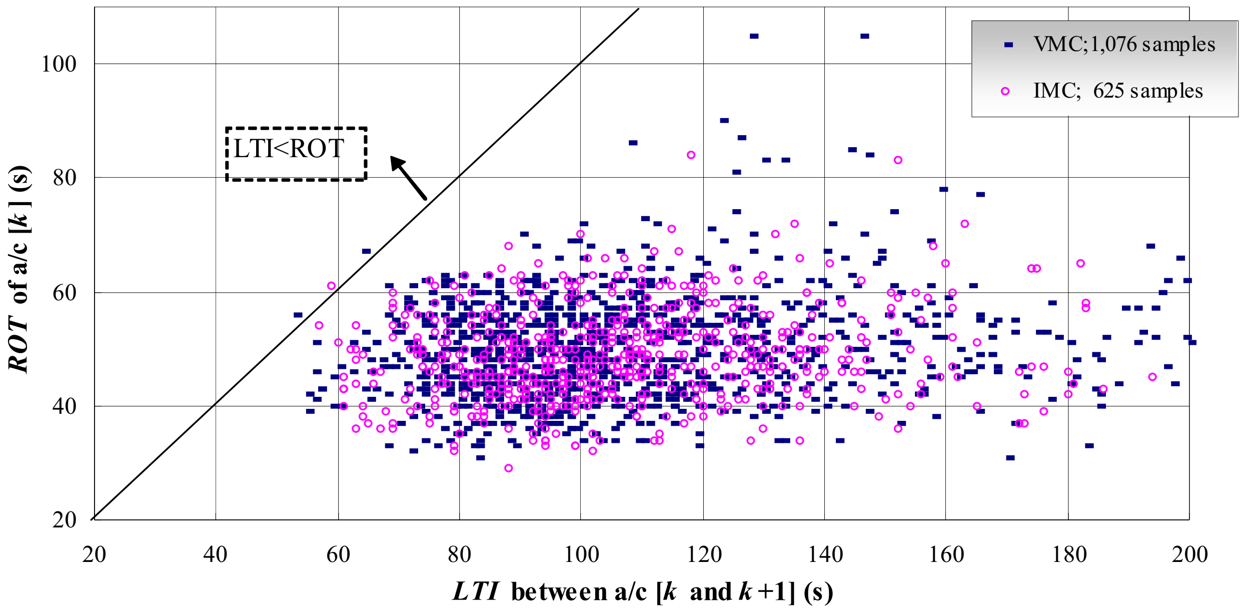


Fig. 10. Runway Occupancy time of aircraft k, ROT_k , versus Landing Time Interval between aircraft k and k+1, $LTI_{k,k+1}$. In this figure, pairs of follow-lead aircraft are not differentiated based on their weight class. For three points above 45 degree line, $LTI_{k,k+1}$ is less than ROT_k one of which has occurred under IMC.

for pairs of aircraft with separation standard 3 nmi in table III.

1) Empirical Method: There are three points above the 45 degree line. These points represent events where $LTI_{k,k+1}<ROT_k$.. The sample frequency is 0.0016 with respect to 1862 peak period landings. (Out of 1862 landings, there were 108 landings for which we could not obtain the *ROT* due to disappearance of the aircraft track over the runway. This might be because the aircraft turned off the transponders or for other reasons. 44 of these lost data happened in IMC and 64 in VMC. We assume that these landings would not have been above the 45 degree line in the figure.)

The frequency of $LTI_{k,k+1} < ROT_k$, for k=1,…,4312, is 0.0007 - that is, 3 out of 4313 landings, assuming that no such event occurred in non-peak periods. As shown in Fig. 10, for 1 (for 2) out of 625 IMC (1076 VMC) landings we have $LTI_{k,k+1}<ROT_k$, i.e. the estimated probability of 0.0016 (0.0019).

2) Theoretical method: We use the probability distribution fits that we calculated as *ROT*~Beta(6.1,14.5) in the range (25,110), and *LTI*~Erlang(40;11,6) to estimate P{*LTI*<*ROT*}. Fig. 11 shows the overlap of these probability distributions. Because there is an overlap between *LTI* and *ROT*, then P{*LTI*<*ROT*} is positive. We note that in fitting the PDF for *LTI*, we have not considered samples of *LTI* that we could not obtain their corresponding *ROT*. Let $g_{ROT}(\cdot)$ represent PDF of *ROT*, and $F_{LTI}(\cdot)$ represent Cumulative Density Function (CDF) of *LTI*. Then,

$$\begin{aligned} p\{LTI < ROT\} &= \int_{-\infty}^{\infty} p\{LTI < ROT \mid ROT = x\} \cdot g_{ROT}(x)dx \\ &= \int_{25}^{110} p\{LTI < x\} \cdot g_{ROT}(x)dx \\ &= \int_{25}^{110} F_{LTI}(x) \cdot g_{ROT}(x)dx. \end{aligned} \quad (4)$$

Equation (4) cannot be evaluated analytically for the distributions we have chosen. We estimate (4) using stochastic simulation. The result is 0.004, as a point estimation for the pairs of interest in peak-IMC period.■

We see that the theoretical estimation 0.004 is about 2.5

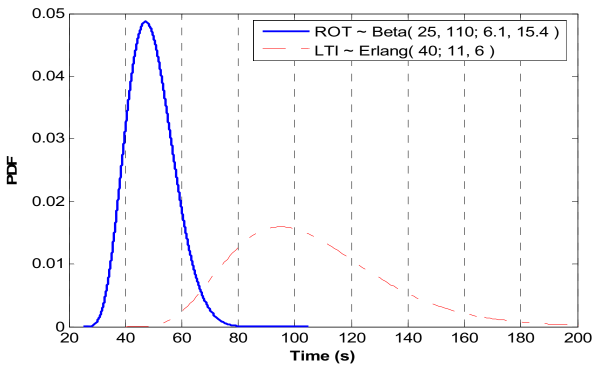


Fig. 11. Overlap of *ROT* and *LTI*. We obtain point estimation 0.004 for P{*LTI*<*ROT*}. The estimation would be slightly smaller if we chose log-logistic distribution for *LTI*.

times of the empirical estimation 0.0016 for peak-IMC periods. We shall note that these estimations are optimistic firstly because we have missed about 3.5% of total landings based on the ASPM, and secondly we could not obtain *ROT* for 44 out of 669 peck-IMC landings. These two effects may have added to P{*LTI*<*ROT*}, i.e. may have had bigger *LTI* than *ROT* of their leading aircraft.

IV. CONCLUSIONS

We presented an efficient way to use multilateration surveillance system data taking into account and analyzing noise, errors, and missing data. We obtained the wake vortex weight class of 98.6% of aircraft landing in peak periods. We added this information to the multilateration data along with the meteorological conditions that we obtained from the ASPM database. We gave an algorithm to extract samples of random variables *LTI*, *IAD*, and *ROT* from the data. The samples were conditioned on IMC times and peak traffic periods in which there were seven or more landings per quarter hour on a given runway. Also, *LTI* and *IAD* were additionally conditioned based on follower-leader wake vortex weight class and aggregated for ones with a minimum separation standard of 3 nmi – namely, pairs S-S, L-S, B757-S, H-S, L-L, B757-L, and H-L.

The data supported our assumption that samples of each random variable were independent. We represented the PDF of *LTI*, *IAD*, and *ROT* by a few known density functions and compared their performance. Fitting distributions to the collected samples showed that *ROT* is best represented by a beta distribution, but not with a normal distribution, which is generally assumed in the literature. *LTI*, and *IAD*, for the F-L pairs of under study, were best fit by log-logistic distributions; however, Erlang (gamma) distribution was also accepted for these random variables. We preferred to use the Erlang distribution rather than the log-logistic distribution because it is better known and has enough accuracy to represent behavior of *LTI* and *IAD*. We also showed that *LTI* between the leading and following aircraft is independent of *ROT* of the leading one. Our overall observations suggested that there was almost no difference of *ROT* between IMC and VMC conditions, for the particular week observed at DTW. We estimated the probability (or frequency) of *LTI*<*ROT* in peak-IMC periods with empirical and theoretical calculations.

Investigation of data for longer time periods, e.g. one month, and for a visual landing system (VFL) at this airport and other airports can be the subject of future studies. Providing methodologies to incorporate incomplete data of (*LTI*, *ROT*) with missing *ROT* in estimation of "runway-related risk" can be a research problem. Distribution of other random variables in the approach process, such as time between exits from the runway, and inter arrival times to the terminal radar approach control (TRACON) area, are subjects for future research.

DISCLAIMER

This paper solely represents the opinions of the authors and does not necessarily reflect the opinion of the United States government or NASA.

ACKNOWLEDGEMENT

The authors would like to thank Wayne Bryant, Wake Program Manager at NASA Langley, for support of this research. We also thank the Volpe National Transportation Systems Center and Sensis Corporation for help in acquiring, manipulating, and understanding the multilateration data. We would like to thank Dr. George Donohue for helpful comments and insights on this research.

REFERENCES

[1] B. Levy, J. Legge, and M. Romano, "Opportunities for improvements in simple models for estimating runway capacity," presented at the 23rd Digital Avionics Systems Conference, Salt Lake City, UT, October 2004.
[2] C.R. Haynie, "An investigation of capacity and safety in near-terminal airspace for guiding information technology adoption," Ph.D. dissertation, Dept. Sys. Eng. and Oper. Res., George Mason Univ., Fairfax, VA, 2002.
[3] Y. Xie, J. Shortle, G. Donohue, "Runway landing safety analysis: a case of Atlanta Hartsfield airport," presented at the 2003 Digital Avionics Systems Conf., Indianapolis, IN, 2003.
[4] Y. Xie, J. Shortle, P. Choroba, "Landing safety analyses of an independent arrival runway," presented at the 2004 Int. Conf. for Res. In Air Trans., Zillina, Slovakia.
[5] Y. Xie, "Quantitative analysis of airport runway capacity and arrival safety using stochastic methods," Ph.D. dissertation, Dept. Sys. Eng. and Oper. Res., George Mason Univ., Fairfax, VA, 2005.
[6] Ballin, M.G., H. Erzberger, "An analysis of landing rates and separations at the Dallas/Fort Worth International Airport," NASA, Tech. Memo. 110397, July 1996.
[7] H.F. Vandevenne, and M.A. Lippert, "Using maximum likelihood estimation to determine statistical model parameters for landing time separations," 92PM-AATT-006, March 2000
[8] J.W. Andrews, and J.E. Robinson, "Radar-based analysis of the efficiency of runway use," AIAA Guidance, Navigation, and Control Conf., Quebec, August 2001
[9] J. Rakas, H. Yin, "Statistical modeling and analysis of landing time intervals: case study of Los Angeles International Airport, California," Transportation Research Record: Journal of the Transportation Research Board, No. 1915, 2005, pp. 69-78.
[10] FAA Airport Diagram, AL-119, 16 Feb 2006, Available http://204.108.4.16/d-tpp/0602/00119AD.PDF
[11] B. Jeddi, "Preliminary Exploration of DROMSIV data on Detroit Airport (DTW)," CATSR Internal report, August 2004
[12] FAA Order 7110.65, Air Traffic Control., Federal Aviation Administration, Sept. 1993, Wake Turbulence, Para 2,1,19
[13] FAA Order 7110.65, Same Runway Separation, Federal Aviation Administration, Sept. 1993, Wake Turbulence, Para 3,9,6
[14] A.H. Bowker, and G.J. Lieberman, *Engineering Statistics*, 2nd Ed, Prentice-Hall, Inc, 1972.
[15] A.M. Law, *ExpertFit: user's guide*, Averill M. Law & Associates, 2000
[16] M. Hollander, D.A. Wolf, *Nonparametric Statistical Methods*, 2nd Ed., John Wiley & Sons, Inc., 1999.

Babak G. Jeddi (B.S.'93–M.S.'96 and '03) is a Ph.D. candidate in the Department of Systems Engineering and Operations Research, and a research assistant in the Center for Air Transportation Systems Research at George Mason University. His experience includes quality and productivity improvement in manufacturing and service organizations. His research interests include analysis of supply chain and air transportation systems via stochastic processes, simulation, and network modeling. He received his masters' degree in Industrial Engineering from University of Cincinnati, Ohio, US.

John F. Shortle, Ph.D., is an assistant professor of Systems Engineering at George Mason University. His experience includes developing stochastic, queuing, and simulation models to optimize networks and operations. His research interests include simulation and queuing applications in telecommunications and air transportation. He received his doctorate degree in operations research from University of California at Berkeley, California, US.

Lance Sherry, Ph.D., is an associate research professor, and executive director of the Center for Air Transportation Systems Research at George Mason University. His research interests include productivity improvement, strategic planning, and dynamic systems. He received his doctorate degree in Industrial Engineering from Arizona State University, Arizona, US.

How to model and optimize the Airport Capacity Allocation Problem

Abdellah Idrissi Chu Min Li
LaRIA, CNRS FRE 2733
Faculté de Mathématiques et d'Informatique
Université de Picardie
33, rue St Leu, 80039, Amiens, France
Email: {idrissi, chu-min.li}@u-picardie.fr

***Abstract*— The Constraint Satisfaction Problem (CSP) is proven more and more promising to model and to solve a large number of real problems. A lot of approaches using constraint reasoning have proposed to solve search problems. In this paper, we propose a modeling of Capacity Allocation Problem of an airport (CAP) and of its fixes ($FCAP$) in form of a Constraint Satisfaction Problem (CSP). This modeling gives a CSP problem called $CSCAP$. In order to assist the airport managers to regulate the arrival/departure predicted demands, and to use efficiently the available capacities of a terminal, we show how to solve the $CSCAP$ and how to optimize the utilization of existing airport capacity. Finally, we present some experimental results followed by its optimized resolution.**

I. Introduction

The Constraint Satisfaction Problem (CSP) [21] is proven more and more promising to model and to solve a large number of real problems. A lot of approaches using constraint reasoning have proposed to solve search problems. Formally, a CSP is defined by the triplet (X, D, C), where $X = \{X_1, ..., X_n\}$, is the set of n variables; $D = \{D_1, ..., D_n\}$, is the set of n domains of values; D_i is the domain of values of the variable X_i and $C = \{C_1, ..., C_e\}$, is the set of e constraints of the problem, specifying compatible values or excluding incompatible values between variables. Solving a CSP consists in assigning values to variables in order to satisfy all the constraints. In general, the CSPs are solved by traditional methods combining a mechanism of search with Backtrack (BT) and a mechanism of reinforcement of consistencies with each node of search.

In this paper, we model the Capacity Allocation Problem (CAP) of an airport and its fixes ($FCAP$) in a CSP and apply CSP techniques to solve it, especially in case of congestion. The congestion occurs at an airport and its fixes when the request of the traffic exceeds the available capacity.

In order to assist the managers of an airport to regulate the arrival and departure demands, and to use efficiently the available capacities of a terminal, we show in this paper how to model and optimize the utilization of existing airport capacity.

This paper is organized as follows. Section II defines the capacity allocation problem. In section III we present a modeling of this problem in form of a CSP which we called $CSCAP$, and propose an optimization model to solve it in section IV. We give an algorithm in section V and expose in section VI an application example of a $CSCAP$. We show in section VII how to assist the managers to regulate traffic and unforeseen events. We conclude in section VIII and discuss some perspectives of this work.

II. Capacity Allocation Problem

A. Presentation

An operational scheme of a single airport system that reflects the arrival and departure processes at the airport is shown in figure 1. The system comprises arrival ways called arrival fixes 'AF', departure ways called departure fixes 'DF', and a runway system. There are two separate sets of arrival and departure fixes located in the near-terminal airspace area (50-70 km off the airport), so that the arrival fixes serve only arrival flow, and the departure fixes serve only departure flow. The runway system on the ground serves both arrival and departure flows.

The arrival flights are assigned to specific arrival fixes, and, before landing, they should go through the fixes. After leaving runways, the arriving flights follow the taxiway to the gates at the terminal. The departing flights are also assigned to the specific fixes. They are directed to the runways, and go through the departure fixes after leaving runways.

The arrival queues are formed before the fixes (see figure 1). This means that the flights which are already in the fixes, must be accepted at the runways. If there is an arrival queue, a certain number of flights should be delayed in the air.

The departure queue is formed before the runway system, and flights can be delayed either at their gates or on the taxiway.

The operational limits on the ground (runways) are characterized by arrival capacity and departure capacity. These capacities are generally variable and interdependent.

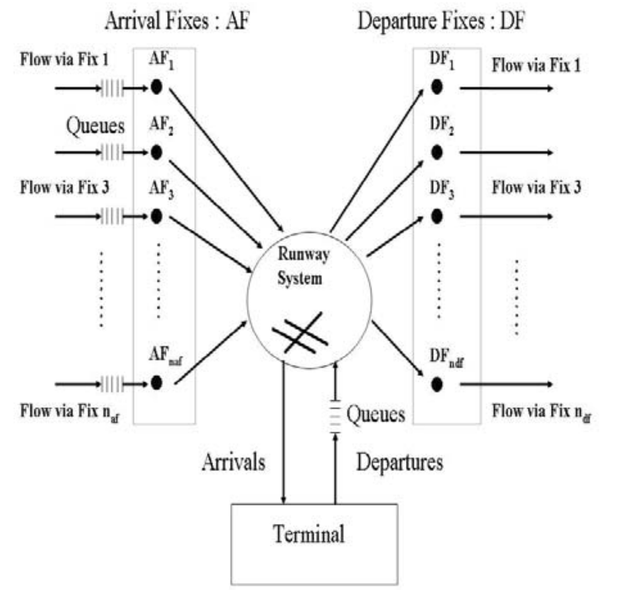


Fig. 1. Example of an airport system

Intervals of time	Predicted arriving	Predicted departing
= i	X_{a_i}	Y_{d_i}
12:00-12:15 = 1	26	36
12:15-12:30 = 2	38	32
12:30-12:45 = 3	42	09
12:45-13:00 = 4	29	15
13:00-13:15 = 5	06	07
13:15-13:30 = 6	13	10
13:30-13:45 = 7	14	17
13:45-14:00 = 8	20	33
14:00-14:15 = 9	40	34
14:15-14:30 = 10	25	22
14:30-14:45 = 11	13	13
14:45-15:00 = 12	12	01
TOTAL	278	229

TABLE I

Example of predicted number of arriving and departing flights per each 15 minutes time interval of a 3 hours period.

The traffic demands for an airport are given by the predicted number of arriving and departing flights per each 15 minutes time interval of a considered period (see table I).

Definition 1. ***Airport capacity*** is defined respectively as the maximum number of operations (arrivals and departures) that can be performed within a fixed time interval (e.g., 15 minutes or other interval) at a given airport under given conditions such as runway configuration and weather conditions.

Definition 2. ***Fix capacity*** is defined as the maximum number of flights that can cross a fix in a 15 minutes interval (or other interval) under given conditions such as weather conditions, etc. These capacities determine the operational constraints in the near-terminal airspace.

Definition 3. ***Airport capacity allocation problem*** consists in determining a balance, between arrivals and departures, to minimize an adapted objective function for a given time period.

The capacity of the airports and the capacity of their fixes become more and more limited compared to the demands and present real problems for the aerial transportation system. The optimal allocation of the existing capacity of an airport and of its fixes during the congestion periods is very important to maximize the departures and to minimize the delay.

B. Related Work

Aerial traffic is continuously increasing and its regulation becomes more and more difficult, given the limited capacities both in ground (airports) and air (aerial sectors). Various solutions can be used to attack this difficulty: building new airports, physically extending existing runway systems, applying new technologies to increase airport capacity, and optimizing utilization of existing airport capacity [4]. See [12] for some possible measures for increasing airport capacity.

On the other hand, aerial traffic regulation should be optimized and automated, especially when treating unforseen events. In [14], we proposed a modeling and a resolution of the aerial conflicts problems by techniques of constraints network (CSP). This modeling and automatic resolution can be possibly introduced into project (FREER) [8] developed by Eurocontrol following the concept “Free Flight” [9].

At an airport, the air traffic managers generally know the range of arrival/departure capacity tradeoff available for each runway configuration and use the tradeoff for strategic planning of arrival and departure traffic. They control and modify the operational arrival and departure limits during the periods of congestion to better serve the traffic demand and to minimize the delay.

However, currently they proceed intuitively because of the lack of automated decision support to determine the best strategies for allocation of airport arrival and departure capacities. An experienced specialist can find the best solutions but, in more complex situations, this task can become impossible.

In [16], we proposed a modeling of the airport capacity allocation problems by techniques of constraints network (CSP). We associated an optimization model to it. But, we had not taken into account the time limit constraint (LT). Because we estimate that in the breakdown event or others, it is necessary to provide a solution even if it is not of very good quality.

In addition, the method presented in [16] find the optimal solution, but spent much time before delivering it, especially when the number of time intervals is greater or equal than 17. To deal with this problem, we propose an improvement of this method by grafting to it a method of local search (LS) combined with the Backtrack (B) beyond a certain Time Limit constraint (TL) fixed by expert. Thus we have constantly, beyond fixed time, at least a solution which perhaps can not be the optimal solution, but which can be acceptable. We called this new method BLSTL (Backtrack + Local Search + Time Limit constraint).

In order to optimize and automate the aerial traffic regulation, we propose in the next section a modeling of capacity allocation problem of an airport and its fixes using constraint network (CSP).

III. *CSCAP*: A CSP model for the Capacity Allocation Problem

A. Problem Formulation

Figure 1 shows a simplified scheme of an airport constituted with several arrival fixes 'AF', several departure fixes 'DF' and one terminal T serving both arrival and departure demands during a time interval 'i' ('i' is a period of 15 minutes). The arrival demands are at the point 'AF'. The departure demands are at the point 'DF'. The point 'T' represents the Terminal.

In the sequel, we consider that there are 'j' arrival fixes $AF_{j_{df_i}}$ with $j = 1, ..., n_{af}$ and 'k' departure fixes $DF_{k_{df_i}}$ with $k = 1, ..., n_{df}$. Each fix 'j' or 'k' has its own capacity noted $C^j_{af_i}$ or $C^k_{df_i}$ during the same interval of time 'i'.
Let us note P_{v_i} the number of empty places at the terminal, P_{oc_i} the number of occupied places at the terminal and C_{T_i} the capacity of the terminal. C_{T_i} is the sum of P_{v_i} and of P_{oc_i}.

$$C_{T_i} = P_{v_i} + P_{oc_i} \tag{1}$$

If $X_{j_{afe_i}}$ represents the number of arrival flights accepted at fix j ($j = 1, ..., n_{af}$) and $Y_{k_{dfe_i}}$ the number of departure flights effectively leaving through fix k ($k = 1, ..., n_{df}$) during time interval 'i', then the number of empty places for the time interval 'i+1' becomes:

$$P_{v_{i+1}} = P_{v_i} + \sum_{k=1}^{n_{df}} Y_{k_{dfe_i}} - \sum_{j=1}^{n_{af}} X_{j_{afe_i}} \tag{2}$$

In the same way, the number of occupied places for the time interval 'i+1' becomes:

$$P_{oc_{i+1}} = P_{oc_i} + \sum_{j=1}^{n_{af}} X_{j_{afe_i}} - \sum_{k=1}^{n_{df}} Y_{k_{dfe_i}} \tag{3}$$

For the sake of simplification, let us admit in the sequel that C_{T_i} is a constant independent of time. We will note it C_T such as:

$$C_T = P_{v_i} + P_{oc_i} = P_{v_{i+1}} + P_{oc_{i+1}} \tag{4}$$

We propose hereafter a modeling of this problem in form of a CSP.

B. Problem Modeling

We propose to model this problem of the arrivals/departures at an airport and at its fixes (figure 1) in form of a CSP, $CSCAP = (X_i, D_i, C_i)$ during a time interval 'i', where:

- $X_i = \{X_{a_i}, X_{ae_i}, Y_{d_i}, Y_{de_i}, C_{a_i}, C_{d_i}, Q_{a_i}, Q_{d_i}, X_{1_{af_i}}, ..., X_{n_{af_i}}, X_{1_{afe_i}}, ..., X_{n_{afe_i}}, Y_{1_{df_i}}, ..., Y_{n_{df_i}}, Y_{1_{dfe_i}}, ..., Y_{n_{dfe_i}}, C_{1_{af_i}}, ..., C_{n_{af_i}}, C_{1_{df_i}}, ..., C_{n_{df_i}}, Q_{1_{af_i}}, ..., Q_{n_{af_i}}, Q_{1_{df_i}}, ..., Q_{n_{df_i}}, X_{t_{a_i}}, Y_{t_{d_i}}, P_{v_i}, P_{oc_i}\}$ is the set of variables such as :
  X_{a_i} is the number of arrival demands at the point 'AF' (figure 1). X_{ae_i} is the number of effective arrivals at 'AF' (the number of planes effectively served). Y_{d_i} is the number of departure demands at the point 'DF' (figure 1). Y_{de_i} is the number of effective departures at 'DF'. C_{a_i} is the arrival capacity at 'AF'. C_{d_i} is the departure capacity at 'DF'. C_{a_i} and C_{d_i} are dictated by the airport managers. Q_{a_i} is the number of arrivals delayed at 'AF'. Q_{d_i} is the number of departures delayed at 'DF'. $X_{j_{af_i}}$ is the number of arrival demands at the fix '*j*' (figure 1) with $j = 1, ..., n_{af}$. $X_{j_{afe_i}}$ is the number of effective arrivals at the fix '*j*' with $j = 1, ..., n_{af}$. $Y_{k_{df_i}}$ is the number of departure demands at the fix '*k*' (figure 1) with $k = 1, ..., n_{df}$. $Y_{k_{dfe_i}}$ is the number of effective departures at the fix '*k*' with $k = 1, ..., n_{df}$. $Q_{j_{af_i}}$ is the number of arrivals delayed at the fix '*j*' with $j = 1, ..., n_{af}$. $Q_{k_{df_i}}$ is the number of departures delayed at the fix '*k*' with $k = 1, ..., n_{df}$. $C_{j_{af_i}}$ is the arrival capacity at the fix '*j*' with $j = 1, ..., n_{af}$. $C_{k_{df_i}}$ is the departure capacity at the fix '*k*' with $k = 1, ..., n_{df}$. $X_{t_{a_i}}$ is the total number of arrival demands for the time interval 'i'. Note that for each value of 'i', we add the delay $Q_{a_{i-1}}$ recorded in the previous interval 'i-1' to the predicted arrival demands X_{a_i}. Then X_{a_i} becomes $X_{a_i} + Q_{a_{i-1}}$ which we noted $X_{t_{a_i}}$ such as $Q_{a_0} = 0$. $Y_{t_{d_i}}$ is the total number of departure demands for the time interval 'i'. For each value of 'i', we add the delay $Q_{d_{i-1}}$ recorded in the previous interval 'i-1' to the predicted departure demands Y_{d_i}. Then Y_{d_i} becomes $Y_{d_i} + Q_{d_{i-1}}$ which we noted $Y_{t_{d_i}}$ such as $Q_{d_0} = 0$. P_{v_i} and P_{oc_i} are respectively the set of empty and occupied places of the terminal at the point 'T' of figure 1 during the same time interval 'i'.
- $D_i = \{D_{X_{a_i}}, D_{X_{ae_i}}, D_{Y_{d_i}}, D_{Y_{de_i}}, D_{C_{a_i}}, D_{C_{d_i}}, D_{Q_{a_i}}, D_{Q_{d_i}}, D_{X_{1_{af_i}}}, ..., D_{X_{n_{af_i}}}, D_{X_{1_{afe_i}}}, ..., D_{X_{n_{afe_i}}}, D_{Y_{1_{df_i}}}, ..., D_{Y_{n_{df_i}}}, D_{Y_{1_{dfe_i}}}, ..., D_{Y_{n_{dfe_i}}}, D_{C_{1_{af_i}}}, ..., D_{C_{n_{af_i}}}, D_{C_{1_{df_i}}}, ..., D_{C_{n_{df_i}}}, D_{Q_{1_{af_i}}}, ..., D_{Q_{n_{af_i}}}, D_{Q_{1_{df_i}}}, ..., D_{Q_{n_{df_i}}}, D_{X_{t_{d_i}}}, D_{Y_{t_{d_i}}}, D_{P_{v_i}}, D_{P_{oc_i}}\}$ is the set of variable domains. We consider that all the domains of these variables are equal and that D_i represents a common domain to all these variables. In other words, $D_i = \{0, 1, 2, ..., C_T\}$.
  Note that the values of arrival and departure capacities are interdependent and are dictated by the airport

managers in form of couples (C_{a_i}, C_{d_i}). They can recommend to use for example, at a given period, the set {(18, 29), (24, 24), (26, 19), (28, 15), (17, 30), (20, 27)}. In this example, the first couple means that if $C_{a_i} = 18$ then $C_{d_i} = 29$ and vice versa. The meaning of other couples is similar.

- C_i is the set of constraints of the problem. We can formulate them as follows :

$$0 \leq P_{v_i} + P_{oc_i} \leq C_T \tag{5}$$

$$0 \leq X_{j_{afe_i}} \leq X_{j_{af_i}}, \forall j \in \{1, ..., n_{af}\} \tag{6}$$

$$0 \leq Y_{k_{dfe_i}} \leq Y_{k_{df_i}}, \forall k \in \{1, ..., n_{df}\} \tag{7}$$

$$0 \leq \sum_{j=1}^{n_{af}} X_{j_{afe_i}} + \sum_{k=1}^{n_{df}} Y_{k_{dfe_i}} \leq C_T \tag{8}$$

$$X_{a_i} = \sum_{j=1}^{n_{af}} X_{j_{af_i}} \tag{9}$$

$$X_{ae_i} = \sum_{j=1}^{n_{af}} X_{j_{afe_i}} \tag{10}$$

$$Y_{d_i} = \sum_{k=1}^{n_{df}} Y_{k_{df_i}} \tag{11}$$

$$Y_{de_i} = \sum_{k=1}^{n_{df}} Y_{k_{dfe_i}} \tag{12}$$

$$C_{a_i} = \sum_{j=1}^{n_{af}} C_{j_{af_i}} \tag{13}$$

$$C_{d_i} = \sum_{k=1}^{n_{df}} C_{k_{df_i}} \tag{14}$$

$$X_{ae_i} + Y_{de_i} \leq X_{a_i} + Y_{d_i} \tag{15}$$

$$Q_{a_i} = \sum_{j=1}^{n_{af}} Q_{j_{af_i}} \tag{16}$$

$$Q_{d_i} = \sum_{k=1}^{n_{df}} Q_{k_{df_i}} \tag{17}$$

$$X_{t_{a_i}} = \sum_{j=1}^{n_{af}} X_{t_{j_{af_i}}} \tag{18}$$

$$Y_{t_{d_i}} = \sum_{k=1}^{n_{df}} Y_{t_{k_{df_i}}} \tag{19}$$

$$X_{t_{j_{af_i}}} = X_{j_{af_i}} + Q_{j_{af_{i-1}}}, \forall j \in \{1, ..., n_{af}\} \tag{20}$$

$$Y_{t_{k_{df_i}}} = Y_{k_{df_i}} + Q_{k_{df_{i-1}}}, \forall k \in \{1, ..., n_{df}\} \tag{21}$$

$$Q_{j_{af_i}} = X_{j_{af_i}} - X_{j_{afe_i}} + Q_{j_{af_{i-1}}}, \forall j \in \{1, ..., n_{af}\} \tag{22}$$

$$Q_{k_{df_i}} = Y_{k_{df_i}} - Y_{k_{dfe_i}} + Q_{k_{df_{i-1}}}, \forall k \in \{1, ..., n_{df}\} \tag{23}$$

$Q_{j_{af_{i-1}}}$, for $j = 1, ..., n_{af}$ (equation (22)) is the number of arrival planes delayed at the fix j for $j = 1, ..., n_{af}$ during the time interval 'i-1' and $Q_{k_{df_{i-1}}}$, for $k = 1, ..., n_{df}$ (equation (23)) is the number of departure planes delayed at the fix k for $k = 1, ..., n_{df}$ during the time interval 'i-1'. These two numbers should be added respectively to the arrival and departure demands during the time interval 'i'. So the total number of arrival demands during 'i' is the sum of demands effectively expressed $X_{j_{afe_i}}$, for $j = 1, ..., n_{af}$ during 'i' increased by the number of planes $Q_{j_{af_{i-1}}}$, for $j = 1, ..., n_{af}$ which were waiting in the queue due to the fact that they were not been served during 'i-1'. Similarly, the total departure demand during 'i' is the sum of demands really expressed $Y_{k_{dfe_i}}$ for $k = 1, ..., n_{df}$ during 'i' increased by the number of planes $Q_{k_{df_{i-1}}}$ at the fix k for $k = 1, ..., n_{df}$ which were waiting in the queue due to the fact that they were not been served during the time interval 'i-1'. The planes which were waiting in the queue during the time interval 'i-1' are served in priority during the time interval 'i'.

C. Problem Solving

Given values of (X_{a_i}, Y_{d_i}) (arrival and departure demands for a time interval 'i') and a number of possible values of $(C_{a_i},\ C_{d_i})$, the problem is to determine the real repartition of the arrival and departure demands among the fixes, and the real repartition of arrival and departure capacities among the fixes, according to the actual demands of each fix and it consists also in assigning a capacity to each demand by respecting the constraints described above. The problem CSCAP is a decision problem and is an optimization problem as well.

Now our constraint network (CSP) is well identified. In particular the variables, the domains of values of these variables and the constraints of the problem are explicitly defined. It remains to apply one of CSP reasoning procedure (Branch and Bound, Local Search, Backtracking, etc.) to solve the capacity allocation problem. We used the Backtrack algorithm combined with Local Search under Time Limit constraint. We called BLSTL this new algorithm.

Let us note that in the case of congestion during a time interval 'i' the sum of numbers of effective arrivals $\sum_{j=1}^{n_{af}}(X_{j_{afe_i}})$ and of effective departures $\sum_{k=1}^{n_{df}}(Y_{k_{dfe_i}})$ is lower than the sum of numbers of arrival demands $\sum_{j=1}^{n_{af}}(X_{j_{af_i}})$ and of departure demands $\sum_{k=1}^{n_{df}}(Y_{k_{df_i}})$.

In other words, only a part of the demands, arrivals or departures, is been served (see equation (15)).

Our objective is to minimize the sum of the numbers of arrival planes delayed ($\sum_{i=1}^{N}[Q_{j_{af_i}}]$) $\forall j \in \{1,...,n_{af}\}$ and to minimize the sum of the numbers of departure planes delayed ($\sum_{i=1}^{N}[Q_{k_{df_i}}]$) $\forall k \in \{1,...,n_{df}\}$ during a given time period (15*N minutes). In other words, we seek to minimize:

$$\sum_{i=1}^{N} [\sum_{j=1}^{n_{af}} Q_{j_{af_i}} + \sum_{k=1}^{n_{df}} Q_{k_{df_i}}] \tag{24}$$

subject to the constraints (5)-(23).

IV. Optimization Model

We propose here an optimization method able to give the best allocation of airport (and/or fix) capacity between arrivals and departures under given operational conditions at the airport (and/or fixes) to satisfy the predicted traffic demand for a time period and to minimize the delay. This optimization method under constraints takes in consideration also the time limit constraint.

When one assigns a capacity to a request consisting of several arrival or departure flight demands, then if this request is totally satisfied, one passes to the next one. If this request is not totally satisfied then we consider that the difference between the request and the available capacity is a penalty p such as $(X - C = p)$, where X is the number of demands in the request, C is the capacity allowed to this request. The problem consists to reduce to the minimum a function of penalty φ under a set of constraints. This function of penalty φ is the sum of the penalties of the unsatisfied demands. Let us note $p(d)$ the penalty of an unsatisfied demand. Then the optimization problem can be characterized by:

$$minimize \quad [\varphi] \quad where \tag{25}$$

$$\varphi = \sum p(d | d \in X \ and \ d \ unsatisfied) \tag{26}$$

In addition, let us note (N_{tar}) the total number of planes delayed over all the period considered. Thus, over one period constituted by N intervals of time one will have:

$$N_{tar} = \sum_{i=1}^{N}[\sum_{j=1}^{n_{af}} Q_{j_{af_i}} + \sum_{k=1}^{n_{df}} Q_{k_{df_i}}] \quad where$$

$$Q_{j_{af_i}} = X_{j_{af_i}} - X_{j_{afe_i}} + Q_{j_{af_{i-1}}}, \forall j \in \{1,...,n_{af_i}\}$$

and

$$Q_{k_{df_i}} = Y_{k_{df_i}} - Y_{k_{dfe_i}} + Q_{k_{df_{i-1}}}, \forall k \in \{1,...,n_{df_i}\}$$

are respectively the number of planes arriving and leaving which are delayed during a time interval 'i=15 min'.

To optimize the function of penalty thus means to minimize the function φ_F where :

$$\varphi_F = \sum_{i=1}^{N} [\sum_{j=1}^{n_{af}}(X_{j_{af_i}} - X_{j_{afe_i}}) + \sum_{k=1}^{n_{df}}(Y_{k_{df_i}} - Y_{k_{dfe_i}})] \tag{27}$$

V. Algorithms

The resolution and the optimization of a $CSCAP$ problem proceed in several stages:

1) Initially they consist of the distribution of the arrival demands of an airport on its arrival fixes and of the distribution of the departure demands on its departure fixes.
2) In the second time they consist to select the pertinent couple of arrival and departure capacity (value) (C_{a_i}, C_{d_i}) among the various couples of capacities (dictated by airport managers) to affect at the same time to arrival and departure demands (variables) of the different fixes. As mentioned above, this choice is selected according to the heuristic which allows to respect all the constraints and to minimize the number of delayed planes.
3) In the third time they consist of the distribution of this pertinent couple of arrival and departure capacity (value) (C_{a_i}, C_{d_i}) found above respectively on the capacities of arrival fixes $C_{1_{af_i}}$, ..., $C_{n_{af_i}}$ and on the capacities of departure fixes $C_{1_{df_i}}$, ..., $C_{n_{df_i}}$.
4) Finally they consist of the assignment of these capacities respectively to the variables $X_{j_{ae_i}}$ for j=1, ..., n_{af_i} and $Y_{k_{de_i}}$ for k=1, ..., n_{df_i} so that all the constraints (5)-(23) are satisfied and that $\sum_{i=1}^{N}[\sum_{j=1}^{n_{af_i}} Q_{j_{af_i}} + \sum_{k=1}^{d_{af_i}} Q_{k_{df_i}}]$ is minimized. This stage proceed as mentioned in the next algorithm.

Algorithm 1: Distribution Algorithm

```
Procedure BLSTL(I, (X_ai, Y_di)_l, a)
  I = I ⋃ {((X_ai, Y_di)_l, a)}
  If ((I checks all the constraints)  AND
  (∑p(d)=Min)  AND (T_execution < T_fixed))
  then
    If ((X_ai, Y_di)_l = (X_ai, Y_di)_n) then
      | I is an optimal solution
    else
      For all b ∈ (D_X_ai, D_Y_di) do
        BLSTL(I,(X_ai, Y_di)_l+1, b)
    end If
  else
    If (T_execution ≥ T_fixed) then
      | Perform Local Search Algorithm
    end If
  end If
End
```

Algorithm 2: BLSTL Algorithm

We called this algorithm BLSTL for (Backtrack + Local Search + Time Limit constraint).

Note that I is the instantiation for intervals 1, 2, ..., i-1; a and b are values in $(D_{X_{a_i}}, D_{Y_{d_i}})$; $(X_{a_i}, Y_{d_i})_l$ is the l^{th} variable; $(X_{a_i}, Y_{d_i})_n$ is the last variable and Min is the best number of delayed flights found so far.

The BLSTL algorithm consists in recursively testing all possible couples of arrival and departure capacity in $(D_{X_{a_i}}, D_{Y_{d_i}})$ dictated by the airport manager to find the minimum delay.

VI. Implementation and Experimental results

We implemented this method with Java (JDK1.5) on an Athlon 1.67GHz with 512MB of RAM. We performed preliminary experiments of the proposed approach and the results obtained for the various instances of the problem are promising.

Note that the local search can be used only if we fixed in advance the $T_{execution}$. In addition, our method presented in [16] find the optimal solution, but spent much time before delivering it, especially when the number of time intervals is greater or equal than 17. So, the local search can be used only when the number of time intervals is greater or equal than 17. For the numbers equal to 16, the execution time is around 2 minutes. For the numbers equal to 12, the execution time is around 60 seconds.

We present below a $CSCAP$ problem and some of its optimal solutions satisfying constraints (5)-(23) and minimizing

$$\sum_{i=1}^{N} \left[\sum_{j=1}^{n_{af}} Q_{j_{af_i}} + \sum_{k=1}^{n_{df}} Q_{k_{df_i}} \right]$$

We consider that there are 4 arrival fixes noted af1, af2, af3 and af4 and 4 departure fixes noted df1, df2, df3 and df4 at the airport (figure 1). In this example, the period is 3 hours (12h00-15h00) and the number of intervals N forming this period is equal to 12.

The Arrival/Departure predicted demands are those given in table I and the available capacities are values (domain) given as an example in subsection B of section III ($D = \{(18,29),(24,24),(26,19),(28,15),(17,30),(20,27)\}$).

In tables II, IV, VI, III, V and VII, column 0 shows the time intervals of 15 minutes. For each interval, column 1A/D shows the number of Arrival/Departure predicted demands X_a/Y_d of the airport. These arrival/departure predicted demands of the airport will be distributed among the arrival/departure predicted demands of its fixes $X1_{af_i}/Y1_{df_i}$, $X2_{af_i}/Y2_{df_i}$, $X3_{af_i}/Y3_{df_i}$ and $X4_{af_i}/Y4_{df_i}$ at each interval of time i for i = 1, 2, 3, ..., 12. We do not show explicitly this table. But it is easy to deduct it for each couple of tables (II, III), (IV, V) and (VI, VII). For each fix it's the sum of the number of the planes delayed added to the number of the planes effectively served.

In tables II, IV and VI, columns 2A/D, 3A/D, 4A/D and 5A/D present respectively arrival/departure flows effectively served (or number of planes effectively arrived/leaved) X_1/Y_1, X_2/Y_2, X_3/Y_3 and X_4/Y_4 at the fixes af1/df1, af2/df2, af3/df3 and af4/df4.

In tables III, V and VII, columns 7A/D, 8A/D, 9A/D and 10A/D present respectively the arrival/depature demands which were not served (queues or delays) $Q1_a/Q1_d$, $Q2_a/Q2_d$, $Q3_a/Q3_d$ and $Q4_a/Q4_d$ at the arrival/departure fixes af1/df1, af2/df2, af3/df3 and af4/df4. The arrival/departure demands which were not served are queues which must be taken into account during the next time interval.

The problem consists in distributing the values of arrival/departure predicted demands of the airport (column 1A/D of each given table) on the arrival/departure predicted demands of its fixes $X1_{af_i}/Y1_{df_i}$, $X2_{af_i}/Y2_{df_i}$, $X3_{af_i}/Y3_{df_i}$ and $X4_{af_i}/Y4_{df_i}$.

Then, it acts of the distribution of the arrival/departure capacities of the terminal (column 6A/D of each given table) on the capacities of arrival/departure fixes af1/df1, af2/df2, af3/df3 and af4/df4 which are in columns 2A/D, 3A/D, 4A/D and 5A/D of tables II, IV and VI. Note that this distribution is for instance arbitrary.

But before that, let us notice, for each line, which capacity (value) to select among the various capacities and to affect to arrival and departure demands (variables) of the different fixes. This choice is selected according to the heuristic which allows to respect all the constraints and to minimize the total number of delayed planes. The resolution is done in an incremental way and follows the law of line per line.

In columns 7A/D, 8A/D, 9A/D and 10A/D of tables III, V and VII we record at each time interval i the number of the planes which had required to arrive/leave at the one of the arrival/departure fixes af1/df1, af2/df2, af3/df3 and af4/df4 but which has not been served. These demands which has not been served will respectively constitute arrival/departure queues for these same fixes for the next time interval i+1.

In a general way, we seek to reduce at the minimum the total sum of all these planes delayed.

The function of penalty φ which is the sum of the numbers of planes delayed (columns 7A/D, 8A/D, 9A/D and 10A/D) at the different arrival/departure fixes will be written:

$$\varphi = \sum_{i=1}^{12} \left(\sum_{j=1}^{4} Q_{j_{a_i}} + \sum_{k=1}^{4} Q_{k_{d_i}} \right) \quad (28)$$

That is to say :

$$\varphi_F = \sum_{i=1}^{12} \left(\sum_{j=1}^{n_{af}} (X_{j_{a_i}} - X_{j_{afe_i}}) + \sum_{k=1}^{n_{df}} (Y_{k_{d_i}} - Y_{k_{dfe_i}}) \right) \quad (29)$$

0	1A/D	2A/D	3A/D	4A/D	5A/D	6A/D
	X_a/Y_d	X_1/Y_1	X_2/Y_2	X_3/Y_3	X_4/Y_4	C_a/C_d
1	26/36	08/09	06/09	06/06	04/00	24/24
2	38/32	09/08	04/00	00/00	11/16	24/24
3	42/09	00/03	14/10	00/08	10/03	24/24
4	29/15	20/06	00/00	08/08	00/01	28/15
5	06/07	01/01	09/04	10/04	08/03	28/15
6	13/10	03/02	04/04	16/02	03/02	28/15
7	14/17	03/04	03/04	03/04	05/05	26/19
8	20/33	05/08	02/06	05/08	05/08	17/30
9	40/34	10/08	13/11	00/00	01/05	24/24
10	25/22	03/05	06/05	00/04	15/10	24/24
11	13/13	06/03	00/04	15/14	03/03	24/24
12	12/01	03/01	07/00	08/00	03/00	28/15
TT	278/229	71/58	68/57	71/58	68/56	299/253

TABLE II

Real flows recorded on the different fixes of the airport.

0	1A/D	2A/D	3A/D	4A/D	5A/D	6A/D
	X_a/Y_d	X_1/Y_1	X_2/Y_2	X_3/Y_3	X_4/Y_4	C_a/C_d
1	26/36	08/09	06/09	04/00	06/06	24/24
2	38/32	06/05	09/08	00/00	09/11	24/24
3	42/09	00/03	10/02	14/19	00/00	24/24
4	29/15	00/06	00/01	14/03	14/05	28/15
5	06/07	15/04	09/06	00/01	04/01	28/15
6	13/10	15/02	04/04	04/02	03/02	28/15
7	14/17	03/04	03/04	05/05	03/04	26/19
8	20/33	02/06	05/08	05/08	05/08	17/30
9	40/34	13/13	10/08	00/00	01/03	24/24
10	25/22	07/07	01/04	16/13	00/00	24/24
11	13/13	03/03	08/04	03/03	10/14	24/24
12	12/01	03/00	03/00	03/00	12/01	28/15
TT	278/229	75/62	68/58	68/54	67/55	299/253

TABLE IV

Real flows recorded on the different fixes of the airport.

0	1A/D	7A/D	8A/D	9A/D	10A/D
	X_a/Y_d	Q_{1_a}/Q_{1_d}	Q_{2_a}/Q_{2_d}	Q_{3_a}/Q_{3_d}	Q_{4_a}/Q_{4_d}
1	26/36	00/00	00/00	00/03	02/09
2	38/32	00/00	05/08	11/11	00/01
3	42/09	12/00	01/00	21/05	00/00
4	29/15	00/00	08/03	20/00	07/02
5	06/07	00/00	00/00	13/00	00/00
6	13/10	00/00	00/00	00/00	00/00
7	14/17	00/00	00/00	00/00	00/00
8	20/33	00/00	03/03	00/00	00/00
9	40/34	00/00	00/00	10/08	09/05
10	25/22	03/00	00/00	17/11	00/00
11	13/13	00/00	04/00	05/00	00/00
12	12/01	00/00	00/00	00/00	00/00
TT	278/229	15/00	21/14	97/38	18/17

TABLE III

Real delays recorded on the different fixes of the airport.

0	1A/D	7A/D	8A/D	9A/D	10A/D
	X_a/Y_d	Q_{1_a}/Q_{1_d}	Q_{2_a}/Q_{2_d}	Q_{3_a}/Q_{3_d}	Q_{4_a}/Q_{4_d}
1	26/36	00/00	00/00	02/09	00/03
2	38/32	05/03	00/00	11/17	00/00
3	42/09	17/03	00/00	07/00	10/02
4	29/15	24/00	08/05	00/00	03/00
5	06/07	12/00	00/00	01/00	00/00
6	13/10	00/00	00/00	00/00	00/00
7	14/17	00/00	00/00	00/00	00/00
8	20/33	03/03	00/00	00/00	00/00
9	40/34	00/00	00/00	10/08	09/05
10	25/22	00/00	05/01	00/00	15/10
11	13/13	00/00	00/00	00/00	09/00
12	12/01	00/00	00/00	00/00	00/00
TT	278/229	61/09	13/06	31/34	46/20

TABLE V

Real delays recorded on the different fixes of the airport.

The number of the arrival planes delayed in table III is: $\sum_{i=1}^{12}(\sum_{j=1}^{n_{af}=4}(X_{j_{af_i}} - X_{j_{afe_i}})) = 15+21+97+18 = 151$ (only the numbers on the left of columns 7A/D, 8A/D, 9A/D and 10A/D), and
The number of the departure planes delayed in table III is: $\sum_{i=1}^{12}(\sum_{k=1}^{n_{df}=4}(Y_{k_{df_i}} - Y_{k_{dfe_i}})) = 00+14+38+17 = 69$ (only the numbers on the right of columns 7A/D, 8A/D, 9A/D and 10A/D).

The total number of planes delayed is $151+69 = 220$. It is an optimal value.

The number of the arrival planes delayed in table V is: $\sum_{i=1}^{12}[\sum_{j=1}^{n_{af}=4}(X_{j_{af_i}} - X_{j_{afe_i}})] = 61+13+31+46 = 151$ (only the numbers on the left of columns 7A/D, 8A/D, 9A/D and 10A/D), and
The number of the departure planes delayed in table V is: $\sum_{i=1}^{12}[\sum_{k=1}^{n_{df}=4}(Y_{k_{df_i}} - Y_{k_{dfe_i}})] = 09+06+34+20 = 69$ (only the numbers on the right of columns 7A/D, 8A/D, 9A/D and 10A/D).

The total number of planes delayed is $151+69 = 220$. It is also an optimal value.

The number of the arrival planes delayed in table VII is: $\sum_{i=1}^{12}[\sum_{j=1}^{n_{af}=4}(X_{j_{af_i}} - X_{j_{afe_i}})] = 36+28+38+49 = 151$ (only the numbers on the left of columns 7A/D, 8A/D, 9A/D and 10A/D), and
The number of the departure planes delayed in table VII is: $\sum_{i=1}^{12}[\sum_{k=1}^{n_{df}=4}(Y_{k_{df_i}} - Y_{k_{dfe_i}})] = 21+18+03+27 = 69$ (only the numbers on the right of columns 7A/D, 8A/D, 9A/D and 10A/D).

The total number of planes delayed is $151+69 = 220$. It is also an optimal value.

VII. Control and regulation of capacities

We saw in inequality (5) that
$0 \leq P_{v_i} + P_{oc_i} \leq C_{T_i}$.

Let us suppose that one wishes to reduce or increase the capacities of an airport and of its fixes for a some reason. In otherwise, one wishes to control and regulate these capacities. Then in this case it is necessary to introduce some control parameters. The formula above becomes :

$$\alpha_i P_{oc_i} + \beta_i P_{v_i} = \lambda_i C_{T_i} \tag{30}$$

0	1A/D	2A/D	3A/D	4A/D	5A/D	6A/D
	X_a/Y_d	X_1/Y_1	X_2/Y_2	X_3/Y_3	X_4/Y_4	C_a/C_d
1	26/36	08/09	04/00	06/09	06/06	24/24
2	38/32	09/08	06/08	09/08	00/00	24/24
3	42/09	10/02	14/11	00/03	00/08	24/24
4	29/15	00/01	00/03	00/00	28/11	28/15
5	06/07	10/06	00/01	16/04	02/01	28/15
6	13/10	03/02	12/02	08/04	03/02	28/15
7	14/17	03/04	05/05	03/04	03/04	26/19
8	20/33	05/08	05/08	05/08	02/06	17/30
9	40/34	00/00	10/08	10/08	04/08	24/24
10	25/22	00/02	06/05	06/05	12/12	24/24
11	13/13	16/14	00/03	00/03	08/04	24/24
12	12/01	06/00	06/01	06/00	03/00	28/15
tt	278/229	70/56	68/55	69/56	71/62	299/253

TABLE VI

Real flows recorded on the different fixes of the airport.

0	1A/D	7A/D	8A/D	9A/D	10A/D
	X_a/Y_d	Q_{1_a}/Q_{1_d}	Q_{2_a}/Q_{2_d}	Q_{3_a}/Q_{3_d}	Q_{4_a}/Q_{4_d}
1	26/36	00/00	02/09	00/00	00/03
2	38/32	00/00	05/09	00/00	11/11
3	42/09	00/00	01/00	12/00	21/05
4	29/15	07/02	08/00	19/03	01/00
5	06/07	00/00	09/00	04/00	00/00
6	13/10	00/00	00/00	00/00	00/00
7	14/17	00/00	00/00	00/00	00/00
8	20/33	00/00	00/00	00/00	03/03
9	40/34	10/08	00/00	00/00	09/05
10	25/22	16/11	00/00	00/00	04/00
11	13/13	03/00	03/00	03/00	00/00
12	12/01	00/00	00/00	00/00	00/00
TT	278/229	36/21	28/18	38/03	49/27

TABLE VII

Real delays recorded on the different fixes of the airport.

The parameters α_i, β_i and λ_i control the capacities. They depend on the wishes of regulation expressed by the airport managers according to certain conditions like weather, priorities, authorities, rush hours, etc. They have values between 0 and 1 such that $\alpha_i + \beta_i = 1$.

For example, if $\alpha_i = 0.5$, $\beta_i = 0.5$ and $\lambda_i = 0.25$, we will have $0.5P_{oc_i} + 0.5P_{v_i} = 0.25C_{T_i}$. This formula becomes $P_{oc_i} + P_{v_i} = 0.5C_{T_i}$. The total capacity is not entirely available: meaning that only half of the airport capacity is available. This restriction can occur because of certain conditions like weather, authorities, etc. It is thus necessary to adapt these new capacities for the arrival and departure demands.

We are thus brought to introduce these parameters into all the formulas which we put until now. More precisely, we will have the following formulas:

$$\alpha_i X_{ae_i} \le \lambda_i C_{a_i} \le \beta_i P_{v_i} \tag{31}$$

$$\alpha'_i Y_{de_i} \le \lambda'_i C_{d_i} \le \beta'_i P_{oc_i} \tag{32}$$

Such that C_{a_i} and C_{d_i} are the total capacities respectively of all arrival and all departure fixes (figure 1) during the time interval 'i' and α'_i, λ'_i and β'_i are parameters of control which can be respectively equal or different of α_i, β_i and λ_i.

Let us note that $C_{a_i} \le C_{T_{a_i}}$ and that $C_{d_i} \le C_{T_{d_i}}$ where $C_{T_{a_i}}$ and $C_{T_{d_i}}$ are respectively the arrival and departure capacities of the terminal 'T' during 'i' such as $C_{T_{a_i}} + C_{T_{d_i}} = C_{T_i}$. We point out that C_{T_i} is the total capacity of the terminal.

Over one period forming by N intervals, we will have :

$$\begin{aligned}\sum_{i=1}^{N}[\alpha_i X_{ae_i} + \alpha'_i Y_{de_i}] &\le \sum_{i=1}^{N}[\lambda_i C_{a_i} + \lambda'_i C_{d_i}] \\ &\le \sum_{i=1}^{N}[\beta_i P_{v_i} + \beta'_i P_{oc_i}]\end{aligned} \tag{33}$$

Concretely if the capacity is rather large so that no congestion takes place, the problem would not arise. But if there are more arrival demands than departure demands during one time period, the problem of the congestion appears. In this case, it is necessary to minimize the number of delayed planes or to maximize departure flows as well as possible.

The following formula allows us to solve partly this problem of congestion and to ensure a certain balance between arrivals and departures.

$$\sum_{i=1}^{N}[\alpha_i P_{oc_i} - \gamma_i Y_{de_i} + \beta_i P_{v_i} - \delta_i X_{ae_i}] \le \sum_{i=1}^{N}[\lambda_i C_{T_{d_i}} + \mu_i C_{T_{a_i}}] \tag{34}$$

N represents the number of intervals over one period of (15*N) minutes;
P_{oc_i} is the number of occupied places during the time interval 'i';
P_{v_i} is the number of empty places during the time interval 'i';
Y_{de_i} is the number of effective departures during 'i';
X_{ae_i} is the number of effective arrivals during 'i';
$C_{T_{d_i}}$ is the departure capacity at the terminal during 'i';
$C_{T_{a_i}}$ is the arrival capacity at the terminal during 'i';
α_i, β_i, γ_i, δ_i, λ_i and μ_i are parameters of regulation.

The response to this load balancing problem between arrivals and departures consists in minimizing the following function which we call the new objective function.

$$min \ \{\sum_{i=1}^{N}[\alpha_i P_{oc_i} - \gamma_i Y_{de_i}] + [\beta_i P_{v_i} - \delta_i X_{ae_i}]\} \tag{35}$$

We will treat in details these concepts of optimization in a forthcoming paper.

Let us note that our approaches of modeling, resolution and optimization presented in this paper remain valid as well for the air traffic as for the road traffic, railway and maritime.

Another possible solution to reduce these problems of congestion of the airports (or any other system) is to carry out a certain cooperation and complementarity between the various systems in particular air, maritime, road and rail transport.

Each mode of transport has its own advantages e.g. potential capacity, high levels of safety, flexibility, low energy consumption, low environmental impact. Intermodal transport allows each mode to play its role in building transport chains which overall are more efficient, cost effective and sustainable [5].

VIII. Conclusion

In this paper, we presented a modeling of the capacity allocation problem of an airport (CAP) and of its fixes ($FCAP$) in form of a constraint satisfaction problem (CSP) called $CSCAP$. Thereafter, we proposed an optimization model under constraints for the capacity allocation problem of the fixes. The method presented in [16] find the optimal solution, but spent much time before delivering it, especially when the number of time intervals is greater or equal than 17. To deal with this problem, we proposed an improvement of this method by grafting to it a method of Local Search (LS) combined with the Backtrack (B) beyond a certain Time Limit constraint (TL) fixed by expert. Thus we have constantly, beyond fixed time, at least a solution which perhaps can not be the optimal solution, but which can be acceptable. We called this new method BLSTL (Backtrack + Local Search + Time Limit constraint). We showed an example of illustration implementing our procedures of modeling and optimization. We also showed how to control and regulate these capacities in order to assist the managers to use the best allocation of capacities and to deal with unforseen events. The experimental results are shown that our approach gives results which are promising.

References

[1] Adams M., S. Kolitz, J. Milner, A.R. Odoni, “Evolutionary concepts for decentralized air traffic flow management”. Air Traffic Control Quarterly 4, (1996).

[2] Andreatta G., L. Brunetta, “Multiairport ground holding problem: A computational evaluation of exact algorithms”. In Operations Research 46 (1998).

[3] Andreatta G., Brunetta L. and Guastalla G., “The flow management problem: Present computational algorithm”. In Proceedings of the Eighth IFAC Symposium on Transportation System, Chania, (1997).

[4] “Airport Capacity Benchmark”. Report FAA, May (2001). http://www.faa.gov/events/benchmarks

[5] “Rail Air Intermodality Facilitation Forum”. Final Report of the European Commission, intermodal transport, Volume 1, (2004). http://europa.eu.int/comm/transport/intermodality/raiff

[6] Dell’Olmo P., G. Lulli. “A Dynamic Programming Approach for the Airport Capacity Allocation Problem”. Technical Paper, Dip. di Statistica, Probabilità e Statistiche Applicate, Univ. di Roma, (2003).

[7] Dell’Olmo P., G. Lulli, “A new hierarchical architecture for Air Traffic Management: Optimisation of airway capacity in a Free Flight scenario”. European Journal of Operational Research 144, (2003).

[8] Duong Vu, “Free route experimental encounter resolution : Initial results”. Technical Paper, Eurocontrol, Bretigny, (1999).

[9] Fearnsides J. J., “Free flight as a system concept”. In Proceedings of CAAASD, (1995).

[10] Gilbo E.P. and K. Howard, “Collaborative Optimization of Airport Arrival and Departure Traffic Flow Management Strategies for CDM”. 3rd USA/Europe Air Traffic Management R&D Seminar, Italy, (2000).

[11] Gilbo E. P., “Arrival/Departure Capacity Tradeoff Optimization: a Case Study at the St. Louis Lambert International Airport (STL)”. 5th USA/Europe Air Traffic Management R&D Seminar, Budapest, (2003).

[12] Hamzawi S. G., “Lack of airport capacity: Exploration of alternative solutions”. Transportation Res.-A, vol. 26-A, no. 1, (1992).

[13] Hoffman R. L., M.O. Ball. “A comparison of formulations for the single airport ground holding problem with banking constraints”. Operations Research 48, 578-590, (2000).

[14] Idrissi A., “CASC : Aide à la décision au contrôleur aérien par satisfaction de contraintes”. In proceedings of International Conference on Computer Sciences, RIVF’05, Can Tho, (2005).

[15] Idrissi A. and Chu Min Li, “Modeling Constraint Optimization Problems: Airport Capacity Allocation as a Case Study”. In proceedings of International Conference on Modeling and Simulation, ICMS’05, Marrakech, (2005).

[16] Idrissi A. and Chu Min Li, “Modeling and Optimization of the Capacity Allocation Problem using Constraints”. In proceedings of IEEE International Conference on Computer Sciences, RIVF’06, Ho Chi Minh City. University of Technology. pp. 107-116, (2006).

[17] Jenny M., “Foreword”. Air Traffic Control Quarterly 5, (1997).

[18] Korf R., “Linear-space best first search”. Artificial Intelligence, 62:41-78, (1993).

[19] Larrosa J. and P. Meseguer, “Exploiting the use of DAC in max-csp”. In proceedings of CP-96, Cambridge, (1996).

[20] Li Chu Min, F. Manya and J. Planes, “Exploiting Unit Propagation to Compute Lower Bounds in Branch and Bound Max-SAT Solvers”. In proceedings of CP-05, (2005).

[21] Montanari U., “Networks of constraints : Fundamental properties and applications to picture processing”. Information Sciences, 7: 95-132, (1974).

[22] Odoni A. R., J. Bowman, et al., “Existing and required modeling capabilities for evaluating ATM system and concepts”. Final Report International Centre for Air Transportation, MIT, Cambridge, MA, (1997).

[23] Schiex T., H. Fargier and G. Verfaillie, “Valued constraint satisfaction problems: hard and easy problems”. In proceedings of IJCAI-95. Montréal, Canada, (1995).

[24] Selman B., H. Kautz and B. Cohen, “Noise Strategies for local Search”. In proceedings of 12^{th} National Conference on Artificial Intelligence (AAAI’94), pp. 337-343, (1994).

[25] Wambsganss M. C., “Collaborative decision making air traffic management”. In L. Bianco, P. Dell’Olmo, A.R. Odoni (Eds.), New Concepts and Methods in Air Traffic Management, Springer, Berlin, (2001).

[26] Zhang W., “Modeling and analyzing soft constraint optimization : Ressource allocation as a case study”. Technical report WU-CIC Working Note 1, Computational Intelligence Center, Washington University, St Louis, MO 63130, (2001).

Airspace block organization with metaheurisitics and partitioning packages

Charles-Edmond Bichot and Nicolas Durand
Laboratoire d'optimisation globale
ENAC - DSNA/DTI-SDER — Toulouse, FRANCE
{lastname}@recherche.enac.fr

***Abstract*— In this paper, different metaheuristics applied on an air traffic control problem. This problem is a graph partitioning problem. It can be solved by classical methods which are spectral and multilevel methods. State-of-the-art public-domain graph partitioning packages, CHACO and METIS are used to resolve it. A comparison between results return by these packages and metaheuristics implementations is made for different objective functions of the literature. Metaheuristics used are simulated annealing, ant colony and a new one called fusion fission developed in the LOG laboratory. Experimental results show that metaheuristics find better results than classical packages.**

I. Introduction

In this paper, different metaheuristics applied on an air traffic control problem are presented. This problem is a graph partitioning problem. It can be solved by classical methods which are spectral and multilevel methods. State-of-the-art public-domain graph partitioning packages, CHACO [1] and METIS [2] are used to resolve this problem. A comparison between results return by these packages and metaheuristics implementations is made for different objective functions. However, whereas classical methods are designed for particular objective function, that can hardly be modified, metaheuristics can easily be changed of objective function, which is an advantage for this problem.

The air traffic control problem is presented in section II. Different objective functions already used to solve partitioning problems are described in section III. In section IV are shortly presented three metaheuristics, which are simulated annealing, ant colonies and a new one developed in the LOG laboratory, fusion fission. The comparison between results of metaheuristics and classical libraries is presented in section V.

II. An Air Traffic Control problem

"The primary purpose of the Air Traffic Control (ATC) system is to prevent a collision between aircraft operating in the system and to organize and expedite the flow of traffic" [3]. The first objective of ATC is safety, the second is efficiency.

The Functional Airspace Block Optimized Process (FABOP) project consists in dividing the European airspace into blocks. Let us explain some air traffic control mechanisms. An air traffic controller supervises the traffic in an limited area, called air traffic sector. Controllers are qualified for working on a set of sectors, which is called a functional airspace block. The FABOP project consists to partition the European airspace into functional airspace blocks. Actually, just a few blocks are crossing countries' frontiers. In this paper we study a new organization of blocks based on flows of aircraft and not on countries' frontiers.

Controllers only know air traffic sectors on which they are qualified, and rarely other sectors. Because "controller-controller coordination is easier and more effective inside an ATC unit (a block) than between ATC units" [4], *we try to maximize flow of aircraft inside blocks and to minimize flow of aircraft between blocks.* For human factors, *each block must have the same "size".* This is the aim of the objective function to minimize.

Let us describe more formerly the problem. We compare each air traffic sector to a vertex and each flow of aircraft between sectors to an edge. We have a graph $G = (V, E)$ with a vertex set V and edge set E. Each edge $e = (v_1, v_2)$ has a weight $w(e)$ which is the flow of aircraft which fly from v_1 to v_2 and *vice versa.* Each vertex v_i has a weight, which is the sum of the weights of connected edges plus aircraft departures and landings if the sector is connected with an airport. A partition of G into k distinct partitions must respect the following constraints : $P_k(G)$ must be a partition of G into k non-empty subgraphs $V_1, \ldots, V_k$ with $\forall\ i, j$ included in $1 \ldots k, i \neq j, V_i \cap V_j = \varnothing$ and $\bigcup_{i=1}^{k} V_i = V$.

The problem is to partition the vertices of the graph G into k roughly equal parts (or blocks), such that a certain objective function is optimized. Intuitively, equal parts are parts with the same number of vertices, *ie.* blocks should have the same number of sectors. Because the number (flow) of aircraft which go through a sector is very different for each sector, we can not use this equality between parts. More probably, comparing the amount of flows of aircraft between the sectors of a block with the same amount of flows with another block is better. The number of conflicts or potential conflicts in a sector increases considerably the difficulty for a controller to coordinate aircraft. Thus, the computation of the equality should take into account the number of conflicts or potential conflicts. The size and the design of a sector influence the number of coordinations, maybe the computation of the equality should take into account this too. Other operationals constraints can be easily find. As we can see, finding what is the equality criteria between parts is difficult.

In the experimental tests, the number of vertices is $|V| = 759$ and the number of edges is $|E| = 3,165$. This number of vertices is the number of air traffic sectors of the European countries core area defined in [5]. This area is the set of countries which have the highest flows of traffic in Europe. The countries core area is composed of Germany, France, United Kingdom, Switzerland, Belgium, Netherlands, Austria, Spain, Denmark, Luxembourg and Italy.

We have seen that finding equal parts for this problem is difficult. Indeed, like image segmentation, we should use objective functions which itself equalize the different parts of the partition.

III. Objective functions

In this section we present objective functions used for partitioning problems. The simplest and the oldest of them is the Cut function, which is the sum of edges' weight between partitions. This objective function is designed for spectral graph partitioning. Let $A \in P_k(G)$, $V - A = \{u \in V, u \notin A\}$, we define

$$cut(A, V - A) = \sum_{u \in A, v \in V - A} w(u, v) \tag{1}$$

and

$$W(A) = \sum_{u \in A, v \in A} w(u, v) \tag{2}$$

Thus,

$$Cut(P_k(G)) = \sum_{A \in P_k(G)} cut(A, V - A) \tag{3}$$

Hagen and Kahng [6] defined the ratio cut :

$$Rcut(P_k(G)) = \sum_{A \in P_k(G)} \left(\sum_{B \in P_k(G) - \{A\}} \frac{cut(A, B)}{|A|} \right) \tag{4}$$

which removes the requirement $|A| = |B|$ and minimizes $cut(A, V - A)$ when the number of vertices in each part is roughly equal.

Shi and Malik [7] propose the normalized cut :

$$Ncut(P_k(G)) = \sum_{A \in P_k(G)} \frac{cut(A, V - A)}{cut(A, V - A) + W(A)} \tag{5}$$

which minimizes $cut(A, V - A)$ while maximizing $Assoc(A, V) = cut(A, V - A) + W(A)$, the sum of the weights of each partition.

And the min-max cut function was introduced in [8] :

$$Mcut(P_k(G)) = \sum_{A \in P_k(G)} \frac{cut(A, V - A)}{W(A)} \tag{6}$$

which minimizes $cut(A, V - A)$ while maximizing $W(A)$ which is the sum of the weights of edges between vertices of the same partition.

The $Mcut$ objective function seems the most appropriate of the objective functions presented, regards to the objective presented in section II : maximizing flows of aircraft (edges weight) inside blocks (partitions) and minimizing flow of aircraft between blocks.

IV. Metaheuristics

The three metaheuristics described in this section are more precisely explained in [9]. More information about a new method called fusion fission can be found in [10].

A. Simulated annealing

This metaheuristic was first introduced in [11]. Let us present very shortly the simulated annealing method. The fundamental idea is to allow moves resulting in solutions of worse quality than the current solution in order to escape from local minima.

The probability of doing such a move is decreased during the search. In metallurgy, a very hot metal is cooled very slowly to increase its solidity. In the same way, vertex are moved among partitions, one by step. The temperature T is decreased when moves are increasing the objective function ($T\frac{t_{max}-t_{min}}{t_{max}}$ with $t_{max}(t_{min})$, maximal (minimal) temperature). The objective function e gives the "energy" of the solution. The result of a move of a vertex which changes of partition, builds a new move s'. s' is kept if $\exp^{\frac{e(s)-e(s')}{T}}$ is upper a random number in $[0, 1]$ or if the objective function is increasing. Else, the old state s is kept. The algorithm stops when $T \leq t_{min}$.

B. Ant colony

Ant colony optimization is inspired by the foraging behavior of real ants. This metaheuristic proposed by Dorigo is explained in [12]. The algorithm uses the ability of ants to find the shortest path between food source and their nest. While walking from food sources to the nest and *vice versa*, ants deposit a substance called pheromone on the ground. When they decide a direction to go, they choose with a higher probability paths that are marked by stronger pheromone concentrations.

The ant colony algorithm is based on three (one optional) steps. While a termination condition is not satisfied, the three steps are executed, but not necessarily in the following order :

- The first step is the ants motion. Ants are moving through nodes of the graph G by applying a stochastic local decision policy which uses pheromone values and a local heuristic. While moving, the ant keeps memory of the path it was walking on the graph.
- The second step consists in updating pheromones. Ants always update the pheromone trails they are using. But if a path lead to "food" (a local solution), the ant can update backward the path it used by using its memory. Finally, like real pheromones, pheromone trail intensity decreases over the time (to avoid convergence into a sub-optimal region).
- The last step is optional. It is used to implement centralized actions which cannot be performed by single ants.

The adaptation of ant colony to the partitioning problem uses k colonies, one for each partition of the graph. These colonies are competing for food. It is important to notice that an ant can only distinguish pheromones of its colony. The weights of vertices correspond to ants food. Ants put down pheromones on edges. A vertex is owned by a colony if the sum of his pheromones on adjacent edges is greater than for other colonies. A local heuristic forces ants to explore edges which have no pheromone. Ants can go where they want, so ants from different colonies can be in the same vertex at the same step. Thus, the connectivity of sets is not forced. As connected sets often produces best results, we do not need to force this connectivity.

To conclude, the approach of k-partitioning with ant colony algorithm is very different than precedent works [13], [14].

C. Fusion fission

The fusion fission method is a new method developed at the LOG laboratory. Its aim is graph partitioning. The fusion fission method is inspired by nuclear fusion and fission. The organization of a molecule can be compared to a k-partition. The molecule is the graph G, nucleons are vertices and atoms are partitions. Fusion and fission are two processes which organize atoms. Fusion is the process which assembles two atoms in one. On the contrary, fission is the process which breaks atoms into two parts. The fusion fission method consists in assembling and breaking atoms successively. Like in the natural process, fusion and fission can eject alone vertices (atoms) from a partition. Such vertices can be added to a different partition, or cut another partition in two (chained fission).

Because the number of partitions continuously change, the energy (returned by the objective function) is smaller for a small number of partitions. Thus we use a function to increase the value of the objective function if the number of partitions is not k.

The fusion fission algorithm is presented figure 1. $cpart$ is the partition of G into sets of vertices. $cpart$ is initialized with a near k-partition of G. t is the temperature. The higher the temperature, the easier the fusion of big atoms and the easier the fission of small atoms. $part_i$ is a partition of $cpart$. $part_i$ is randomly chosen between all partitions of $cpart$. A new partition $npart$ of G is made by $fusion$ or $fission$. The $laws$ correspond to probability lists, which are probabilities to eject zero, one, two or three nucleons as described upper. Then nucleons (in ln, a list of nucleons) are added to another partition, or, in the fission case, if the temperature is high (*high_energy* function), these nucleons can cut partitions. The *laws* are updated if the result of the objective function is lower than the preceding partition. The function which decreases temperature is $decrease(t) = t\ \frac{t_{max}-t_{min}}{nbt}$ where nbt is the number of steps of the temperature decrease. The best result is memorized in *best_part*. If the temperature is lower than a minimal temperature, the algorithm is run with the best partition *best_part* and the upper temperature. Else, it is run with the partition *npart* find and the new temperature *new_t*.

Algorithm 1 Fusion / Fission

```
cpart ⇐ initial_partion;
t ⇐ t_max;
while Stop condition is not fulfilled do
  part_i ⇐ choose_partition(cpart);
  if choice(part_i) < random number then
    — fusion —
    (npart, ln) ⇐ fusion(part_i, cpart, laws);
    for all n ∈ ln do
      npart ⇐ nfusion(n, npart, laws);
  else
    — fission —
    (npart, ln) ⇐ fission(part_i, cpart, laws);
    for all n ∈ ln do
      if high_energy(n, t) then
        npart ⇐ nfission(n, npart, laws);
      else
        npart ⇐ nfusion(n, npart, laws);
  new_laws ⇐ update(laws, t);
  new_t ⇐ decrease(t);
  if Energy(npart) < Energy(cpart) and
  Energy(npart) < Energy(best_part) then
    best_part ⇐ npart;
  if low_temperature(t) then
    cpart ⇐ best_part; t ⇐ t_max;
  else
    cpart ⇐ npart; t ⇐ new_t;
```

The process of choice between fusion and fission is function of the number of nucleons x of the atom choose :

$$choice(x) = \begin{cases} 1 & if\ x > n + \frac{1}{2\alpha(t)} \\ 0 & if\ x < n - \frac{1}{2\alpha(t)} \\ (x-n)\alpha(t) + \frac{1}{2} & else \end{cases}$$

where $n = \frac{nbv}{k}$, nbv is the number of vertices of the graph, k is the number of partitions, and $\alpha(t) = q\frac{t_{max}-t}{t_{max}-t_{min}} + r$ where q, r are adjusted by the user.

TABLE I
OBJECTIVE FUNCTIONS RESULTS FOR EACH ALGORITHM

	Cut	$Rcut$	$Ncut$	$Mcut$
CHACO Spectral	202,353	8504	22.31	75.45
CHACO Multi.	202,095	**8492**	22.42	76.93
METIS Multi.	208,224	9962	22.76	80.49
S. annealing	203,946	9385	22.34	74.44
Ant colony	203,308	9689	22.30	74.22
Fusion Fission	**197,955**	8508	**21.83**	**69.03**

V. RESULTS

Classical partitioning methods were first used to solve the air traffic control problem. Kenighan and Lin [15] have created a very efficient algorithm, which uses a local optimization strategy. Spectral methods have been popularized by the work of Pothen, Simon and Liu [16]. Multilevel methods have been developed by Hendrickson and Leland [17] and by Karypis and Kumar [18]. In this example, the CHACO [1] library was used. This library includes Kernighan-Lin algorithm, a spectral method and the multilevel method of Hendrickson and Leland. The METIS [2] library was also used. This library includes the multilevel method of Karypis and Kumar.

CHACO's objective is to minimize the Cut objective function, and minimize the difference between the number of vertices in each partition. METIS's objective is to partition the vertices into k disjoint subsets such that the sum of the vertex weights in each subset is the same, and to minimize the Cut objective function. The metaheuristics algorithms use the $Mcut$ objective function which is the most appropriate for the air traffic problem (see section III). In table I the best results of the different methods are presented. These results are computed for a partition into 32 sets, on a 3 GHz Intel Pentium 4 running with Linux. The objective function of the three metaheuristics is $Mcut$. CHACO spectral method uses RQI/Symmlq eigen solver, with the octasection partitioning method and multiple Kernighan-Lin refinement. CHACO multilevel method uses the bisection partitioning method. METIS multilevel method uses *pmetis* with default parameters. *kmetis* was also tested but without better results. Note that, if spectral and multilevel algorithms use local refinement (Kernighan-Lin), metaheuristics do not. The best results for each objective functions are strewn in bold.

TABLE II
VARIANCES

	$\sigma(spb)$	$\sigma(vw)$	$\sigma(ewbb)$
CHACO Spectral	0.4	1448	1680
CHACO Multi.	0.4	1573	1685
METIS Multi.	8.5	384	779
S. annealing	11.4	1392	1778
Ant colony	12.8	1500	1901
Fusion Fission	15.7	2620	2567

Regards to this results, the fusion fission is the best algorithm (with $Rcut$ nearest the multilevel method). If METIS and CHACO software are extremely fast (a few seconds to compute), metaheuristics are running 30 minutes to give these results.

Figure 1 and 5 details the map of the partitioning result of the fusion fission algorithm corresponding to table I result. In the same way, figure 2, 6 and 3, 7 present the partitioning results of the CHACO spectral algorithm and the METIS multilevel algorithm. Figure 4 and 8 present real partitions of the European airspace on February 2002. The three maps are vertical cuts of the European countries core are, at flight level 380 (11,000 m). Regarding to blocks of other figures, blocks of figure 5 have disproportionate shapes and sizes. Especially, the block in the center of France, which has a lot of sectors (88, compared to an expected value which is around 23.8). This result is irrelevant for air traffic control, because no controller can be qualified for so many sectors.

Table II presents the variance $\sigma(spb)$ of the number of sectors per block, the variance $\sigma(vw)$ of the sum of vertex weights $W(A)$, per block, and the variance $\sigma(ewbb)$ of the sum of edges weights between blocks $cut(A, V-A)$, per block. Expected values corresponding to these variances are: $E(spb) = 23.8$, $E(vw) \simeq 2700$, and $E(ewbb) \simeq 3500$. Because of its goal, the CHACO software produces the smallest variance of the number of sectors per block. But with 1,446 aircraft to control for the lowest block and 10,346 aircraft for the highest, it seems that this result is irrelevant for air traffic control too. METIS multilevel partitioning methods have the lowest variance of $W(A)$ and $cut(A, V - A)$. These results correspond to METIS goal. Considering METIS results, the maximal number of sectors in a block is 48, which makes these results more relevant.

Thus, a new approach of the problem is necessary. Future work on the air traffic problem will deal with minimizing the $Mcut$ objective function, but with the following constraints :

1) $\sigma(spb) < 6$
2) $\sigma(vw) < 500$
3) $\sigma(ewbb) < 500$

This defines a multi-objective problem.

VI. Conclusion

An application of three metaheuristics (simulated annealing, ant colony and fusion fission) to an air traffic control problem was presented in this paper. This air traffic control problem is a k-partitioning problem. These metaheuristics were compared to the CHACO and the METIS libraries which are classical methods to solve partitioning problems (Kernighan-Lin, spectral and multilevel methods). We compare results of these methods to results of metaheuristics with the $Mcut$ function. Compared to all of the objectives functions, metaheuristics return better results, especially the fusion fission algorithm. But the best result found with fusion fission is irrelevant for air traffic control. The analysis of CHACO and METIS partitioning results showed that these results are not relevant either. Consequently, to be applied, a solution of the air traffic control problem must respect some constraints, while minimizing the $Mcut$ objective function.

References

[1] B. Hendrickson and R. Leland. *The Chaco user's guide*. Sandia National Laboratories, 2 edition, 1994.

[2] George Karypis and Vipin Kumar. *MeTis: A software package for partitioning*, 4 edition, 1998.

[3] Federal Aviation Administration (U.S. Department of Transportation). *Air Traffic ControL : FAA Order 7110.65K*, July 1997.

[4] Anders Hallgren. Restructuring european airspace: functional airspace blocks. *Skyway*, pages 20–22, autumn 2005.

[5] Charles-Edmond Bichot and Jean-Marc Alliot. A theoretical approach to defining the european core area. Technical report, LOG - ENAC/CENA, 2005.

[6] Lars Hagen and Andrew Kahng. New spectral methods for ratio cut partitioning and clustering. *IEEE Transactions on Computer-Aided Design*, 11(9):1074–1086, 1992.

[7] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[8] Chris H. Q. Ding, Xiaofeng He, Hongyuan Zha, Ming Gu, and Horst D. Simon. A min-max cut algorithm for graph partitioning and data clustering. In *Proceedings of ICDM 2001*, pages 107–114, 2001.

[9] Charles-Edmond Bichot, Jean-Marc Alliot, Nicolas Durand, and Pascal Brisset. Optimisation par fusion et fission. application au problème du découpage aérien européen. *Journal Européen des Systèmes Automatisés*, 38(9-10):1141–1173, 2004.

[10] Charles-Edmond Bichot. A metaheuristic based on fusion and fission for partitioning problems. In *9th International Workshop on Nature Inspired Distributed Computing, In conjunction with IEEE IPDPS 2006*, Rhodes Island, Greece, 2006.

[11] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, May 1983.

[12] M. Dorigo, V. Maniezzo, and A. Colorni. The ant system: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, 26(1):29–41, 1996.

[13] P. Kuntz, P. Layzell, and D. Snyers. A colony of ant-like agents for partitioning in vlsi technology. In *the Fourth European Conference on Artificial Life*, pages 417–424. MIT Press, 1997.

[14] A. E. Langham and P. W. Grant. A multilevel k-way partitioning algorithm for finite element meshes using competing ant colonies. In *the Genetic and Evolutionary Computation Conf.*, volume 2, pages 1602–1608, Orlando, Florida, USA, 1999.

[15] B. W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *The Bell System Technical Journal*, 49(2):291–307, 1970.

[16] Alex Pothen, Horst D. Simon, and Kang-Pu Liou. Partitioning sparse matrices with eigenvectors of graphs. *SIAM J. Matrix Anal. Appl.*, 11(3):430–452, 1990.

[17] Bruce Hendrickson and Robert Leland. A multi-level algorithm for partitioning graphs. In *Supercomputing*, 1995.

[18] George Karypis and Vipin Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, 20(1):359–392, 1998.

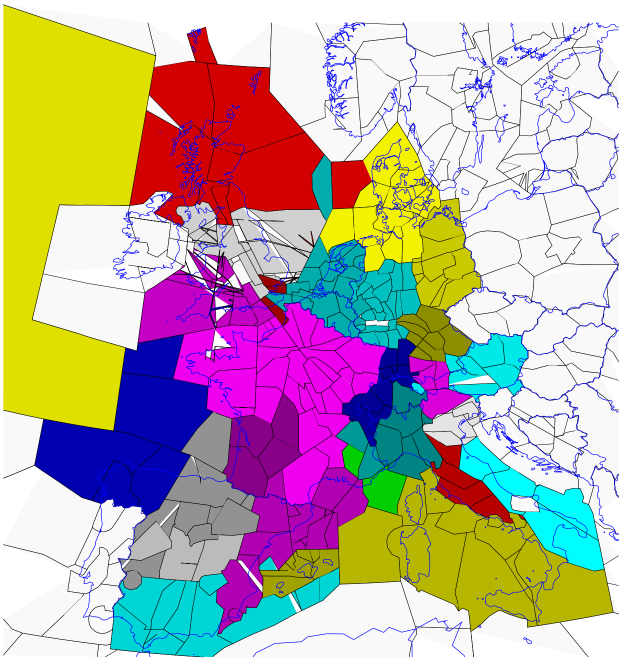


Fig. 1. Partitioning result of fusion fission (FL240)

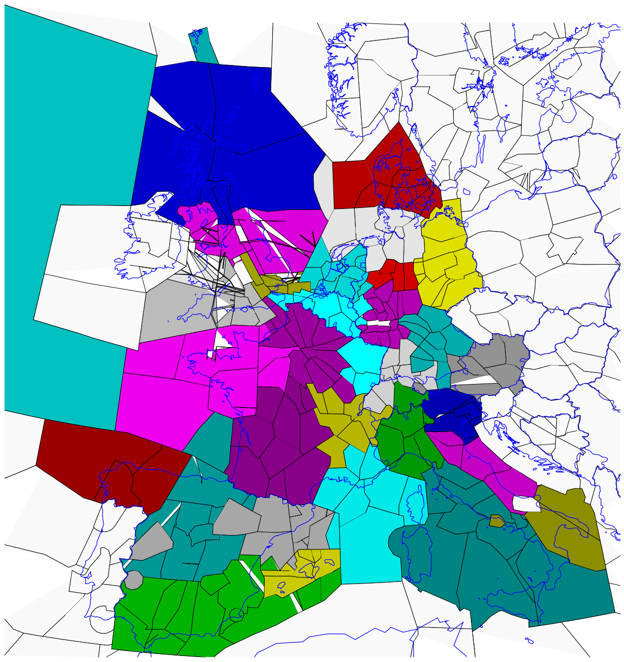


Fig. 2. Partitioning result of the CHACO package (FL240)

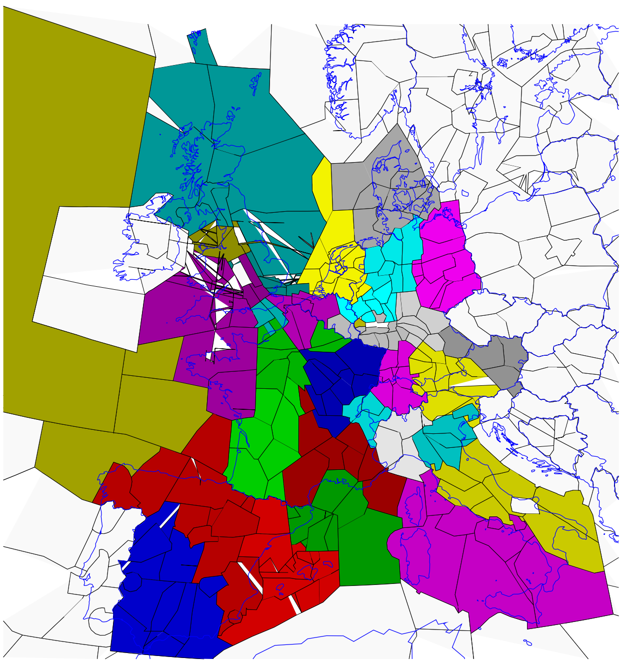


Fig. 3. Partitioning result of the METIS package (FL240)

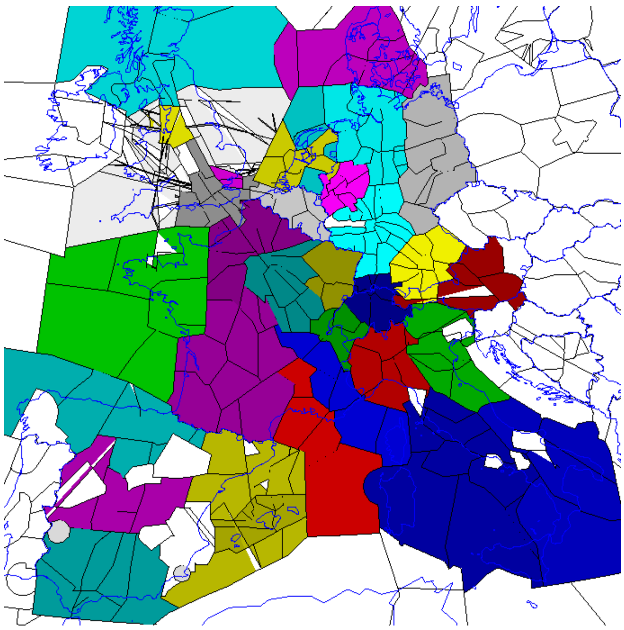


Fig. 4. Real block partitioning in February 2002 (FL240)

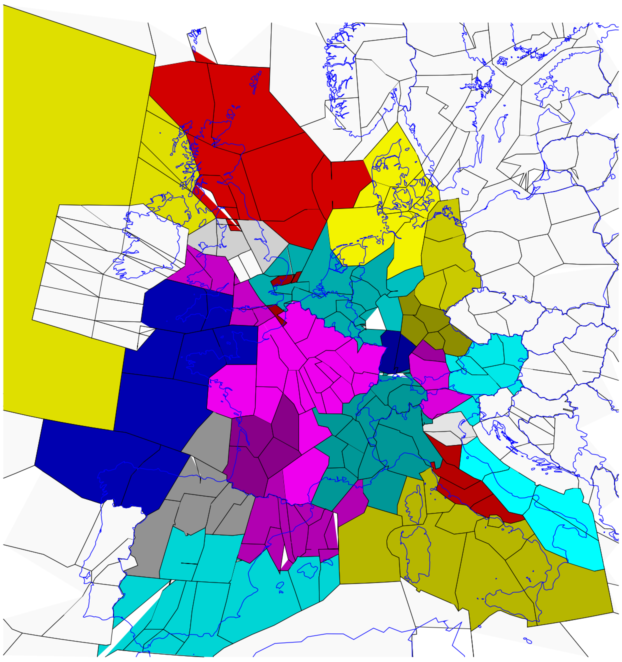


Fig. 5. Partitioning result of fusion fission (FL320)

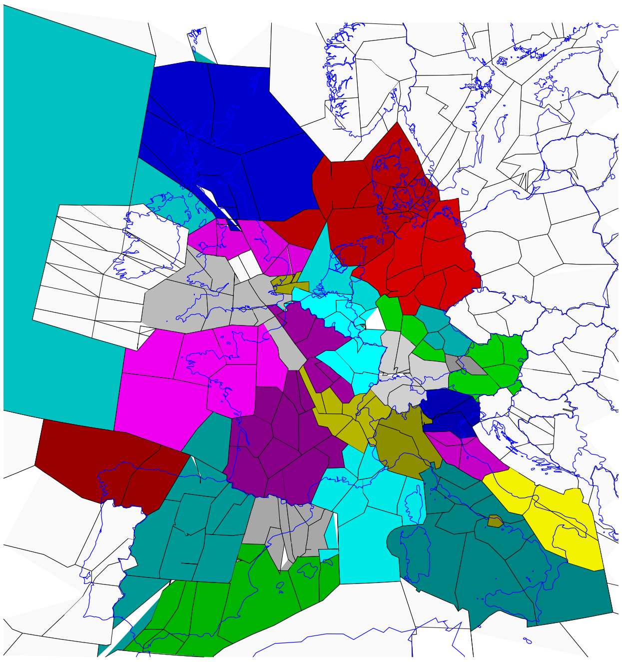


Fig. 6. Partitioning result of the CHACO package (FL320)

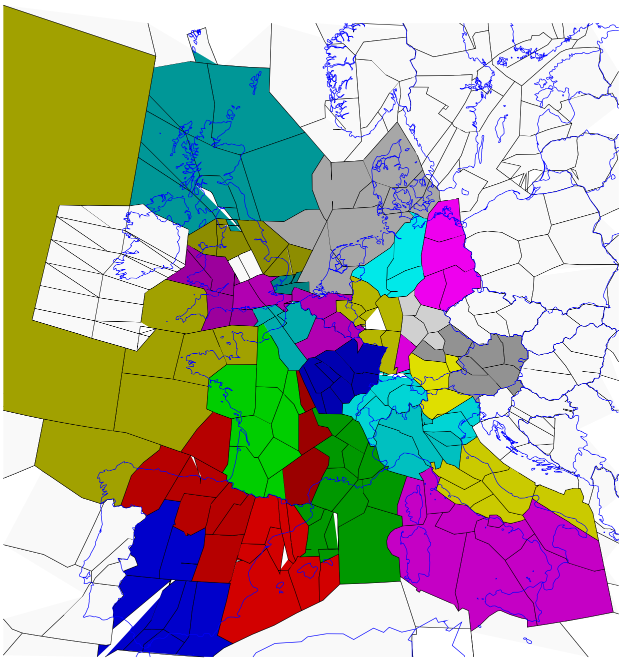


Fig. 7. Partitioning result of the METIS package (FL320)

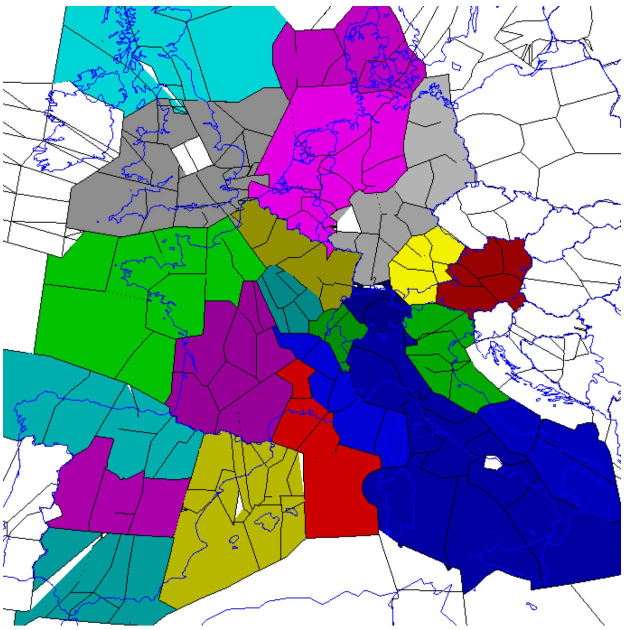


Fig. 8. Real block partitioning in February 2002 (FL320)

Session 4

Analytical Modeling 2

Evaluation of air traffic complexity metrics using neural networks and sector status

David Gianazza
Laboratoire d'Optimisation Globale

Kévin Guittet
Laboratoire d'Economie et d'Econométrie de l'Aérien

DSNA/DTI/SDER (former CENA)
7, avenue Edouard Belin 31055 Toulouse Cedex
Email : {lastname}@recherche.enac.fr

***Abstract*— This paper presents an original method to evaluate air traffic complexity metrics. Several complexity indicators, found in the litterature, were implemented and computed, using recorded radar data as input. A principal component analysis (PCA) provides some results on the correlations between these indicators. Neural networks are then used to find a relationship between complexity indicators and the actual sector configurations. Assuming that the decisions to group or split sectors are somewhat related to the controllers workload, this method allows to identify which types of complexity indicators are significantly related to the actual workload.**

I. Introduction

Much research has been conducted over the last decade to help understanding air traffic complexity and controller workload. The inadequacy of the *aircraft count* to appropriately reflect the traffic complexity has now been acknowledged for a long-time, and complementary indicators such as "traffic mix", "number of potential conflicts" and others, have been (and still are) designed. A linear combination of these variables, often referred to as *dynamic density*, is likely to better fit traffic complexity than individual indicators. It is used throughout most studies, where the correllation of a set of indicators with a quantifiable variable, assumed to represent the actual traffic complexity, is maximized.

A possible shortcoming of this methodology is that potentially non-linear relations between indicators are missed (see [1] and the concern of Eurocontrol when writing calls for proposals). But, more importantly, the choice of the dependent variable is crucial to determine how well complexity is actually measured. Indeed, physical activity, as used in [2] and [3], miss the important cognitive part of the controller activity. On the other hand, physiological indicators ([4], [5]) seem difficult to exploit and how well they relate to traffic complexity is unclear. Finally, widely used subjective ratings ([6], [7]) provide high quality data (as they obviously relate to the kind of complexity investigated), but are often seen as subject to biases (such as the recency effect denounced in [5], and the possibility of raters errors in the case of "over-the shoulder workload ratings" [8]). In all of these cases, data are very expensive to collect, as they require the active participation of controllers. Databases are often small and might exhibit low variability, which may in turn harm the statistical relevance of the results. This phenomenom is acknowledged in [7], where the overfitting of data is clearly a consequence of a lack of observations rather than a misspecification of the neural network. Finally, as these complexity metrics may be used to design computer-assisted control tools or traffic management tools, and to help organizing airspace, it is surprising to notice that the question of the relevance of the complexity measured to the final goal is scarcely discussed. The question really is to understand which complexity is measured and how well it relates to the foreseen application (benchmarking, improvements in airspace organization, design of new tools...).

This paper is motivated by former studies on optimal airspace sector configurations ([9], [10]) and intends to improve the criterion used therein to evaluate sector configurations. The basic idea, introduced in [11], is that the decisions to split a sector, mostly taken when the controller is close to overload, are linked to traffic complexity and may therefore provide an acceptable dependent variable. Interestingly enough, collecting data on sector configurations does not require controllers active participation, as current outcomes from control centers can be used, while related flight informations are available from recorded radar tracks. As such, raw data needed in our study are noticeably cheap to collect and might be produced in large quantities. The price to pay is that these data are noisy, as we may not be sure that a sector splitting (resp. merging) decision is directly related to overload (resp. underload). Other factors might distort data, such as training of unexperienced controllers, meteorological hazards, military airspace use... However, we will assume that the impact of these phenomena on the accuracy of the results is limited, particularly because of the kind of complexity we are looking at here. Indeed, this work is conducted in the perspective of future pre-tactical applications (e.g. sector planning) and thus does not ask for as much details as studies of instantaneous workload would (on the opposite, benchmarking of ATC centers would require an even coarser granularity, as indicators are averaged on wide temporal and geographical horizons [12], [13]). To investigate the link between complexity indicators and sector configurations, we use neural networks, as non-linear interactions are suspected.

The paper is organized as follows. Section II briefly describes the indicators used throughout the study, while section III presents the raw data from which the final database is built. A Principal Component Analysis (PCA) is then performed in section IV to restrain the dimensionality of the data. Neural networks are introduced in section V and their results are presented and discussed in section VI. Section VII concludes.

II. Air traffic complexity indicators

The accuracy of the results of a study related to air traffic complexity is strongly dependent of the diversity and quality of the chosen individual complexity indicators. Many have been suggested to help describe the controllers workload, and it is hardly possible to implement the entire pool. In order to limit the number of variables to be (re)programmed and present indicators that are representative of the *dynamic density* litterature, we focused on the ones selected by Kopardekar [6] in its unified complexity metric[1]. These indicators, such as references of studies where they were used and where definitions may be found, are presented in Table I. We also implemented several indicators inspired by studies conducted elsewhere in the SDER (former CENA). Definitions are indicated in appendix[2]. Finally, we also used incoming flows as explanatory variables, as they may be a significant factor in the decision to split (or merge) a sector.

III. Input data

The indicators are computed every round minute of the day, using recorded radar data, environnement data (sector description), and recorded sector configurations of the five french ATC centers. The sector configurations are recorded every round minute of the day, which explains our choice concerning the frequency at which we compute the indicators.

Indicator	Definition	Used in
Nb	Number of aircraft	[16] [7] [17] [6]
Nb^2	Squared number of aircraft	[17] [6]
σ^2_{gs}	Variance of ground speed	[7] [6]
N_{ds}	Number of descending aircraft	[2] [16] [7] [6]
N_{cl}	Number of climbing aircraft	[6] [2] [16] [7]
$\frac{\sigma^2_{gs}}{gs}$	Ratio of standard deviation of speed to average speed	[7] [6]
F_5	Incoming flow (hozizon 5mn)	[11]
F_{15}	Incoming flow (hozizon 15mn)	[11]
F_{30}	Incoming flow (hozizon 30mn)	[11]
F_{60}	Incoming flow (hozizon 60mn)	[11]
vprox_1	Vertical proximity	[7] [17] [6]
vprox_2	*See appendix*	[7] [17] [6]
hprox_1	Horizontal proximity	[7] [17] [6]
Dens	*See appendix*	[18]
track_disorder		[18]
speed_disorder		[18]
Div		[18]
Conv		[18]
sensi_d		[18]
insen_d		[18]
sensi_c		[18]
insen_c		[18]
inter_vert	*See appendix*	[13]
avg_vs		[13]
inter_hori		[13]
creed_ok	*See appendix*	[19], [20]
creed_pb		[19], [20]

TABLE I

Chosen subset of air traffic complexity indicators

Radar data is available in several forms: records made by each center, with one position every twelve seconds, in average, and a global record of the five centers, with one position every three minutes. Several months of global records were available, whereas the centers local records were not readily available, at least for a sufficiently long period of time. So we used the global records (made by the IMAGE system), and interpolated the aircraft positions in order to get one position per minute. As many trajectory changes may occur within three minutes of flight, the computed positions are not highly accurate, and this may introduce a bias in the indicators values. However, this bias is most probably of small importance in our problem: we just want to predict when a sector will be merged into another one, or split in several smaller sectors. We are not considering the instant workload, which may require a very high level of accuracy on the aircraft position, speed, and so on. To be sure that this bias is small, we should compare the computed positions, and maybe also the indicators values, using local centers records, and global records, on small data samples. This is left for future work.

Several months of recorded traffic are available. However, considering the volume of data, it would be tedious to run several experimentations on very large

[1]Though we were not always able to find an explicit formula, and thus missed seamingly important indicators like, e.g., "MET_airspace structure". Note that this difficulty to get clear definitions is also reported by Eurocontrol in [14].

[2]Further informations and discussions about indicators are to be found in the internal note [15].

data samples. So, we have restricted our choice, at least for the moment, to one day of traffic (1st june, 2003). Once we have found the most significative complexity indicators, it will be possible to re-train the neural network on larger data samples.

On the chosen day, 103 different sectors were armed. The term "sector" means here either an elementary sector, or a set of elementary sectors merged together, and handled on a single controller's working position. The air traffic complexity indicators were computed for each of these sectors, every minute of the day, together with the sector status (*merged*, *armed*, or *split*). This data was split into two sets : about sixty percent was randomly selected in order to *train* the neural network, and the rest was used to *test* the trained network on fresh data.

This single day of traffic already provides a big volume of data, as detailed in table II, with a great diversity of geographic sectors, and with enough data in each class of sector status.

	Total	Merged	Armed	Split
Train	71270	46.6%	27.0%	26.4%
Test	47513	46.4%	27.0%	26.6%

TABLE II

NUMBER OF MEASURES AVAILABLE, ON THE 1ST JUNE OF 2003.

Before applying the neural network to complexity indicators and sector statuses, let us first discuss the correlations between the indicators, using a *principal component analysis method*.

IV. PRINCIPAL COMPONENT ANALYSIS

Including incoming flows, we end up with 27 complexity indicators. Given the neural network greediness in numbers of parameters and the high multicollinearity of the data, we will use principal components rather than individual indicators for our experiments. 6 main components were identified (corresponding to eigenvalues greater than 1), that covered more than 76 % of the variance of the data set. These components are interpreted below.

C1: Eigenvalue 12.6, 46.7 % of the variance of the data set. Appart from *vprox_1* and *vprox_2*, all variables are strongly (and positively) correlated with this component, explaining its high associated eigenvalue. This component may be seen as a "size factor", and we follow [6] on the term "Overall monitoring". This component is strongly representative of the *aircraft count*.

C2: Eigenvalue 2.78, 10 % of the variance of the data set. This component is strongly correlated with avg_vs, σ_{gs}^2 and $\sigma_{gs}/\overline{gs}$ (resp. 0.70, 0.69 and 0.68). Accounting for the impact on ground speed of vertical evolution of aircraft, this component may be seen as related to the *ground speed variance*, and the *aircraft vertical evolutions*.

C3: Eigenvalue 1.96, 7.3 % of the variance of the data set. This component is mainly correlated with *incoming flows*.

C4: Eigenvalue 1.25, 4.6 % of the variance of the data set. Appart from variables directly related to the traffic volume (N, *inter_hori*), the correlation on this component is high with *insen_c* and CREED indicators, and thus might be related to *converging flows* and *anticipation of conflicts*.

C5: Eigenvalue 1.06, 3.9 % of the variance of the data set. This component is strongly correlated with Div and *insen_p*, and seems therefore mostly linked with *divergent flows*.

C6: Eigenvalue 1.03, 3.8 % of the variance of the data set. This component is strongly correlated with the horizontal proximity measures ([7]), and could stand for the *monitoring of horizontal separation* (near the minimas).

Notice that we extracted only 6 components, thus significatively less than the 12 components (briefly) described in [6]. This might be explained by the lack, at the point of the project, of indicators related to the sector geometry.

V. NEURAL NETWORKS

A. General presentation

Artificial neural networks are algorithms inspired from the biological neurons and synaptic links. An artificial neural network is a graph, with vertices (neurons, or units) and edges (connections) between vertices. There are many types of such networks, associated to a wide range of applications: pattern recognition (see [21] and [22]), control theory,...

Beyond the similarities with the biological model, an artificial neural network may be viewed as a statistical processor, making probabilistic assumptions about data ([23]). Some *train* data is used to determine a statistical model of the process which produced this data. Once correctly trained, the neural network uses this model to make predictions on new data.

Neural networks are closely related to the Bayesian probabilities. They may be used for unsupervised learning (*density estimation* problems), and, mainly, supervised learning problems (*regression*, *classification*). Density estimation is not in the scope of our paper, so we will not detail it. The aim of *regression* is to find a statistical model producing an *output* y from *input variables* (let us denote them by x), so that the output y is as close as possible to a *target variable*, which we shall denote by t. In the case of *classification* problems, the target variables represent class labels, and the aim is to assign each input vector x to a class.

We will use a specific class of neural networks, referred to as feed-forward networks, or multi-layer perceptrons (when the activation function is logistic). In such networks, the units (neurons) are arranged in fully-connected layers: an *input layer*, one or several *hidden layers*, and an *output layer*. Figure 1 shows an example of such a network.

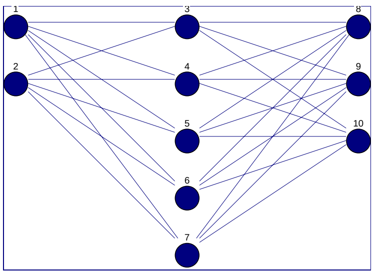


Fig. 1. Example of a feed-forward network with one hidden layer

For a network with one hidden layer, the output vector $y = (y_1, ..., y_k, ..., y_q)^T$ is expressed as a function of the input vector $x = (x_1, ..., x_i, ..., x_p)^T$ as follows:

$$y_k = \Psi(\sum_{j=1}^{q} w_{jk}\Phi(\sum_{i=1}^{p} w_{ij}x_i + w_{0j}) + w_{0k}) \quad (1)$$

where the w_{ij} and w_{jk} are weights assigned to the connections between the input layer and the hidden layer, and between the hidden layer and the output layer, respectively, and where w_0j and w_0k are biases (or threshold values in the activation of a unit). Φ is an *activation function*, applied to the weighted output of the preceding layer (in that case, the input layer), and Ψ is a function applied, by each output unit, to the weighted sum of the activations of the hidden layer. This expression can be generalized to networks with several hidden layers.

The output error – *i.e.* the difference between the target values t and the output y computed by the network – will depend on the parameters w (weights and biases). The *training* aims at choosing these parameters, so as to minimize a chosen function of the output error.

In the case of *regression*, the minimized function is the sum of quadratic errors. For classification problems, it is best to consider a log-likelihood function. The network is then designed with one output unit per class. When the output of such an unit is 1, the input x is assigned to the class corresponding to the unit, and when the output is 0, it is not. Let us consider a problem with C classes. The log-likelihood function minimized during training is the following, known as *cross entropy*:

$$E(w) = -\sum_{n=1}^{N}\sum_{k=1}^{C} t_k^{(n)} ln(y_k^{(n)}) \quad (2)$$

where $t^{(n)}$ and $y^{(n)}$ are the n^{th} target and output vectors, respectively.

Several optimization methods may be used to minimize $E(w)$, when training the network. Let us cite *backpropagation*, which consists in successive modifications of the weights assigned to the connections between the layers, starting with the output layer. These modifications take account of the relative importance of each weight in the output error variations. Other local optimizations using the gradient of the error (BFGS, conjugate gradients, for example) are also widely used. A variety of global optimization methods (simulated annealing, evolutionary algorithms) are also proposed in the literature. These global methods usually perform better than the local methods when there are many local minima for the error function, but they are generally much slower.

B. Neural networks applied to our problem

For our problem, we have chosen three-layers feed-forward networks, denoted $I_\alpha H_\beta O_\gamma$ in the rest of the paper, with α units in the input layer, β units in the hidden layer, and γ units in the output layer. The input variables are normalized, by substracting the mean value and dividing by the standard deviation.

There are many possible choices for the functions Ψ and Φ, depending on the problem being adressed. A common choice for Φ is the logistic function :

$$\Phi(z) = \frac{1}{1 + e^{-z}} \quad (3)$$

This is the activation function that was used in our experiments. As we adress a classification problem – assign each input vector (a list of complexity indicators values) to a class representing the sector status (*merged*, *armed*, or *split*) – we have chosen to minimize the *cross entropy* function. Therefore, the transfer function Ψ applied to the output layer, must be the *softmax* function:

$$\Psi(z_k) = \frac{e^{z_k}}{\sum_{m=1}^{C} e^{z_m}} \quad (4)$$

The `nnet` package of the `R` language was used (see `http://www.r-project.org/` for details on the `R` language and environnment). In this package, developped by Pr B. D. Ripley, a quasi-newton minimization method (BFGS) is used for the network's learning. The parameters of the `nnet` tool are the range parameter (default 0.5), defining the range into which the initial random weights are chosen, the parameter for weight decay (default 0), and the maximum number of iterations. The training stops either if the fit criterion (the *cross entropy* in our case) falls below a chosen parameter $abstol$ (default 1.0e-4), or if the improvement of the fit criterion is less than $1 - reltol$ (the default value for $reltol$ is 1.0e-8).

Several combinations of air traffic complexity indicators, or of principal components, will be tested. The number of input units of the network will be chosen equal to the cardinal of the evaluated set of indicators (or components). The number of hidden units is 15 (this choice is discussed later). The output layer is made of three units, one for each class (*merged*, *armed*, or *split*).

So a target vector $t^{(n)}$ with value $(1, 0, 0)$ means that the considered sector was merged with other sectors when the n^{th} measure of the vector of complexity indicators was made. Armed sectors will be represented by $(0, 1, 0)$. A value of $(0, 0, 1)$ will mean that the sector was previously split in two or more sectors at the time x was measured. Of course, the actual output of the neural network will not be exact values 0 or 1. It will be triples (a, b, c) of floating-point values between 0 and 1, each value being the probability to belong to a class. The input vector $x^{(n)}$ will be assigned to the class of highest probability.

C. Evaluation of the neural network's outputs

A well-known problem, when using neural networks (or other regression methods), is *overfitting*: with enough parameters and enough training cycles, it is always possible to find a good fit for a given data set. So one may find a perfect fit for a chosen data sample, and then feel disappointed when the trained network makes wrong predictions on fresh data. So, we will systematically proceed as follows: train the network on a randomly chosen data sample (called *train*), then check the results, first on the same data sample, and second on a fresh data sample (called *test*), that was not used for the training.

In order to evaluate the outputs of several different models, we have to compare the neural networks predictions to the actual target values. We may use the fit criterion (*cross entropy*) but it does not reflect the influence of the number of weights (and biases) in the neural network. It is known (see [23]) that a network with too few weights may not be able to capture all the variations of the response to the input x, whereas a network with too many weights will more likely be subject to *overfitting*. In the next sections, we will compare several sets of input variables, of various sizes. Consequently, *the number of weights in the network will not remain constant*, and this variation will bias the results.

We will therefore use the *Akaike information criterion* ([24]): $AIC = 2\lambda - 2ln(L)$, where λ is the number of unadjusted parameters of the model (i.e. the number of weights and biases of the network), and $ln(L)$ is the log-likelihood. This criterion is strongly related to information theory, and more specifically to the Kullback-Leibler distance (K-L) between a candidate model and the "true" model. In our case, the AIC is written as follows:

$$AIC = 2\lambda - 2\sum_{n=1}^{N}\sum_{k=1}^{C} t_k^{(n)} ln(y_k^{(n)}) \qquad (5)$$

One should be aware that AIC is a *relative* criterion, which can only be used to compare a set of candidate models relative to a same "true" model: as this true model is unknown, the corresponding term in the K-L distance was considered as a constant and dropped, in the AIC. As we would like to compare predictions made on the *train* test and on the *test* set, which are of different size, we will divide the AIC by N, the number of data items, and use: $AIC_{avg} = \frac{AIC}{N}$.

In addition to the numerical results provided by the *Akaike information criterion*, we shall also consider the global proportion of correctly classified input vectors, and also the percentage of correct classifications for each class. One must be aware, however, that the rate of correct classifications *is not* the criterion being maximized by the neural network, so we should remain cautious when comparing the different classification rates. However, these percentages are easily understandable and may allow us to make some interesting statements on the results.

VI. Results

A. Preliminary results, discussion on the parameters

Before making statements about how significant the indicators (or combinations of indicators) are, let us first check if the chosen network is efficient on our problem, and consider how to choose the network and training parameters.

Figure 2 shows the evolution of the cross entropy criterion during the network's training. The input variables were the six main components found by the PCA, and also the sector's volume, so the input layer of the network had 7 units. We have chosen a hidden layer with 15 units. The output layer had 3 units, each one representing a class, as explained in section V. The maximum number of cycles was set to 1500, but the training actually stopped at 1010 cycles, because the algorithm was unable to significantly improve the cross entropy criterion.

Several questions may arise, concerning the parameters choices, and their influence on the results. In this

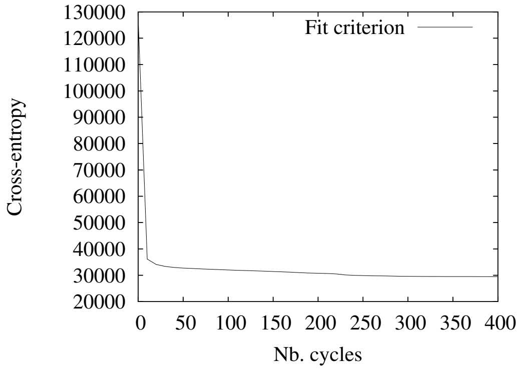


Fig. 2. Evolution of the cross entropy criterion with the number of training cycles, using the main 6 components and the volume as input (network $I_2H_{15}O_3$)

study, we will use the default values for the following `nnet` parameters: the stop criteria $abstol$ and $reltol$, and the weight decay parameter, set to $decay = 0$. The range in which the network's weights and biases are initialized depends on the range of the input variables, that we have normalized. A parametric study with all components plus the sector volume as input and with 15 units in the hidden layer shows that the best results, averaged on two runs for each value of the range parameter, were obtained with a range value of 0.4. So we shall use 0.4 for the `nnet` range parameter.

It may have been useful to make another parametric study, trying to find the optimum number of hidden units by minimizing the AIC, but this is rather time-consuming (it should be repeated for each set of input variables). So we have chosen, after a few trials, a hidden layer of 15 units.

B. Model selection

Let us now select the best model among several candidate models. A *model* is a trained neural network and set of input variables that we expect to provide a good explanation of the sector status (*merged*, *armed*, or *split*). We use the PCA components plus the sector volume as input variables. The sector volume was not analyzed in the PCA, but as we have not implemeted any indicator using the sector geometry, we will use the volume as a proxy for metrics such as "space available around conflicts" or "distance to sector boundary".

An iterative approach is used: we shall first use component C_1 (representative of the number of aircraft) as input to the neural network, then add the volume, and continue with the five other main components, successively added in the order found by the PCA. At last, we will use all the 27 components, and the sector volume, as input. The AIC_{avg} criterion (see section V-C) is used to select the best model.

For each set of input variables, five training runs are made. The reason for this is the following: the training method of `nnet` is a local optimization method which starts at a randomly chosen point (the initial weights), and which follows the steepest descent of the error function being optimized. This error function may have several local minima, so choosing different initial weight values may lead to different results. Although these local minima are often fairly close, several runs will comfort our results.

Figure 3 shows, for *train* and *test* data, how the AIC_{avg} criterion evolves when adding components to the set of input variables. The mean values, averaged on the five runs, are presented. This figure should be interpreted as follows: when the AIC significantly decreases when introducing a new variable in the model, this means that this variable improves the prediction of the sector status. When the AIC increases or remains constant, this means that the benefit provided by the additional variables is offset by the complexity it implies on the model (increase in the number of parameters)[3].

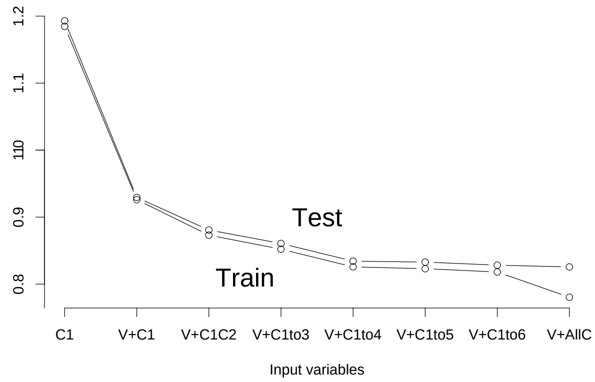


Fig. 3. Values of the AIC_{avg} criterion for the different sets of input variables

The two curves, *train* and *test*, are fairly close, except at the last point ($V + AllC$), which corresponds to the model $\{V; C_1; ...; C_{27}\}$. For this last point, the AIC criterion is improved on *train* data, whereas it remains nearly the same on *test* data, which shows a little overfitting of the neural network on *train* data. Otherwise, the results on *train* and *test* data are quite consistent, so the neural networks are able to generalize efficiently on fresh data.

The AIC criterion is significantly improved when adding the *sector volume* to the component C_1 (*aircraft count*). Component C_2, which is mostly related the

[3]We could have used the Schwartz BIC criterion instead of the AIC. As it assigns a heavier penalty to additionnal parameters, it may have been easier to interpret.

speed variance and the *aircraft vertical evolutions* also improves the criterion, although much less than the sector volume. Small improvements are brought by component C_3 (*incoming flows*), and also by C_4, mostly related to the *converging flows* and the *conflict anticipation*. The components C_5 (*diverging flows*) and C_6 (*monitoring of horizontal proximity*) bring no significant improvement. The use of the other components does not improve the prediction, as shown by the last point of the *test* curve.

It might look surprising that the aircraft horizontal proximity is of little influence on the sector status explanation. But, as was already stated in [7], the aircraft have already been separated – vertically, we may guess – before the proximity situation occurs, which explains why it is not very useful. The *anticipation* of future aircraft proximity (component C_4), is more significant.

C. Classification rates

Tables III and IV give the proportions of correct classifications made by the neural networks (the best of the five runs), using *train* data as input, or *test* data, respectively. The second column of each table shows the global rate of correct classifications. The three last columns detail the results for each class (*merged*, *armed*, or *split*). As previously stated on figure 3, the results on *test* data are quite consistent with the results on *train* data.

Set	Global	Merged	Armed	Split
$\{C_1\}$	72.91%	81.92%	44.90%	85.61%
$\{V;C_1\}$	79.34%	86.26%	59.01%	87.89%
$\{V;C_1;C_2\}$	80.70%	86.93%	60.79%	90.04%
$\{V;C_1;C_2;C_3\}$	80.84%	87.32%	60.23%	90.43%
$\{V;C_1;...;C_4\}$	82.03%	87.87%	63.02%	91.15%
$\{V;C_1;...;C_5\}$	81.80%	88.23%	61.77%	90.88%
$\{V;C_1;...;C_6\}$	81.83%	88.09%	62.16%	90.85%
$\{V;C_1;...;C_{27}\}$	83.36%	88.63%	65.61%	92.19%

TABLE III

CORRECT CLASSIFICATIONS ON TRAIN DATA

Set	Global	Merged	Armed	Split
$\{C_1\}$	72.61%	81.64%	44.41%	85.40%
$\{V;C_1\}$	79.12%	85.56%	59.14%	88.12%
$\{V;C_1;C_2\}$	80.48%	86.58%	60.74%	89.82%
$\{V;C_1;C_2;C_3\}$	80.42%	86.71%	60.04%	90.08%
$\{V;C_1;...;C_4\}$	81.82%	87.78%	62.77%	90.74%
$\{V;C_1;...;C_5\}$	81.59%	88.19%	61.44%	90.49%
$\{V;C_1;...;C_6\}$	81.65%	87.98%	61.88%	90.64%
$\{V;C_1;...;C_{27}\}$	82.67%	88.04%	64.51%	91.68%

TABLE IV

CORRECT CLASSIFICATIONS ON TEST DATA

When using only the component C_1 (*aircraft count*) as input, the neural network already makes more then 72% of correct predictions. However, the rate for the *armed* class is less than 45%, which is not very good. Once again, we see that the use of the sector volume (denoted V) greatly improves the prediction of the sector status, with more than 79% of correct classifications, globally, and a rate of about 59% for the *armed* class. The classification rates when adding C_2 climb to 81% for the global rate, and about 60% for the *armed* class. Adding C_3 does not increase the rates, although it improved the AIC. C_4 slightly improves the results, whereas C_5 and C_6 provide no improvement. At most, when considering the sector volume and the 4 first components, we have about 82% of correct classifications, and about 63% of correct classifications for the *armed* class.

The last line of both tables, with all components and the volume, shows the best classification rates, whereas the AIC showed no improvement in the same case. The apparently good results of the last line are certainly due to the higher number of network's parameters than in the previous models. They do not mean that the model with all components is the best one.

In order to verify this assumption, we have tried another network for the $\{V;C_1;...;C_4\}$ model, with about the same number of parameters than the $\{V;C_1;...;C_{27}\}$ model. This last model was assessed with a network $I_{28}H_{15}O_3$, where there were 483 weights and biases. Let us take a $I_5H_{53}O_3$ network, with 53 units in the hidden layer, for the $\{V;C_1;...;C_4\}$ model. This network has 480 parameters. Table V shows that the classification results, with this network, are better than the ones obtained with all components and the volume in table III and IV.

Set	Global	Merged	Armed	Split
$\{V;C_1;..;C_4\}_{train}$	84.34%	89.74%	67.61%	91.88%
$\{V;C_1;..;C_4\}_{test}$	83.84%	89.32%	67.26%	91.09%

TABLE V

CORRECT CLASSIFICATION RATES FOR THE $\{V;C_1;...;C_4\}$ MODEL, WITH A $I_5H_{53}O_3$ NETWORK

When considering the detailed results, for each class, on the three tables, we see that the *merged* and *split* classes have better classification rates than the *armed* class. This is not a surprise, as the cloud of points representing the measures of the *armed class*, in the variables space, is located "between" the clouds representing the two other classes. The neural network aims at finding the frontiers between these clouds, and this is more difficult with two frontiers instead of one.

To conclude this section, we have seen that the best model is able to classify correctly about 84% of the input vectors. We may still slightly improve these results by running a parametric study on the number of hidden units, for this model, still minimizing the AIC criterion. But this is rather time-consuming, and our results are already quite good, considering that our data is rather noisy. Let us remind that we have assumed that the decisions to split or to merge sectors had a single cause, that is the workload. This may not always be the case in reality, where there could be other reasons: controllers training, hardware failures, and so on. In a next step of our research, we may try to improve our results by filtering the data – at least the sector *split* decisions – using an approximation of the workload (the number of aircraft for example).

D. Discussion on the results, comparison with other works

To summarize the previous results, the AIC criterion (figure3) allowed us to select the best model among the ones we have tested. This model uses a subset of only 5 input variables $\{V; C_1; ...; C_4\}$, among the 28 that we have considered.

The proportions of correct classifications (tables III and IV) show that the use of the sector volume significantly improves the network's prediction, by 6.5%, when compared to the sole *aircraft count* C_1. There is a relatively small improvement (around 3%), between the *sector volume + aircraft count* model and the best model. The other components bring no significant improvement.

Other works, like [7] and [6] for example, have already stated that the *aircraft count* model is not very efficient in predicting the controller's workload. In these studies, the improvement brought by other complexity indicators was much higher than what is shown in our results. There may be several explanations to this.

In [7], Chatterji and Sridhar show workload prediction rates (on test data) ranging from less than 16% to 54%, for their *medium* workload class, and from 0% to 100% for their *high* workload class. However, their results with *test* data were not consistent with the ones obtained with *train* data. They honestly state that they were unable to draw reasonable conclusions from these results, as they had too few measures in the medium and high workload classes.

In [6], Kopardekar and Magyarits apply a linear regression on subjective complexity ratings. The results were not given in percentages of correct classifications. The R^2 criterion of the regression was used to compare the candidate models. The results presented in [6] show some significant differences of R^2 values between the *aircraft count* model and the model based on *dynamic density* (which is viewed as a linear combination of several indicators). Kopardekar and Magyarits also lacked measures, but for very low and high workload traffic.

Both studies ([7], [6]) use subjective complexity ratings, provided by air traffic controllers who assessed the traffic complexity of several traffic samples, only during periods when the sector was armed[4]. In our study, we use the actual sector status, recorded for several elementary sectors and groups of sectors, assuming that the sector status is related to the controllers workload. So we have only three levels of workload: *low*, when the sector is merged with other sectors, *normal use*, when the sector is armed, and *too high* when the sector has been split into smaller sectors. We have the feeling that the studies [7] and [6] focus on the *normal use* load interval, and that the use of complexity ratings within this domain magnifies the observed phenomenons.

Aside from these considerations, another, and more straightforward, explanation of the relatively small differences between our candidate models – when comparing the classification rates – resides in the nature of the data used in our study. We simply have a great number of measures for which there is no doubt as to which class they belong to, even when using the worst possible model. For example, when there are 2 aircraft in a given sector, it is most likely merged with other sectors. On the opposite, when there are 60 aircraft in a sector, it has certainly been split into several sectors. We may have exhibited much higher variations in the results, by computing the classification rates using only measures collected around the times when the sector configurations changed.

VII. Conclusion

In conclusion, we were able to select the best model, among several candidate models, establishing a functional relationship (equation V-A with the weights of the trained network) between the air traffic complexity indicators and the sector status. Assuming that the decisions to merge or to split a sector are statistically related to the controllers workload, this original method provides an objective way to validate the complexity metrics. Our method also has the advantage, in comparison to other methods, to use widely available data (sector configurations and radar tracks recorded by the ATC centers), noticeably cheap to collect, as the active participation of air traffic controllers is not required.

Neural networks, minimizing the *cross entropy* function of the output error, showed good results, consistent on *train* and *test* data. The *Akaike information criterion* proved useful in selecting the best model, avoiding the bias due to the different number of parameters in the candidate models. In the iterative approach that

[4]This is not explicitly stated in the papers, but we assume so, as they lacked measures for high workload traffic situation.

was used, the highest improvement, when comparing to the *aircraft count* model, was brought when introducing the *sector volume* as a new input variable. Smaller improvements were provided by components C_2 (*speed variance* and *aircraft vertical evolutions*), C_3 (*incoming flows*), and C_4 (*converging flows* and *conflict anticipation*).

So far, we have only considered the PCA components in our study. The next step of our research may be to select a subset of individual indicators, issued from the components of the best model, and re-iterate the approach presented in this paper. This would provide a more direct and simple relationship between the indicators and the sector status, by avoiding to compute the components from all the indicators.

The neural network approach used in this study seems appropriate for the granularity we are interested in and the foreseen applications, either strategical (sector design) or pre-tactical (sector planning). We are also fairly confident that decisions to split or merge sectors may allow to assess the instantaneous workload as well, and could therefore be used to improve tactical tools (PRESAGE). To this end, other statistical methods should be investigated to take into account the serial correlation of sector status, looking closely at the sector splitting times. We plan to tackle this issue in a close future, using dynamic discrete choice models.

Finally, another issue that we intend to address, in relation to the complexity indicators, is the prediction of optimal sector configurations. Previous works ([9], [10]) proposed several algorithms to compute optimal sector configurations, using sector capacities and *incoming flows*. The output of the neural network is a triple of probabilities, allowing to decide when a sector should be split, or merged. We may derive a realistic workload indicator – and also threshold values – from these probabilities, which could be used to compute optimal sector configurations.

REFERENCES

[1] B. Hilburn and G. Flynn. Toward a non-linear approach to modeling air traffic complexity. In *2nd Human Performance Situation Awareness and Automation Conference*, 2004.

[2] I. V. Laudeman, S. G. Shelden, R. Branstrom, and C. L. Brasil. Dynamic density: An air traffic management metric. Technical report, 1999.

[3] A. Majumdar, W. Y. Ochieng, G. McAuley, J.M. Lenzi, and C. Lepadetu. The factors affecting airspace capacity in europe: A framework methodology based on cross sectional time-series analysis using simulated controller workload data. In *Proceedings of the 6th USA/Europe Air Traffic Management R & D Seminar*, 2005.

[4] J.H. Crump. Review of stress in air traffic control: Its measurement and effects. *Aviation, Space and Environmental Medecice*, 1979.

[5] P. Averty, S. Athènes, C. Collet, and A. Dittmar. Evaluating a new index of mental workload in real atc situation using psychological measures. Note cena nr02-763, CENA, 2002.

[6] P. Kopardekar and S. Magyarits. Measurement and prediction of dynamic density. In *Proceedings of the 5th USA/Europe Air Traffic Management R & D Seminar*, 2003.

[7] G.B. Chatterji and B. Sridhar. Measures for air traffic controller workload prediction. In *Proceedings of the First AIAA Aircraft Technology, Integration, and Operations Forum*, 2001.

[8] C. Mannings, S. Mill, C. Fox, E. Pfleiderer, and H. Mogilka. The relationship between air traffic control events and measures of controller taskload and workload. In *Proceedings of the 4th Air Traffic Management Research & Developpment Seminar*, 2001.

[9] D. Gianazza and J. M. Alliot. Optimization of air traffic control sector configurations using tree search methods and genetic algorithms. In *Proceedings of the 21st Digital Avionics Systems Conference*, 2002.

[10] D. Gianazza, J. M. Alliot, and G. Granger. Optimal combinations of air traffic control sectors using classical and stochastic methods. In *Proceedings of the 2002 International Conference on Artificial Intelligence*, 2002.

[11] D. Gianazza. *Optimisation des flux de trafic aérien*. PhD thesis, Institut National Polytechnique de Toulouse, 2004.

[12] G. M. Flynn, C. Leleu, and L. Zerrouki. Traffic complexity indicators and sector typology analysis of u.s. and european centres. Technical report, Eurocontrol, 2003.

[13] Note de synthèse sur l'indicateur de complexité pru. Technical report, DTI/SDER (ex CENA), 2005.

[14] Cognitive complexity in air traffic control, a litterature review. Technical report, Eurocontrol experimental centre, 2004.

[15] K. Guittet and D. Gianazza. Analyse descriptive des indicateurs de complexité du trafic aérien à partir des données image et courage. Note nr05-905, DSNA/DTI/SDER, Décembre 2005.

[16] B. Sridhar, K. S. Sheth, and S. Grabbe. Airspace complexity and its application in air traffic management. In *Proceedings of the 2nd USA/Europe Air trafic Management R&D Seminar*.

[17] P. Kopardekar. Dynamic density: A review of proposed variables. Faa wjhtc internal document. overall conclusions and recommendations, Federal Aviation Administration, 2000.

[18] D. Delahaye and S. Puechmorel. Air traffic complexity: towards intrinsic metrics. In *Proceedings of the third USA/Europe Air Traffic Management R & D Seminar*, 2000.

[19] P. Averty. Conflit perception by atcs admits doubt but not inconsistency. In *Proceedings of the 6th Air Traffic Management Research & Developpment Seminar*, 2005.

[20] P. Averty, K. Guittet, and P. Lezaud. *Work in progress, presented at an internal SDER seminar*. Technical report, DTI/SDER (former CENA), 2005.

[21] C. M. Bishop. *Neural networks for pattern recognition*. Oxford University Press, 1996. ISBN: 0-198-53864-2.

[22] B. D. Ripley. *Pattern recognition and neural networks*. Cambridge University Press, 1996. ISBN: 0-521-46086-7.

[23] M. I. Jordan and C. Bishop. *Neural Networks*. CRC Press, 1997.

[24] W. H. Greene. *Econometrics analysis*. Prentice Hall International. ISBN: 0-13-015679-5.

[25] F. Chatton. Etudes de nouvelles métriques de complexité de la circulation aérienne. Master's thesis, Ecole Nationale de l'Aviation Civile (ENAC), 2001.

APPENDIX: COMPLEXITY METRICS

Delahaye and Puechmorel metrics

To present the geometrical indicators introduced in [18], we need to define several quantities:

- The vector representing the distance between two aircraft is denoted by $\overrightarrow{X_iX_j}$ where X_i (resp. X_j) stands for the location of aircraft i (resp. j).
- The "oblical" distance between two aircraft (i and j) is denoted by

$$d_{ij}^{ob} = \sqrt{< \overrightarrow{X_iX_j}, \overrightarrow{X_iX_j} >}\ , \tag{6}$$

where $< .,. >$ stands for the appropriate scalar product.

- We denote by $\vec{v}_{ij} = \vec{v}_j - \vec{v}_i$ the speed difference between two aircraft.
- The derivative of the "oblical" distance between two aircraft is denoted by v_{ij} and writes

$$v_{ij} = \frac{< \overrightarrow{X_iX_j}, \vec{v}_{ij} >}{d_{ij}^{ob}} . \quad (7)$$

- We introduce a weighting function f. As suggested in [25], we used

$$f(d_{ij}^{ob}) = \frac{e^{-\alpha (d_{ij}^{ob})^2} + e^{-\beta d_{ij}^{ob}}}{2} , \quad (8)$$

with $\alpha = 0.002$, $\beta = 0.01$, the d^{ob} being expressed in nautical miles.

These indicators are defined pointwise. To get a value on the controlled airspace, they have to be averaged on the different aircraft. In [18], a density indicator is defined as follows

$$Dens(i) = \sum_{j=1}^{N} b f(d_{ij}^{ob}) . \quad (9)$$

Two indicators are introduced to reflect the variability in headings (*track_disorder*) and speed (*speed_disorder*). There are defined as

$$track\_disorder(i) = \sum_{j \neq i} |\theta_i - \theta_j| f(d_{ij}^{ob}) . \quad (10)$$

$$speed\_disorder(i) = \sum_{j \neq i} \|\vec{v}_{ij}\| f(d_{ij}^{ob}) . \quad (11)$$

Indicators *Div* et *Conv* respectively describe convergency and divergency of the aircraft in the controlled sector.

$$Div(i) = \sum_{\substack{j=1 \\ j \neq i}}^{Nb} 1_{\mathbb{R}^-}(v_{ij}) \, . \, |v_{ij}| \, f(d_{ij}^{ob}) , \quad (12)$$

$$Conv(i) = \sum_{\substack{j=1 \\ j \neq i}}^{Nb} 1_{\mathbb{R}^+}(v_{ij}) \, . \, |v_{ij}| \, f(d_{ij}^{ob}) . \quad (13)$$

Indicators Sd_+ and Sd_- are designed to set a weight on potential conflicts that are difficult to solve. These "sensitivity" indicators are defined by

$$Sd_-(i) = \sum_{\substack{j=1 \\ j \neq i}}^{Nb} 1_{\mathbb{R}^-}(v_{ij}) \, \|\vec{\nabla} v_{ij}\| f(d_{ij}) , \quad (14)$$

$$Sd_+(i) = \sum_{\substack{j=1 \\ j \neq i}}^{Nb} 1_{\mathbb{R}^+}(v_{ij}) \, \|\vec{\nabla} v_{ij}\| f(d_{ij}) . \quad (15)$$

Note that components of the gradient are weighted so as to reflect the difficulty of the respective manoeuvers[5]. As observed in [18], a situation with a high "sensitivity" is easier to resolve for the air controller than one with a low "sensitivity". As these indicators "increase" with the number of aircraft, it is unclear whether they actually are "complexity" or "simplicity" indicators. We thus define a last pair of indicators, *insen_c* and *insen_d* as

$$insen\_c = \frac{Conv^2}{Sd_+} \quad \text{and} \quad insen\_d = \frac{Div^2}{Sd_-} . \quad (16)$$

Modified PRU metrics

The work conducted by SDER-RFM for the Performance Research Unit (citeRFM), though initially designed to compare ATC centers on a daily basis, inspired the following indicators :

- *inter_hori*: number of potential crossings (irrespective of the aircraft direction on their trajectories) with angle greater than 20 degrees.
- *inter_vert*: denote by n_1, n_2 et n_3 the numbers of stable/climbing/descending aircraft. The indicator is then defined as

$$\texttt{inter\_vert} = \frac{(n_1 n_2 + n_2 n_3 + n_1 n_3)}{(n_1 + n_2 + n_3)} . \quad (17)$$

- *avg_vs*: this is simply the average vertical speed of controlled aircraft.

Metrics inspired from the CREED project

The work of P. Averty on conflict detection [19] inspired a set of indicators. One of the ideas in [19] is that conflict perception is "plannar". The author thus defines for converging pairs of aircraft the following quantities

- *Ed* : minimum horizontal distance between aircraft.
- *Efl* : horizontal distance when the aircraft are vertically separated (after the crossing).
- *Da* : the "anticipation degree", i.e. the distance between the faster aircraft and the intersection of the aircraft trajectories (in the horizontal plan). We replace this variable to a modified *Da*, *DaC*, which stands for the greater distance between one of the aircraft and the point where, horizontally, the distance between aircraft is the smallest. For explanations about this substitution, we refer to [15].

Originally, these quantities are defined to describe conflict perception. To translate the idea of [19] in terms of traffic complexity, we assume that a conflict is all the more critic that the expected separation (*Ed* and *Efl*) and the anticipation (*DaC*) are small. We thus set

$$\texttt{creed} = \frac{1}{\alpha \ \texttt{Da} + (1-\alpha)(\beta \ \texttt{Ed} + (1-\beta) \ \texttt{Efl})} , \quad (18)$$

where α and β are parameters in [0; 1][6]. Finally, aircraft pairs considered in [19] are such that vertical separation occurs prior to separation, as the converse situation is avoided as much as possible by controllers. Accordingly, the complexity associated with these latter pairs is likely to be greater and we distinguished the two kind of conflicts by summing the quantity fintroduced in (18) on both sets of aircraft, thus creating two distinct indicators, *creed_ok* ("good pairs") and *creed_pb* ("bad pairs").

[5] Reasonable weights were given by P. Averty and M. Tognoni.

[6] As for now, these parameters are set equal to 0.5, but are meant to be adjusted and possibly vary with *DaC* to reflect the results of ongoing research [20].

Averages, Uncertainties and Interpretation in Flow Planning

Claus Gwiggner
LIX, Ecole Polytechnique
Palaiseau, France
Email: gwiggner@lix.polytechnique.fr

Vu Duong
EUROCONTROL Experimental Center
Bretigny-sur-Orge, France
Email: vu.duong@eurocontrol.int

***Abstract*— In an analysis of flight data, we found a relationship between the planned number of sector entries and the average realized number. It suggests that there are systematically more aircraft than planned arriving for few planned traffic and less for high numbers of planned traffic. This is counter-intuitive since one would expect random fluctuation around the planned number if the planning procedure were accurate. The relationship we found can be described concisely by logarithmic, square root or reciprocal functions. Moreover, we show that it can be seen as the mean values of Poisson distributions.**
Taking both together, the uncertainty about the real number of aircraft arriving at a sector can be characterized.
We validate the findings on a large number of sectors, randomly chosen from the Central European Upper Airspace.
The results are empirical but they give insight into how controllers deal with their workload.

I. Introduction

In European airspace, the main strategy to balance demand with the available capacity is to distribute departure slots among aircraft. As today, this idea assumes that trajectory and speed of all aircraft are known in advance. Experience has shown that another number of aircraft than planned (the planned traffic) sometimes arrives at sectors (the realized traffic) (figure 1). This may cause safety problems on the one hand and non-optimally used sector capacity on the other hand. A main reason for this phenomenon is *uncertainty* about the behavior of users of the airspace: passenger delays, controller behavior or others. Until now, it is unknown how all these uncertainties play together [1]. Are there propagations of delays that lead to congestion? Or do pilots and air traffic controllers compensate them successfully?
The aim of our study is to find out whether there are, or not, situations, in which systematically another number of aircraft than planned arrives at sectors. One would expect that the realized traffic equals in average the planned traffic, if the flow planning were accurate.
There are several possible approaches to this problem. One could *build* models taking into account the 'stochastic' behavior of airspace users. An important question would then be on which assumptions such models are built.
Our approach is to analyze past flight data recordings in order to *describe* this behavior. In more detail, we analyze the relationship between planned and realized traffic. An important question in this approach is how to interpret and generalize possible findings.

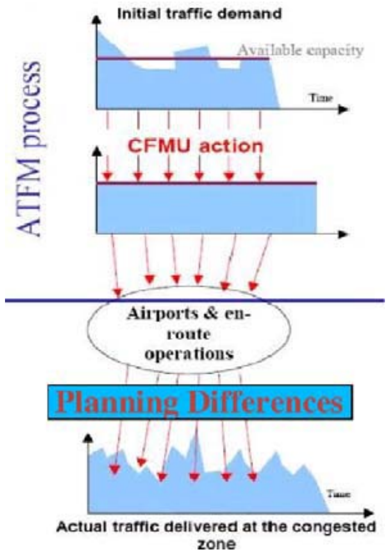


Fig. 1. Current ATFM Procedure

This paper is organized as follows: after reviewing related work we present a relationship that we discovered during data analysis and its main properties. In section IV we validate our findings. We interpret the results in section V before concluding with a short summary of the article.

II. Related Work

Major uncertainty factors (e.g. lost slots, reactionary delay, etc.) are identified in [1]. A statistical analysis of departure delays can be found in [2]. Simulation studies conclude that differences between planned and realized traffic can appear under 'normal' conditions [3].
What is unknown today is how the interaction of uncertainty factors affect the *real* situation at en-route sectors [1].
Formal statistical approaches to analyze data with the above characteristics can be divided into two parts: count data analysis (e.g. [4]) and Point process analysis (e.g. [5], [6], [7]). The former studies relationships between discrete variables and the latter between series of events.

III. Analysis

Notation
$REAL_t^S$: number of real entries in sector S in time interval t.
PLN_t^S: number of planned entries in sector S in interval t.
$t \in \{1, ..., t_{max}\}$: time intervals over a day (please see below for the choice of t_{max}).
When it is clear from the context, we omit the indices. In general, we consider $REAL_t^S$ as random variables with unknown distributions and the data as realizations thereof.

a) Relationship between planned and realized traffic: A typical daily pattern of sector entries is shown in figure 2: few traffic in the morning and night; peak hours around noon and in the late afternoon. We superpose a fourth order polynomial time trend (bold line). Note that there are many possibilities to model a time trend, for example with harmonic curves.
We are more interested in how the real number of arrivals in a time interval t depends on the planned number of arrivals ($REAL_t^S$ vs PLN_t^S). In other words, we are looking for a trend that is not a function of time but a function of planned traffic.
In order to gain insight into this relationship, please look at figure 3. It plots the planned number of entries against the *average* number of real entries (bold line). The diagonal corresponds to the cases where exactly the same number than planned arrives. The averages lie above the diagonal for small values of PLN and below the diagonal for large values. This suggests that controllers avoid a high number of aircraft and accept more than planned, when their workload is low, which seems quite natural. Our point is that this relation is not our hypothesis about controllers behavior. It is the visualization of the real data. We use 95 days of data, leading to > 9000 sample points to obtain these averages. But, plotting the sample means like in figure 3 is statistically unreasonable because the underlying points might not be independent. This may result in misleading estimates (e.g. [8]). This means that we cannot draw conclusions from figure 3 alone.
To overcome this, we do the following: the trend looks like a concave, monotonic increasing function. Examples for such functions are logarithm-, square root- or reciprocal functions. They are the dotted functions in figure 3. In the range of our interest ($0 \leq PLN \leq 12$) they have similar behavior. We will analyze how regression models like $\mathbf{E}\{REAL_t^S\} = \alpha log(PLN_t^S)$ explain the variation in the real data. This particular example is displayed as the dotted lines in figure 2. A good agreement with the data is perceivable. We will verify the conditions to draw valid conclusions from our analysis in section IV.

b) Uncertainty: Counting the number of arrivals $REAL_t^S$ results in right skewed distributions for every time interval t (except early morning and late night). Also, in all cases in which r aircraft are planned to enter the sector ($PLN_t = r$), the distribution of the real number of arrivals is right skewed (figure 4). The trend curve in figure 3 corresponds to the mean values of these distributions. More formally, we are dealing with the (unknown) conditional distributions

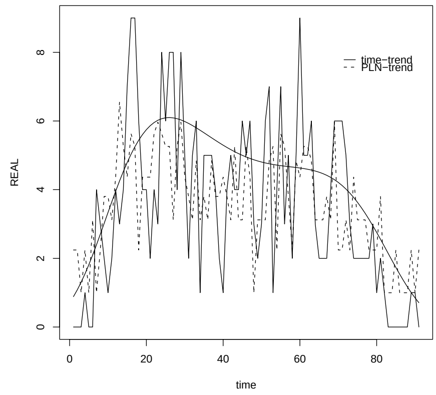


Fig. 2. Arrivals in fifteen minutes intervals with t^4 time-trend (bold) and logarithmic PLN trend (dotted)

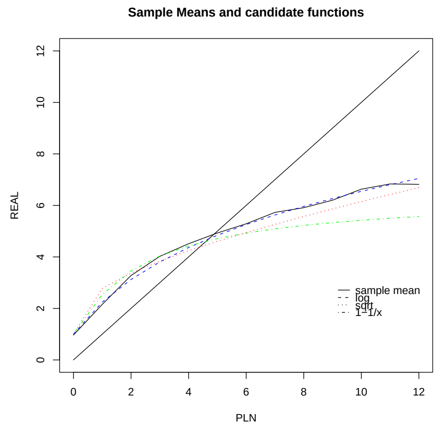


Fig. 3. Relationship between number of planned arrivals PLN (x-axis) and average number of real arrivals $REAL$ (y axis) in 15 minutes time intervals. The bold line corresponds to the sample means. The dotted functions are candidates for modeling this trend: logarithmic, square-root and 1-1/PLN.

$Pr(REAL_t^S | PLN_t^S)$ and their expected values.
The hypotheses that these distributions correspond to Poisson distributions where the parameter is the sample mean could not be rejected in nearly all the cases. No interesting departures from a Poisson distribution could be found neither. This gives an idea of the variation of the number of arrivals in a sector because mean and variance are the same for a Poisson distribution. However, it is not enough information to unambiguously draw conclusions about the underlying mechanism of the phenomenon. For further information please consult [9], [7], [10] or [11].

IV. Validation

In this part we collect evidence that our findings are a reasonable basis for interpretation and that they can be generalized to other sectors and other days. We analyze 31 sectors, covering the the Upper Airspace between London,

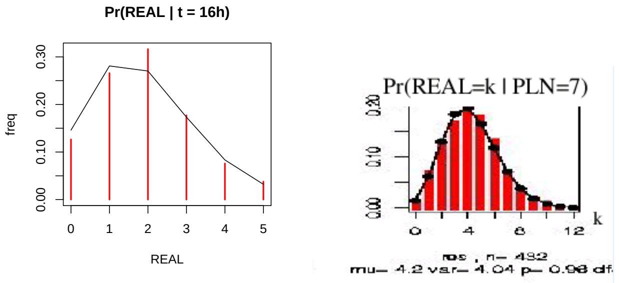


Fig. 4. Examples for distributions of number of real arrivals ($REAL_t$) dependent on time of day (left) and on number of planned entries (PLN_t) (right). Superposed are Poisson distributions. Their parameter is the sample mean.

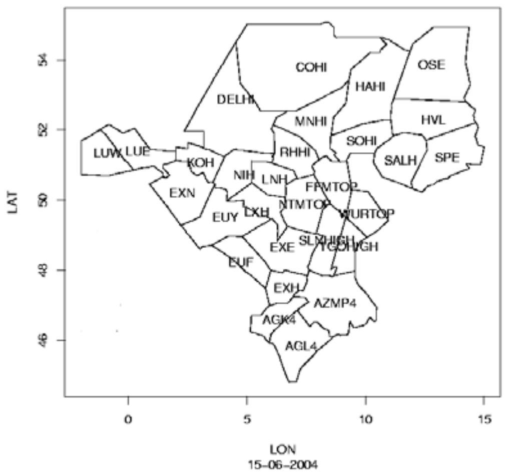


Fig. 5. Central European Upper Airspace. 31 sectors covering the area above London, Zurich and Berlin

Zurich and Berlin (figure 5). Data is available for 95 week days (Mon-Thu) in the period 13.5.-29.9.2004.
The validation procedure consists of two parts. First, we show graphically that the same type of trend appears in all sectors under study. Then, we randomly select sectors and days and analyze regression models of this trend.

c) Data and Time intervals: We experimented with 5 minutes and 15 minutes time intervals. Interesting results were obtained for the 15 minutes intervals. For the graphical validation we use the whole data set, leading to > 9000 points for each sector to estimate the mean values.
For each of the regressions we select randomly one sector S and one day of week data d (Mon-Thu), denoted by
$D_d^S = \{(REAL_{t_1}^S, PLN_{t_1}^S), ..., (REAL_{t_{96}}^S, (PLN_{t_{96}}^S)\}$.

d) Graphical Validation: Figure 6 shows the relationship between PLN and $REAL$ for 12 sectors from Central Upper Airspace. They all show the same logarithm-like shape. The fluctuations at the end of the intervals can be explained by few underlying data. For all the 31 sectors, 68 % of the asymptotes lie in the interval $(8, 10]$. Since we are

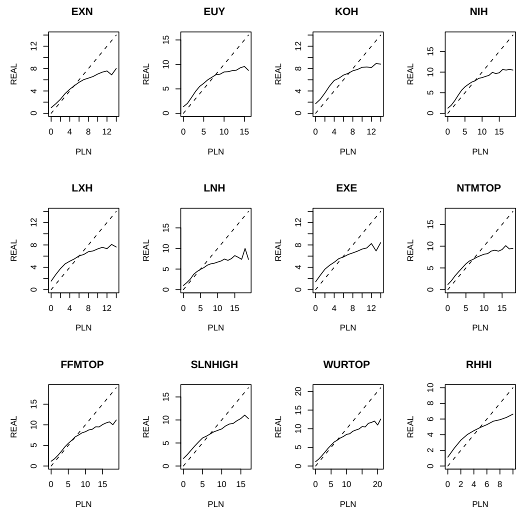


Fig. 6. Same shape of trend for all sectors: plots of the sample means

working with 15 minutes intervals, this corresponds to hourly workloads between 36 and 40 aircraft. This is roughly the declared capacity of many sectors. 26 % lie below and 6 % above this interval
In 63 % of the cases, the points, where $PLN = REAL$ lie between $6 \leq PLN < 8$. In 35 % of the cases, these points are below 6. Only in one case, point where PLN equals $REAL$ lies above 8.
This suggests that controllers accept this 'traffic pattern'; no rerouting is necessary. For other plannings, controllers re-organize the flows to improve their working conditions.

As we said above, plotting the sample means like in figures 3 or 6 is statistically unreasonable because points $(PLN_{t_i}^S, REAL_{t_i}^S), (PLN_{t_j}^S, REAL_{t_j}^S), i \neq j$ might not be independent. Estimators for the mean value under the false assumption of independence have high variance [8]. This means that we cannot draw conclusions from figures 3 and 6 alone.

e) Regression Analysis: Beside the logarithm-like relationship that we investigate, are there other factors that influence the variables $REAL_t^S$? Please remind that we are not modeling the traffic flows themselves, but the relationship between planned and real traffic.

We analyze three models for this aim:

1) A simple model is to assume

$$\mathbf{E}\{REAL_t^S\} = \alpha f(PLN_t^S)$$

It models the mean value of $REAL_t^S$ as a simple function of PLN_t^S, but one that has similar shape than the sample means in figure 3 (e.g. $sqrt$, $1-1/x$ or log). It assumes that the real traffic is *only* dependent on the planned traffic of the same sector and in the same time interval.

Model	$\mu(\mathbf{R^2})$	$\sigma(\mathbf{R^2})$
$log(PLN)$	0.94	0.02
$sqrt(PLN)$	0.94	0.02
$1-1/PLN$	0.93	0.03
$\sum_{i=0}^{4} t^i$	0.94	0.02

TABLE I

COMPARISON OF DIFFERENT TREND MODELS

2) A more realistic model takes into account the neighborhood of a sector. Let $S^{N_i}, i = 1, .., m_S$ be the neighbors of sector S. Let $PLN_t^{\mathbf{N}(S)} = \sum_{i=1}^{m_S} PLN_t^{S^{N_i}}$ be the sum of planned traffic in the neighborhood of S.

$$\mathbf{E}\{REAL_t^S\} = \alpha_1 f(PLN_t^S) + \alpha_2 PLN_t^{\mathbf{N}(S)}$$

This model expresses that the arrivals at a sector S depend additionally on the traffic density in the neighborhood.

3) Finally, one may think that the impact of the two independent variables in model 2 is not independent (the additivity assumption is not satisfied). To overcome this, we allow for interaction:

$$\mathbf{E}\{REAL_t^S\} = \alpha_1 f(PLN_t^S) + \alpha_2 PLN_t^{\mathbf{N}(S)} + \alpha_3 (f(PLN_t^S) \cdot PLN_t^{\mathbf{N}(S)})$$

Please see below for details on the interaction term.

All models are 'static': they explain the number of real arrivals in fifteen minutes windows independently of the past. Behind this is the assumption that, from a controllers point of view, it is of no importance what happened 15 minutes ago. More complex models would take into account spatio-temporal dependencies.

The main assumption of the models is that the variables PLN contain all of the information to explain the variable $REAL$. A covariate that explains all structure of a time varying variable is not unusual in time-series contexts [7]. This is a difference to ARMA type models, which would assume dependencies between the variables $REAL_t$ themselves.

For validation we proceed as follows: we select randomly sectors and days and analyze the resulting regression models. We will perform residual analysis to check our assumptions.

Estimation All models are linear. The variables $REAL_t^S$ count the number of arrivals in a time interval t. Their distributions are right skewed (figure 4). A natural regression technique for such variables is Poisson regression (e.g. [4]). Another possibility to treat skewed variables is to transform them logarithmically and to use ordinary least squares for estimation. This is what we do for this paper because formal inference is not our priority and because the technique is known in a wide audience.

Results

We fit every candidate model to 30 randomly selected sectors and days. Figure 7 shows a typical instance. In Residual vs fitted and Normal QQ plot (first and second plots) the points should be homogeneously distributed around 0 and build a straight line respectively. These are graphical methods to test for homogeneity of variance and normality, two assumptions of a linear model. The third plot is explained below. In the fourth plot, the sample auto-correlation function is shown. Since we are analyzing a phenomenon in time, the absence of correlation in the residuals is crucial to verify before making inferences.

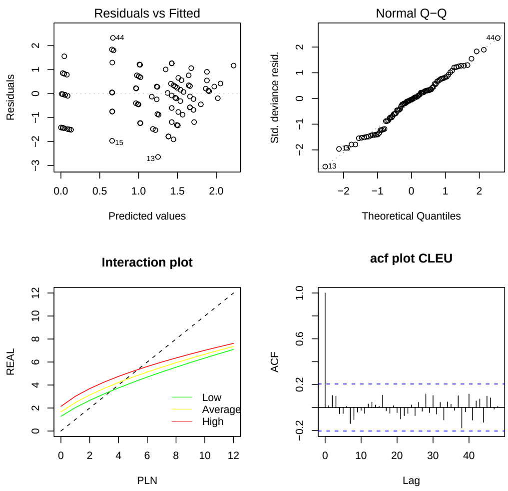


Fig. 7. Typical instance for validation of residuals

It turns out that negative auto-correlation of -0.2 to -0.3 at lag 1 appears in about 10 % of the residuals of the simple type models (first and second). Thus, it is likely that other factors influence the variable $REAL$ and inference about the validity of the logarithm-like trend cannot be made.

For the third model, auto-correlation appears in less than 3 % of the cases. The other residual checks are valid, as well. This is evidence that the model assumptions are correct. The coefficients for this model are almost always significant on a 1 % level. For this model, we explain the third plot in figure 7. It shows the predicted number of real arrivals against planned arrivals for three different levels ('low', 'average', and 'high') of the 'neighborhood' variable. The levels correspond to the mean value plus minus one standard deviation respectively. To analyze this in more detail, the predictions can be grouped into three classes: either they follow the the same line (90 %), they intersect in the second half (6 %) or their starting values diverge (figure 8). Here, we multiplied the variables in the interaction term, because they are both numeric. When the neighborhood variable is transformed to a factor with three levels, the resulting model exhibits sometimes neighborhood levels for which the predicted values are always above or always below the diagonal. On the same hand, in such a model, not all parameters are systematically significant. The reasons for this have to be found in another study. For our purpose,

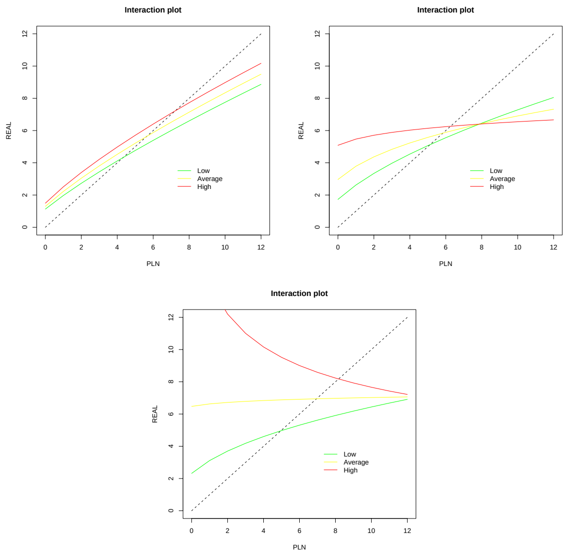


Fig. 8. Three types of predictions. Same (left), with changing effects (middle) and diverging starting values (right)

we can say that there is still systematically another number of arrivals than planned. A finer analysis will determine whether there are some exceptions from this. Table I shows mean values and standard deviations of R^2 for each of the 30 regressions. The first three models are the logarithmic, square root and reciprocal functions of PLN_t. They all have $\mu(R^2) \approx 94\%$ with small standard deviation. This means that independently of sector and day, the explained variation is roughly the same. The fourth order polynomial time trend is shown in the last line $(\mu(R^2) = 93\%)$. This is similar to the others. The advantage of the logarithmic like trend models is that they lead to insight in how controllers deal with their workload, whereas the polynomial time trend can be less easily interpreted. For a discussion on model formulation and selection we refer the reader to chapter 1.9. in [12] and the references therein.

As a conclusion we can say that there is evidence for our initial observation: that systematically another number of aircraft than planned enters flight sectors. Moreover, the traffic density in the neighborhood seems to be a useful exlanatory variable.

V. Interpretation

We remind that our study is 'observational'. Interpreting regression models in such contexts can be misleading, for example due to unobserved variables or correlation between regressors (e.g. [12]).
However, our results show some evidence that there are systematically more aircraft than planned arriving for few planned traffic and less for high numbers of planned traffic. One would expect rather a fluctuation around the planned number, if the flow planning were accurate. This is not the case. We found thus a relationship between the planned traffic and the average realized traffic that asks to be interpreted. It gives insight into how controllers treat their workload: they reroute aircraft (in time or in space) for high numbers of planned traffic and accept rerouted vehicles in periods with low traffic density. This seem quite natural. The point is that we "learned" from our data analysis the nature of this relation: its shape can accurately be described by a logarithmic, square root or reciprocal function (with different regression coefficients for different sectors). We found dependency with the traffic density in the neighboring sectors. We have seen that there is exactly one point, where correspondence between planned and realized traffic exists in average. This suggests that controllers accept this 'traffic pattern'; no rerouting is necessary. For other plannings, controllers re-organize the flows to improve their working conditions.
We have also seen the existence of asymptotes for realized traffic. No matter how many aircraft are planned to enter a sector, controllers will take actions to maximally let this amount of aircraft enter. These asymptotic values are in agreement with the declared capacities.
Concerning uncertainties we showed that these average values of realized traffic can be seen as the mean values of Poisson distributions. This gives us an idea about the variation of realized traffic around the mean values because mean and variance are the same for a Poisson distribution.
Last but not least, our models are 'static'. They explain the number of real traffic in fifteen minutes windows independently of the past. This suggests that from a controllers point of view, it is of no importance what happened 15 minutes ago.
All these findings are empirical. We found ways to accurately *describe* the data and its variation. But the question is not whether one description 'performs better' than another. We should confirm *why* the data shows exactly this behavior before drawing further conclusions.

VI. Conclusion and Future Work

We found a relationship between the planned number of sector entries and the realized number. It suggests that in average, more aircraft than planned arrive when few were planned and that less than planned arrive when many were planned. This relationship can be described accurately by logarithmic, square root or reciprocal functions. Moreover, we showed that these averages can be seen as the mean values of Poisson distributions.
Taking both together, the uncertainty about the real number of aircraft arriving at a sector can be described.
We validated the findings on a large number of sectors, randomly chosen from the Central European Upper Airspace. Our results are empirical but they lead to insight in how controllers deal with their workload. However, we should confirm why the data shows exactly these characteristics before drawing any further conclusions. Elements of answering this question could be:

- Time invariance: we did not look into the question whether the trend is the same over the whole day. A positive answer to this question would simplify any future model. A more detailed statistical analysis should answer this question.

- Micro-dependencies: our approach is macroscopic; we aggregate variables of sector entries. This leads to a loss of information about dependencies between successive aircraft in a flow. A finer analysis, based on stochastic Point processes for example, can lead to insight into such dependencies.
- Interaction with neighboring sectors: based on our findings more concrete hypotheses of how controllers share workload with neighboring sectors can be formulated.

This work is a contribution in the characterization of uncertainties in current traffic flow planning.

VII. Acknowledgements

The authors would like to thank Jean-Marie Marion for a discussion about the validity of the models.

References

[1] Eurocontrol. Independent Report for the Improvement of ATFM. *Eurocontrol*, 2002.
[2] Eurocontrol. Toulouse-Blagnac Delay Model. Synthesis of the Study. Technical report, EEC Note 14/02, 2002.
[3] S. Stoltz and P. Ky. Reducing traffic bunching through a more flexible air traffic flow management. In *Fourth International Air Traffic Management R&D Seminar ATM-2001, Santa Fe (New-Mexico, USA)*, 2001.
[4] J.K. Lindsey. *Modelling Frequency and Count Data*. Oxford University Press, 1995.
[5] D.R. Cox and P.A.W. Lewis. *The Statistical analysis of series of events*. London, Methuen, 1966.
[6] P. Guttorp. *Stochastic modeling of scientific data*. Chapman and Hall/CRC, London, 1995.
[7] J.K. Lindsey. *Statistical Analysis of Stochastic Processes in Time*. Cambridge University Press, 2004.
[8] P.J. Diggle. *Time Series: A Biostatistical Introduction*. Oxford University Press, 1990.
[9] D.J. Daley and D. Vere-Jones. *An Introduction to the theory of Point Processes*. Springer, 1988.
[10] D.R. Cox and V. Isham. *Point Processes*. Chapman and Hall, 1980.
[11] C. Gwiggner. Implicit structures in ATFM. In EUROCONTROL, editor, *INO Annual Report 2005*, Bretigny, France, 2005.
[12] D.R. Cox and N. Wermuth. *Multivariate Dependencies. Models, Analysis and Interpretation*. Chapman and Hall/ CRC, 1998.

Demand Modeling for Pre-scheduled Group Travel: Charter Service vs. Scheduled Service

Gautam Gupta, Mark Hansen

***Abstract*—We consider the case of airline travel made by groups of people, where the itinerary of the group is known in advance. A known demand and schedule coupled with numerous such groups moving in network validates the possibility of a profitable charter service for this scenario. We develop a demand model that models consumer decision making on a network level, identifying flights in the network that are served by the charter service, and subsequent routing of the aircraft to minimize consumer expenditure. We develop an Integer Programming based model for this purpose, that resembles a fleet assignment problem, and test it for a particular case.**

***Index Terms*—Charter Service, Demand, Integer Programming**

I. Introduction

Aost September 11 and the subsequent carnage and uncertainty in the airline industry ,certain new business concepts in aviation are being explored, based on the availability of Regional Jet aircrafts and Very Light Jet aircrafts (VLJ's). These include fractional ownership concepts like Netjets, and on demand services like Dayjet. One possible service area for such aircraft is the charter industry. Specifically, in cases where a group of people travel frequently in order to perform (athletes, performing troupes etc), a charter service based on up to 50 seat aircraft has a lot of appeal because of the following reasons:

- Assigning groups instead of individuals to small or medium sized aircraft overcomes the question of "filling" the aircraft to its capacity to increase the revenue
- Since the focus is on pre-scheduled group travel, the demand is deterministic. This creates opportunities to fine tune schedules and operations.
- "Tailoring" air travel to customer needs could mean utilizing an airport closer to the actual destination, which would not have been used earlier due to infrequent scheduled flights. Usage of closer airports, coupled with "anytime" and convenient service has the capability to attract a lot of passengers.
- A large variety of groups would benefit from such a service. These include professional and collegiate athletes, performing troupes etc.
- Last but no the least, with a large number of unemployed RJ's in desert storage, an entrepreneurial airline could derive a lot of benefit from employing such aircraft.

The focus of this paper is to analyze the demand for such a charter service. The work was motivated by the interaction with MVP Air (http://www.mvpair.com), a charter service company focusing on student athlete travel. The objective is to model consumer decision making on a network level, given the prices for charter service and the price and schedule of a scheduled carrier. This work is part of a broader research agenda to analyze the competition between a charter airline and scheduled airline, specifically for pre-scheduled group travel. Such a charter service is part of a future trend in "personalization" of airline industry. Such shifts in trends become all the more important given the cascading effect that they have on the entire aviation system.

Any potential customer for a charter service would compare the charter option with the commercial scheduled service option. The customer will chose the option from which he or she can derive the maximum benefits. These benefits include both cost and convenience.

Besides tackling large group sizes (10 people or more), the modeling framework presented in this paper can be extended to "personalized" air transportation service, where an individual or small group of people have a high value of time. For the rest of the paper, charter airline refers to a charter airline which flies the customer whenever they want, and a scheduled airline represents an established airline which has a predefined set of prices and schedules.

II. Model Formulation

Let us consider the operations of a charter operator. The operator will have a limited fleet, and we assume that the price to the consumer for flying between any origin and destination is pre-defined by the operator. Besides the price for flying between any origin and destination, the consumer would also be charged the cost of dead-heading the empty aircraft. For the given price and fleet size, there is a certain subset of the consumers' demand that the charter operator would be able to satisfy. The key question is, what subset of the total demand would be assigned to the charter service for a given price and fleet size. This problem of assigning a limited fleet to demand

Manuscript received February 26, 2006.

Gautam Gupta is a Graduate Student and Researcher at the National Center for Excellence in Aviation Operations Research (NEXTOR), University of California, Berkeley, CA 94720-1720 USA. (phone: 510-282-3860; fax: 510-643-8919; e-mail: ggupta@berkleey.edu)

Mark Hansen is co-director, NEXTOR, and professor at University of California, Berkeley, CA 94720-1720 USA (phone: 510-642-2880; fax: 510-643-8919; e-mail: mhansen@ce.berkeley.edu)

spread over a network resembles a typical airline fleet assignment problem, as discussed below.

Commercial airlines use demand estimates over the various links of a network, and assign their fleet to the various links at different times in order to minimize their cost of operation. This involves identifying the demand estimates, planning a schedule to best serve the demand, and routing the fleet through the network to satisfy this schedule with a minimum cost of operation.

Schedule planning involves determining when and where to offer flights so that profits are maximized, while fleet assignment involves assigning aircraft types to flight legs to minimize operating cost. Schedule planning and fleet assignment are not independent of each other, since scheduling a certain flight requires the availability of an aircraft with sufficient capacity to adequately serve the demand for the flight. However, given the large size of the problem of, historically the decision making step has been broken down into steps, where each step is optimized in a sequential manner. Most of the time, schedule planning starts from an existing schedule with a well developed route structure. Historical data, marketing estimates and allocation models may all be used to project probability distribution of demand for each flight leg in the schedule (Rushmeier et al, 1995). Once legs and their departure times are finalized, fleet assignment is done. Barnhart and Lohatepanont (2004) present integrated models that optimize selection of flight legs and assignment of aircraft types to each flight leg. The approach to schedule design is incremental, i.e., modifications are introduced to a base schedule.

Once a flight schedule is developed, the next step is to determine which aircraft would serve each flight. This is referred to as the fleet assignment problem. An airline would like to maximize its profit, which would involve minimizing the assignment cost. Assignment cost is a function of the which aircraft is assigned to each flight, and includes flight operating cost, passenger carrying related cost and spill cost. Spill cost is the revenue lost when assigned aircraft cannot accommodate every passenger (Barnhart et al, 2002).

Coming back to our initial problem of determining what subset of group travel demand would be served by charter aircraft for a given price and fleet composition, we can draw parallels in airline fleet assignment and our problem. Assuming that the scheduled service is another “option” or “fleet component” to serve all the demand, the problem decomposes into a fleet assignment problem. The various flights are assigned to either charter aircraft or scheduled carrier, such that the cost of satisfying the entire demand is minimized. The option of flying on a scheduled service is treated as another fleet component, although with no constraints on availability. The constraints on a flying on a scheduled service flight would be in terms of time when the flight is scheduled, whereas the constraints flying charter would be the availability of aircraft. Since flying on a scheduled service would involve “schedule delays” due to waiting for the flight, some penalty for serving a particular flight through a scheduled service would be incorporated in the consumer price for serving that flight through a scheduled service. The difference between airline fleet assignment and our problem would be that there would be no spill cost, since we assume that all the demand would be fulfilled.

Typically, fleet assignment is modeled as a multi-commodity network flow problem. The nodes in the underlying network flow problem correspond to time intervals in which interconnection activity occurs. The equipment types are the commodities, and the decision variables correspond to flow along the network arcs. Operations like maintenance and crew scheduling provides restrictions on the type of aircraft that can serve a flight leg. Barnhart et al (2002) propose an itinerary based model that captures network effects and gives better estimates of spill and recapture of passengers. A noticeable weakness is the requirement of fixed departure times. Rexing et al (2000) assign time windows to each flight, and discretize this window to allow for flexibility in schedules. This approach is computationally expensive, and they present two algorithmic approaches for solving this model.

Given that previous approaches to fleet assignment assume fixed schedules, or at the most assume time windows of serving flights at the expense of computational time, we looked at an alternate way to formulate the problem of assigning aircraft to charter flights. Assigning charter aircraft to fixed schedules is not realistic in the charter case, as groups traveling together for prescheduled activities would not be averse to reasonable adjustments on the departure times, when these adjustments are done in the planning stage. For example, a college baseball team intending to leave on Friday for a game on Saturday would be willing to leave after classes end on Friday, and reach the destination in order that it gets a full nights rest. The desired departure times would not be rigid, but lie in a certain interval. Thus, we develop a method in which flights are assigned in intervals. These intervals or “window of indifference” are continuous time intervals, and the customer is indifferent to departing anytime within this window. Thus, the formulation here differs from typical fleet assignment formulations in two ways:

- It does not assume fixed schedules for flight departure, but considers intervals during which the flight can depart, giving better utilization of available fleet.
- The departure interval is treated in a continuous manner, instead of discretization as done before (Rexing et al, 2000).

A charter service differs from a scheduled service primarily in travel time (possibly due to stops due to aircraft change) as well as “schedule delay” (difference in the preferred departure/arrival time and actual departure/arrival time). We internalize these costs in our formulation, to realistically model the consumer decision. The time components have been internalized by assigning a value of time and adding them to the cost of using that option. For the scheduled option, the “time cost” involves the time spent in flying and value of schedule delay beyond the window of indifference. The charter time cost involves the cost of time spent in traveling. Since the costs are calculated on a flight-by-flight basis, different value of time can be assumed for different groups.

The formulation minimizes the cost of serving the entire demand using both charter and commercial service. The formulation presented here assumes only one kind of aircraft for the charter service, based on the size of the group. This model can easily be extended to multiple group sizes and multiple aircraft types.

We further introduce the concept of *level of service* (*d*) (window of indifference). A group might not be averse to leaving some time before the latest departure time (in case of venue-bound flights) or after the earliest departure time (in case of home-bound flights). We define this flexibility in scheduling as level of service. A level of service of 2 hours means that a venue bound flight can served anytime between the latest departure time and 2 hours prior to it. Similarly, home bound flight can be served anytime between the earliest departure time and 2 hours after it. The integer programming formulation is as follows:

$$\min\left(\begin{array}{l}\sum_{f}\sum_{k} y_{k,f}\times C_f+\sum_{f} x_f\times S_f\\ +\sum_{k}\sum_{f1}\sum_{f2} z_{k,f1,f2}\times C_{e,f1,f2}\end{array}\right) \tag{1}$$

such that:

$$x_f, y_{k,f}, z_{k,f1,f2}\in\{0,1\} \quad \forall f, f1, f2\in Flights, \forall k\in Planes \tag{2}$$

$$\sum_{k} y_{k,f}+x_f=1 \quad \forall f\in Flights \tag{3}$$

$$z_{k,f1,f2}\geq y_{k,f1}+y_{k,f2}-\sum_{i=f1+1}^{f2-1} y_{k,i}-1 \quad \forall f1, f2\in Flights\ (f2>f1), \forall k\in Planes \tag{4}$$

$$t_{k,f}\leq y_{k,f}\left(T_f+d_f\right) \quad \forall f\in Flights, \forall k\in Planes \tag{5}$$

$$t_{k,f}\geq y_{k,f}\times T_f \quad \forall f\in Flights, \forall k\in Planes \tag{6}$$

$$\begin{array}{l}t_{k,f2}-t_{k,f1}-y_{k,f1}\left(H_{f1}+L_{f1,f2}\right)\\ +\left(T_{f1}+d_{f1}+H_{f1}+L_{f1,f2}\right)\left(1-y_{k,f2}\right)\geq 0\end{array} \quad \forall f1, f2\in Flights\ (f2>f1), \forall k\in Planes \tag{7}$$

$$z_{k,f1,f2}\left(T_{f2}+d_{f2}-T_{f1}-H_{f1}-L_{f1,f2}\right)\geq 0 \quad \forall f1, f2\in Flights\ (f2>f1), \forall k\in Planes \tag{8}$$

$$\left(T_{f2}+d_{f2}-T_{f1}-H_{f1}-L_{f1,f2}\right)/M\geq y_{k,f1}+y_{k,f2}-2 \quad \forall f1, f2\in Flights\ (f2>f1), \forall k\in Planes \tag{9}$$

where:

Planes	is the set of all the charter aircraft,
Flights	is the set of all flights, ordered on the basis of desired departure time
$y_{k,f}$	is 1 if flight *f* is flown by plane *k*, zero otherwise
x_f	is 1 if flight *f* is flown by scheduled service, zero otherwise
$z_{k,f1,f2}$	is 1 if plane *k* is used to flights *f*1 and *f*2, and flies empty in-between to relocate
C_f	is the charter flying cost for flight *f*
S_f	is the scheduled flying cost for flight *f* (including cost of delays)
$C_{e,f1,f2}$	is the charter flying cost for flying empty between flights *f1* and *f2*
T_f	is the earliest desired flight time for flight *f*
$t_{k,f}$	is the actual departure time of flying for flight *f* if flown by plane *k*.
d_f	is the “window of indifference” (or the level of service) after T_f for flight *f*
H_f	is the flying time for flight *f* using charter service
$L_{f1,f2}$	is the empty flying time for charter aircraft for relocating between flights *f*1 and *f*2
M	is a large number

x, *y* and *z* are the binary decision variables. The objective function incorporates all the cost incurred in serving all the demand. As shown in equation (1), it consists of 3 terms. The first term represents the cost of flying charter for any demanded flight, whereas the second term represents the cost of flying commercial. The third term represent the cost for flying the charter aircraft empty to re-position it for the next flight.

Equation (3) represents the constraint that all demand has to be served, by using either a charter aircraft or scheduled service. Equation (4) gives constraints on the binary variables used to identify the empty flights made by the charter aircraft for repositioning between consecutive flight of the same aircraft. $z1_{k,f1,f2}$ is a binary variable which is one only if the flights *f*1 and *f*2 made by plane *k*, and plane *k* serves no other flight between these 2 flights.

Equation (5) and (6) give bound on the value of $t_{k,f}$. Equation (5) states that $t_{k,f}$ is less than the a desired departure time plus the window of indifference, and is 0 if flight *f* is not served by plane *k*. Equation (6) states that $t_{k,f}$ is greater than or equal to the desired departure time if flight *f* is served by plane *k*, and zero otherwise.

Equation (7) is the overlap constraint. It states that if a particular plane serves two flights, then there should be sufficient time for the plane to serve the first flight and dead-head to the origin of the next flight.

Equation (8) and (9) are further overlap constraints, incorporated for better bounds in the solution. They state that between any two flights, if the time of deadheading and serving previous flight is such that the next flight cant be served at any time in its "window of indifference", then no plane will serve both the flights.

III. Model Application

As an application of the model developed in section 3, we tested it for student athlete travel, specifically for college athletic conferences. The highlight of team travel for athletic conferences is that there are numerous movements of students among the various members of the conference. Servicing athletic conferences has some operational benefits, namely

- Limited geographical expanse of the various origins and destinations
- The possibilities of "shuttling" the teams, i.e., avoid dead heads or repositioning empty aircraft.

We analyze the case of student travel for the Pacific 10 athletic conference, the details of which follow.

A. Data Details

The Pacific 10 (PAC 10) athletic conference is comprised of 10 schools. The member schools, their location and closest airports are given in table 1. We had historical schedule data for all the teams of the 10 schools, which we converted into set of demand of flights between certain origins and destinations. There were some instances of *trip chaining* (if the origin is A, the teams flies from A to B, B to C and then back to A). We developed a code in PERL that analyzed such scenarios and generated the appropriate flight demand data. The cost data for flying the charter aircraft was taken for the Dornier D328 aircraft. This aircraft has 30 seats, and is suitable for most teams besides football. The fares for commercial flying were sought from the travel website sidestep (http://www.sidestep.com).

As discussed before, the cost component for each flight includes the cost of time. Quantifying the value of time involves assigning some value of time for each flight. Estimating the amount of money a customer would be willing to spend to save a certain amount of time is not easy. Value of time would depend on many factors, some of them being the purpose of travel, time of day etc. Moreover, in order to check the profitability of the enterprise, an airline would look at the worst-case scenario, which would be that time does not translate into money. Based on this premise, we use zero as the value of time for this analysis. As for the price to the consumer of using the charter service, we assume that the PAC-10 conference owns the aircraft. Thus, the price of flying on the charter service is the operating cost of the aircraft, and the charter service is operated on a no-profit basis.

The operating cost of the charter aircrafts would include the hourly cost of operating the aircraft (including crew cost, fuel cost etc) and the cost of ownership of the aircraft. In order to determine the ownership cost of each plane, we divided the yearly lease cost of each aircraft into daily ownership cost. This daily ownership cost was included in the total cost of serving all the demand on the days of the flights only. This is a conservative approximation of the cost, since the ownership cost for the rest of the year has not been included. A worst case approach would be to include this cost for the entire year. We use the conservative approach since the assigning the cost for the entire year is not realistic, as the charter aircraft might be utilized elsewhere when not serving PAC-10.

TABLE I
MEMBER UNIVERSITIES AND RELEVANT DATA FOR PACIFIC 10 ATHLETIC CONFERENCE

School	Code	Location	Airport
Arizona State	AZS	Tucson, AZ	TUS
UC Berkeley	CAL	Berkeley, CA	OAK
Oregon State	ORS	Corvallis, OR	EUG
Stanford	STA	Stanford, CA	OAK
Univ. of Arizona	UAZ	Tempe, AZ	PHX
UC Los Angeles	UCLA	LA, CA	LAX
Univ. of Oregon	UOR	Eugene, OR	EUG
Univ. of Southern California	USC	LA, CA	LAX
Univ. of Washington	UWA	Seattle, WA	SEA
Washington State	WAS	Pullman, WA	PUW

B. Results

Given the huge size of the problem (for 10 days with 10 flights each and fleet size of 3, the number of variables is nearly exceeds 15,000), we decided to split the entire schedule into 7 day periods, and optimized for each of these 7 days periods separately. We looked at historical schedule of all the games in the PAC-10 conference, and used all the games from May 2004 to April 2005 for generating the demand for flights. We wrote a script in PERL that converted each sports event in to location and time data. This script also checked for trip chaining as mentioned before. We encoded the formulation and data in AMPL, and used CPLEX 8.0 to solve it for the above demand data. As a test case, we considered two fleet sizes, of 3 aircraft and 5 aircraft. We ran the formulation for the d values of zero and 2. Figure 1 and 2 give the results.

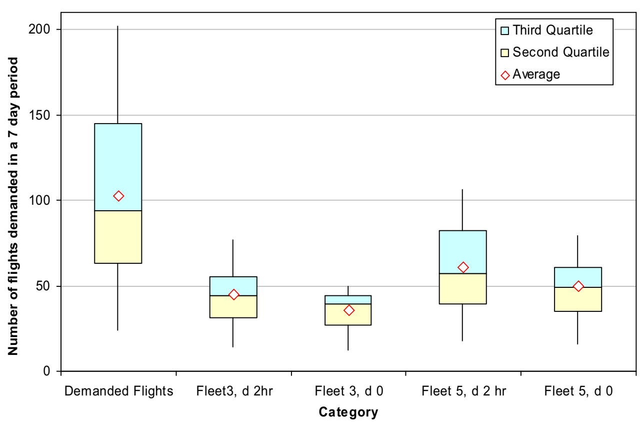


Fig. 1. Box and whisker plot depicting the results from running the model in PAC 10 conference. The leftmost box and whiskers depict the total demand and its variation over different weeks. The rest depict the demand served by the charter service for four different scenarios.

Figure 1 is a box and whisker plot that shows the variation in the number of demanded flights in the different 7-day periods. The box and whiskers on the left depict the number of flights demanded over a week. The next four box and whiskers depict the number of flights served by the charter service for four different scenarios. The number flights that can be served by a given fleet size changes substantially with increase in *los* value. A higher fleet size, obviously, allows more flights to be served by charter service.

Figure 2 does the cost comparison of the costs involved. The first bar on the left shows the cost of serving all the flights using only the scheduled service. The subsequent bars represent the costs for different alternatives in charter service, namely fleet size of 3 and 5, and *d* values of 0 and 2 hours.

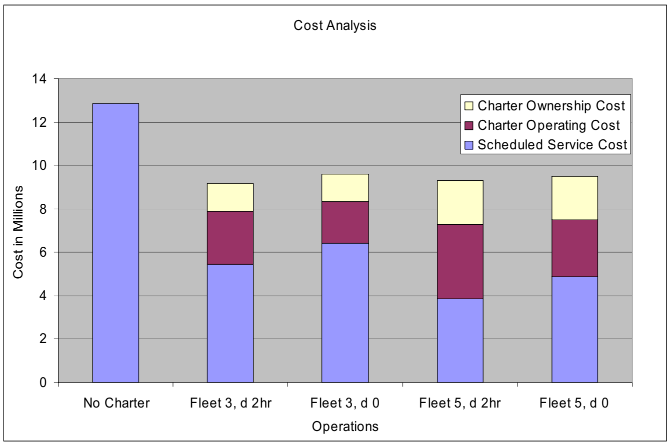


Fig. 2. Comparison of total travel cost. The leftmost bar indicates cost if only scheduled service is used. The rest of the bars represent the cost of serving all the demand for four different scenarios of the charter service.

The difference in the left-most column and the subsequent columns can be regarded as the savings in cost, if the PAC-10 conference decides to operate its own aircraft. This makes the business case for any prospective charter operator, where the difference described above can be regarded as profit. Further, it can be seen from figure 1 and 2 that there is not a substantial change in the number of flights served by charter service between the options of (d = 0) and (d = 2 hrs). This is primarily due to the high number of games that occur in the evening, making it difficult to serve numerous flights with a limited fleet. A large fleet allows for more of the demand to be served by the charter service, but due to the presence of ownership cost, there is an "optimal" fleet size for a given demand. As can be seen from figure 2, the cost of operating with 5 aircraft is similar to operating with 3 aircraft. Cost for operating with more than 5 aircraft is higher than both 3 and 5 aircraft.

IV. Future Work

The model described in this paper is demand model, that simulates consumer decision making given the prices for the charter and scheduled services, and the schedules. It minimizes the consumer expenditure. The obvious question is, how should the charter service be priced in order to maximize the profit of the charter operator. Different prices will lead to different choices by the consumer, and the effect of these choices over the entire network needs to be considered as a whole. One of the future goals of this research is to embed this model in a larger model that identifies the optimal pricing strategy for the charter operator, given the price and schedule of the scheduled service operator.

In the present formulation, the number of constraints explodes with the number of flights and the number of planes. In the application studied here, we have divided the schedule into parts, and analyzed each part separately. In the future, we intend to develop methods to tackle a larger part of the problem in a single go. Given the huge number of constraints, Lagrangian Relaxation based techniques appeals as a logical option. We intend to look into the application of such techniques to approach larger problem sizes.

Lastly, it remains to be seen how an established scheduled service operator will compete with a charter operator. For example, what happens if the scheduled service operator decreases the price of tickets. A realistic model of the competition between charter service and scheduled service will go a long way in identifying the future trends in the airline industry.

References

[1] Aboulafia, R., "Crisis of Faith in Regional Jets", Industry Insights, *Aerospace America*, May 2003

[2] Barnhart, Cynthia, Kniker, Timothy S. and Lohatepanont, Manoj, "Itinerary Based Fleet Assignment", *Transportation Science*, Vol. 36, No. 2, pp 199-217, May 2002.

[3] Bishop, M. and Thompson, D., "Peak-load Pricing in Aviation: The Case of Charter Air Fares", *Journal of Transport Economics and Policy*, Vol. 26, No. 1, January 1992

[4] Lee, J. J., Lukachko, S. P., Waitz, I. A. and Schafer, A., "Historical and Future Trends in Aircraft Performance, Cost and Emissions", *Annual Review Energy Environment*, 2001, 26:167-200

[5] Lohatepanont, Manoj and Barnhart, Cynthia, "Airline Schedule Planning, Intergrated Models and Algorithms for Schedule Design and Fleet Assignment", *Transportation Science*, Vol. 38, No. 1, pp 19-32, February 2004

[6] Mehndiratta, S. R., "Time-of-day Effects in Inter-city Business Travel," *PhD Dissertation* submitted to Civil Engineering, University of California, Berkeley, 1996

[7] Poole, R. W., Jr. and Butler, Viggo. "Airline Deregulation: The Unfinished Revolution", *Regulation,* vol. 22, no. 1, Spring 1999, pp. 8.

[8] Rexing, Brian, Barnhart, Cynthia, Kniker, Timothy S., Jarrah, Ahmad and Krishnamurthy, Nirup, "Airline Fleet Assignment with Time Windows", *Transportation Science*, Vol. 34, No. 1, pp 1-20, February 2000.

[9] Rushmeir, R. A. and Kontogiorgis, S. A., "Advances in the Optimization of Airline Fleet Assignment", *Transportation science*. Vol. 31, No. 2 May 1997

Passenger Trip Time Metric for Air Transportation

Danyi Wang, Dr. Lance Sherry, and Dr. George Donohue

Abstract— **the traditional Air Transportation metrics, such as flight delays and flight cancellations, do not accurately and completely reflect the actual flight experience of passengers, and underestimate time penalties of cancellations on passenger trip time.**

This paper describes and demonstrates algorithms calculating statistical measures of total passenger flight times between an origin/destination pair. The algorithms convert flight data to passenger trip data, based on a large quantity of public assessable historical data. A passenger-based metric, estimated passenger delay, is proposed to measure on-time performance from a passenger's perspective. Ranking at origin-destination pairs, airport, and other results of analysis of 2004 data are provided.

Index Terms—**Passenger Flow, On-Time Performance Measurement, Metrics, Network Properties**

I. Introduction

The flight-based on-time performance metrics do not accurately reflect delays on passengers [1]. Bratu and Barnhart collected passenger complaints and negative reports for period 1995 to 2000, and compared them with flight-based performance metrics. They discovered the flight-based metrics could not explain the sharp increase of complaints and negative reports in 2000.

In Figure 1, we compare the flight-based metrics, such as number of delayed flights and number of flight cancellations, with passenger complaints collected by DOT on flight delays and cancellations in period 2002-2004. In year 2002, the air transportation system has 868,225 delayed[1] arrivals, 65,143 cancellations, 829 complaints on flight delays, and 720 complaints on cancellations. We use those metric values in 2002 as baseline and plot the ratios of year 2003 and 2004. As the number of delayed and cancelled flights (the first two bars) increased in 2003 and 2004, the corresponding complaints (the third and fourth bars) decreased. The discrepancy between flight-based metrics and passenger' feedback is clearly observed.

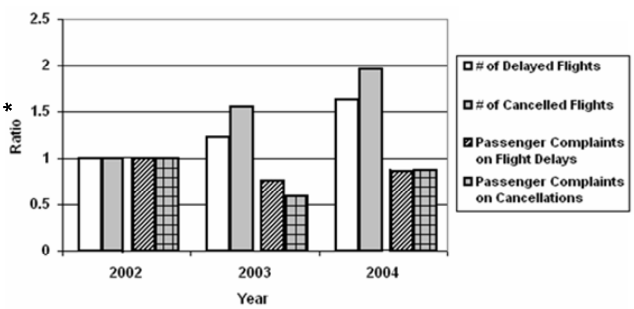


Fig. 1 Comparison between flight-based metrics and passenger's complaints on disrupted activities

* Y axis, "ratio" means the metric value in year 2003 or 2004 divided by the value of the same metric in baseline 2002. For example: # of delayed flights in year 2003 / # of delayed flight in baseline 2002 = 1.25

Delays, missed connections and cancellations are the reasons caused difference between flight experience and passenger experience. Firstly, flight-based metrics are constrained by the unit they use (per flight). They do not consider passenger related factors, like load factor, aircraft size, etc. So it is hard for them to explain the complicated scenarios of passenger flow and travel disruptions. Secondly, flight-based metrics underestimated the serious impacts of cancellations on passenger trip time. In August 2004, passengers on cancelled flight had been delayed by 13 hours averagely according to our estimation.

Consider the following two cases. Day 1 and day 2 both have 1000 scheduled flights and 10 of them are cancelled. The on-time performance of these two days will have no difference according to flight-based metrics. However, if the ten cancellations in day 1 happened in the late afternoon, within a short time period, and from one single airline, they will generate enormous amount of passenger delays due to shortage of resources. The ten cancellations in day 2 have a much scattered distribution through the day, and they are from different airlines that have lower load factors. Then passenger delays in day 2 could be much less compared with day 1.

"Flight delay has been one of the key indicators of system performance, and will continue to be an important indicator, more sophisticated delay metrics are needed to provide a more complete picture of performance (Bolczak and. Hoffman 1997)". To passengers, delay and cancellation are essentially the same. They both impose time penalties on passenger's travel time. In this paper, we proposed a passenger-based on-time performance metric, *passenger delay*, to evaluate the on-time performance of domestic non-stop flights.

This paper is organized as follows. The background information and previous research are included in Section 2. Detailed algorithms involved in converting flight data to passenger trip data are discussed in Section 3. Major results are presented in Section 4, and finally conclusion and references are presented in Section 5 and 6.

II. BACKGROUND

Transportation systems are commonly represented using networks as analogy for their structure and flows. Physically, the air transportation network consists of routes and airports. Functionally, it has route characteristics like flight frequency, distance, etc associated with each route, and airport characteristics like airport capacity, runway layouts, etc. associated with each airport. The network as a whole can be affected by systematic characteristics like delay propagation, system capacity limitation, etc. The routes in the network have specific behaviors and these behavior patterns form the network properties.

The operational view of the air transportation network performance has been well developed. There are:

-- Flight metrics such as flight delays, flight cancellations, etc. used to measure the performance.

-- Simulation tools, such as TAAM, Vensim, etc., built to simulate network structure and to predict impacts on network performance given potential changes.

-- Research paper concerned in network structure and performance [3]-[12], etc.

The passenger or user's view of the network performance is deficient. Major reason could be the lack of passenger trip data, which can only provided by airlines.

Bratu's research on flight schedule reliability [2] has made a breakthrough in passenger on-time experience. He was provided with a few months' passenger booking data by a single legacy airline. Combined these passenger booking data with ASPM data, he analyzed the passenger delays generated by missed connections and flight cancellations, and did sensitivity analysis of the relationship between factors like load factor, flight frequency, and passenger delays. Unfortunately, Bratu's research is constrained by a few months' data from a single airline, hence is not sufficient to be expanded to a system level analysis.

Table 1 shows the difference between Bratu's research and the algorithms described in this paper. They are differentiated by four major aspects.

TABLE I
DIFFERENCE BETWEEN BRATU AND WANG

Aspects	Bratu	Algorithms in this paper
Data	Few months' passenger booking data from an single legacy airline, combined with ASPM flight operation data	Public accessible databases on BTS (Bureau of Transportation Statistics): AOTP*, T-100*, and etc.
Goal	Airline flight schedule reliability	Passenger view of system on-time performance
Perspective	Airline's perspective	Passenger's perspective
Model	Testify different airline strategies for disrupted relocating passengers	Understand and predict passenger flow through the network

* AOTP is the Airline On-Time Performance data (from BTS)
* T-100 is the Air Carrier Statistics (Form 41 Traffic) domestic data (from BTS)

III. METHODOLOGY

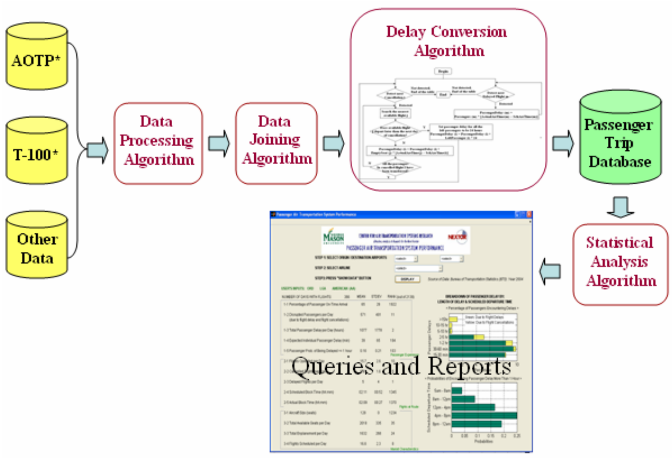
Passenger delays are caused by three kinds of disrupted activities: delays, missed connections and cancellations. The goal of methodologies is to estimate passenger delay caused by these disrupted activities, using public accessible flight-based databases.

Figure 2 illustrates the converting process from flight data to passenger trip data. Raw data is collected from BTS public accessible databases:

- Airline on-time performance database – It contains "departure delays and arrival delays for non-stop domestic flights by major air carriers, and provides such additional items as origin and destination airports, flight numbers, cancelled or diverted flights and etc." (Bureau of Transportation Statistics)

- Air carrier statistics database – It contains domestic non-stop segment data by aircraft type and service class for passengers, freight and mail transported, available capacity, scheduled departures, departures performed and aircraft hours, etc." .(Bureau of Transportation Statistics)

The ATOP data is flight-based, which means one record in the database represents a single flight in the system. It contributes operational flight data to the final passenger trip



* ATOP = Airline On-Time Performance data (from BTS)
* T-100 = Air Carrier Statistics (Form 41 Traffic) domestic data (from BTS)

Fig. 2 Converting Process from flight data into passenger trip data database. The T-100 data is aggregated data, which contributes passenger and seat data to the final passenger trip database.

Data processing and data joining algorithms have the following functions, as showed in Table 2.

TABLE 2
FUNCTIONS OF DATA PROCESSING AND DATA JOINING ALGORITHMS

Algorithms	*Functions*		
Data process algorithm	Extract useful data	Detect errors or abnormal data	Repair or reformat data
Data joining algorithm	Restructure databases	Link databases	Join databases

After the data has been filtered, reorganized and rejoined, the most important algorithm, delay conversion algorithm is applied on it to convert flight data into passenger trip data. Passenger delays are estimated in this process.

A. Delay Conversion Algorithm

Total passenger trip delay is a function of aircraft flight delays and cancellations. Tracing each passenger, we are able to find whether he/she arrived on time. If his/her flight was delayed, how long he/she has been delayed. If the flight was cancelled, which flight he/she was relocated to, and compared to the original schedule, how late he/she arrived at destination. Unfortunately, passenger travel data is not available to public. Therefore, we developed delay conversion algorithm, which is designed to estimates passenger delays caused by flight delays and flight cancellations.

A flight here defined as a departure-arrival process. If a flight i is delayed, the associated passenger delays are calculated as:

PassengerDelay(i) = (# of passenger loaded on flight i) * (ActualArrTime(i) – SchArrTime(i)) (1)

We assume passengers in cancelled flights will be relocated to the nearest available flights belong to the same carrier and have the same origin-destination pair, if the available flights still have empty seats to fit in more passengers. Generally passengers from a cancelled flight will be relocated to 2 or 3 different flights due to limited empty seats on each available flight. For a specific passenger from the cancelled flight, the delay he or she experienced is calculated as the time difference between actual arrival time of the flight he or she relocated to, and scheduled arrival time of the original scheduled flight, that is the cancelled flight. Finally, the total passenger delays caused by this cancellation equal to the summation of passenger delays each passenger in the cancelled flight experiences separately.

$$TotalPassengerDelay(i) = \sum_{j=1}^{n} \{p(i) * [ActualArrTime(j) - SchArrTime(i)]\} \quad (2)$$

Here, i = cancelled flight
j = available flight
n = total number of available flights needed to finish relocating passengers on cancelled flight i
P(i) = number of passengers relocated from flight i to flight j
ActualArrTime(j) = actual arrival time of flight j
SchArrTime(i) = scheduled arrival time of cancelled flight i

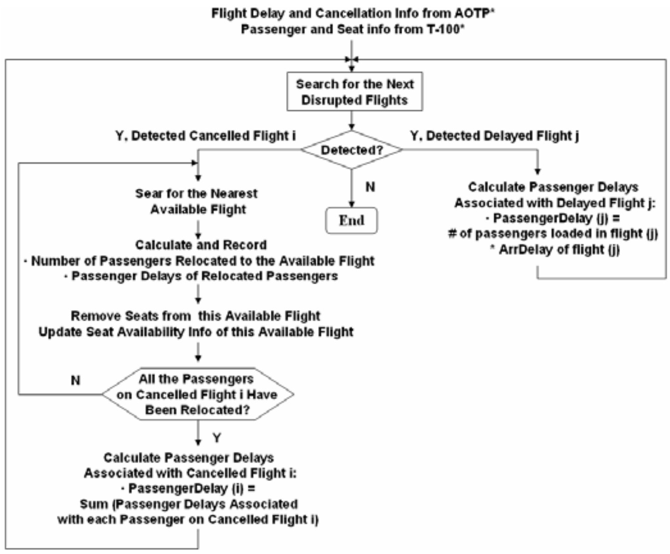


Fig. 3 Main process of delay conversion algorithm

The major process of the algorithm is shown in Figure 3.

Table 3 gives an example of estimating passenger delays caused by cancellation. Assume a flight with 100 passengers is cancelled. Its scheduled arrival time is 12:00 pm. The first available fight has 30 empty seats and it arrives at 2:00 pm. The second available flight has 45 empty seats and it arrivals at 3:00 pm. The third available flight has 40 empty seats and it arrivals at 4:00 pm. So the passengers relocated to the first available flight will be delayed by 2 hours each. The passengers relocated to the second available flight will be delayed by 3 hours each, and 4 hours each for those relocated to the third available flight. Therefore the passenger delay caused by the cancelled flight is

TABLE 3
EXAMPLE: ESTIMATE PASSENGER DELAYS CAUSED BY CANCELLATION

Available Flights	*Empty Seats*	*Relocated Pax from cancelled flight*	*Empty Seats Left*	*Delay expd by each pax relocated to this flight*	*Total delay expd by pax relocated to this flight*
Available Flights1	30	30	0	2hrs	30×2=60 hrs
Available Flights 2	45	45	0	3 hrs	45×3=135 hrs
Available Flights3	40	25	15	4 hrs	25×4=100 hrs
		Total = 100 relocated passengers			Total = ***295*** hrs

* expd = experienced; pax = passenger(s)

2(hr)*30 + 3(hr)*45 + 4(hr)*25 = 295 hours.

Assumptions for delay conversion algorithms are:

- To be conservative, we set 15 hours as the cap of passenger delays [1].
- Carrier and its subsidiaries will help each other relocating passengers from cancelled flights. For example, if a flight of American Airline (AA) is cancelled, passengers in this flight will be relocated to the nearest available flights, no matter it is from American Airline (AA) or its subsidiary American Eagle (MQ).

IV. RESULTS

We focus our research on domestic flights through OEP35[2] airports. OEP35 airports have the greatest number of operations and are heavily traveled. They account for 73% of total enplanements and 69% of total operations in the air transportation system [12]. The closed network formed by OEP35 airports generates 1044 directed routes in 2004.

Note: results shown in this section are all from the closed network formed by OEP35 airports in 2004.

A. Cancellations Disproportionally Generate High Passenger Delays

The disruption to passenger trip time caused by cancellations is underestimated. Simple metric, like number of cancellations, does not tell the complexity of passenger relocation process and the huge delays on passengers. According to our estimation results, passengers on cancelled flights in August 2004 were averagely delayed by 13 hours.

We depict the passenger delays at OEP 35 airports in 2004 in Figure 5. Total passenger delays consist of two parts: passenger delays caused by cancellations and passenger delays caused by flight delays. Averagely, 40% of total passenger delays were caused by flight cancellations, while cancellations only accounted for 1.7% of total flights. Among OEP35 airports, ORD generated the largest number of highest total passenger delays. At some of the OEP35 airports, such as MCO, TPA, CVG, IAD, and etc., cancelled flight which accounted for less than 3% of total flights, produced more than half of the total passenger delays.

As noted in Figure 5, passenger delays at different airports have specific patterns. For a single airport, passenger delays are seasonal dependent. We normalized the passenger delays by enplanements in order to prevent impacts from different level of enplanements in each season. Figure 6 shows the monthly normalized passenger delays at ORD in 2004.

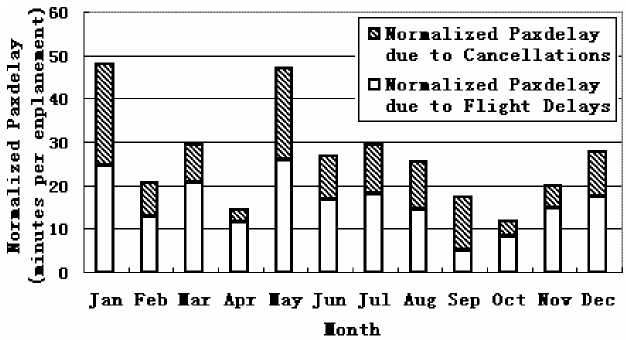


Fig. 6 Example: seasonal change of passenger delays at ORD in 2004

As shown in Figure 6, normalized passenger fell below 400 thousand hours in low season such as April and October, but began to grow through the next 3 months until they peaked in midseason like January and May. The seasonal changes in passenger delays caused by cancellations are clearly evident. In general, passenger delays caused by cancellations are more sensitive to seasonal changes. Cancellations happen in midseason might have stronger impacts on passenger delays than cancellations in low season.

Flight experience describes the operational view of the air transportation system performance, while passenger experience describes the passenger view of the air transportation system performance. They review the system performance from different perspectives. Combining them together could provide us more complete and accurate views of the air transportation system performance.

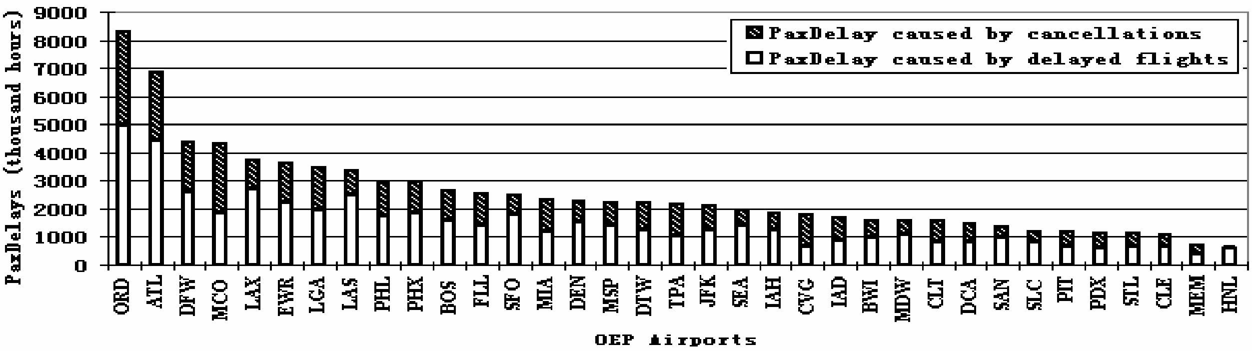


Fig.5 Passenger delays caused by flight delays and flight cancellations

[2] According to Airport Capacity Benchmark Report published by FAA in 2004, the OEP35 airports includes: ATL, BOS, BWI, CLE, CLT, CVG, DCA, DEN, DFW, DTW, EWR, FLL, HNL, IAD, IAH, JFK, LAS, LAX, LGA, MCO, MDW, MEM, MIA, MSP, ORD, PDX, PHL, PHX, PIT, SAN, SEA, SFO, SLC, STL, TPA

B. Validate Consistency between Passenger's Complaints and Estimated Passenger Delays

In Section 1, we depicted discrepancy between flight-based metrics and passenger's flight experience. We proposed a passenger-based metric, "passenger delay", to measure on-time performance from a passenger's perspective. Can this metric accurately reflect passenger's experience? A consistency check is shown in Table 4.

TABLE 4
EXAMPLE: CONSISTENCY CHECK BETWEEN P ESTIMATED PASSENGER DELAYS AND PASSENGER'S COMPLAINTS (AUGUST 2004)

Flight-Based Metrics v.s. Passenger's Complaints		*Estimated Passenger Delays v.s. Passenger's Complaints*	
# of cancellations per complaint on cancellations	94	*Average PaxDelays caused by cancellations per complaint on cancellations*	98
# of delayed flights per complaint on flight delays	1896	*Average PaxDelays caused by delayed flights per complaint on flight delays*	98

Averagely there was one complaint on flight cancellations reported to DOT for every 94 cancellations, and one complaint on flight delays reported for every 1896 delayed flights. Though the number of delayed flights was much higher than the number of cancellations, average complaints on cancellations were 20 times more than complaints on flight delays. This can be well explained using passenger delays. As long as passengers experience an average delay of 98 minutes, they reported complaints on the disrupted activities, not matter the delays were caused by flight delays or by cancellations. Figure 6 confirms that passenger delays consist well with passenger's on-time experience.

C. *Differentiate Service on Routes*

Routes in the system are not homogeneous. Each route has specific behavior. Routes have different levels of enplanements and flights due to limited airport and airspace capacities. We normalized the passenger delays by enplanements and flight delays by flights in order to achieve a fair comparison among routes. Figure 7 shows rankings of routes in terms of normalized passenger delays. Route JFK-ORD generated the biggest normalized passenger delays, 98 minutes per enplanement, in 2004. Large normalized passenger delays don't necessarily indicate large normalized flight delays, since factors affecting passenger delays, like cancellations, aircraft size and load factor, do not affect flight delays. As observed in Figure 7, routes with the largest normalized passenger delays, such as JFK-ORD and PHL-EWR, are not necessary the routes with the largest normalized flight delays.

Rank	Airport Code	Normalized Passenger Delays (minutes per enplanement)	Normalized Flight Delays (minutes per flight)	Prob. (Passenger Delay >=1 hour)
1	JFK-ORD	98	38	0.053
2	PHL-EWR	80	14	0.073
3	JFK-PHL	79	25	0.090
4	EWR-PHL	74	25	0.075
5	PHL-JFK	59	22	0.071
6	BWI-EWR	57	18	0.060
7	JFK-BWI	57	21	0.061
8	STL-PIT	56	26	0.080
9	DTW-HNL	56	54	0.085
0	PHL-STL	52	26	0.139
11	PHL-IAD	49	15	0.064
12	JFK-CLE	49	17	0.056
13	PIT-EWR	47	20	0.098
14	JFK-PIT	47	17	0.086
15	EWR-BWI	47	17	0.084
16	STL-PHL	45	19	0.060
17	IAD-EWR	45	21	0.053
18	BWI-LGA	45	45	0.109
19	MDW-CVG	44	8	0.068
20	CVG-ORD	44	18	0.079
...	...	...	...	...

Fig. 7 Sample rankings of routes in terms of normalized passenger delays

Scheduled trips of passengers might be seriously disrupted when the passenger delays are large, especially for connecting passengers. Risk analysis of passenger being delayed more than an hour is provided in the last column of Figure 7. Risk of long delays is a different concept to normalized passenger delays. It emphasizes more on number of passengers being severely delayed, instead of mean of passenger delays. Figure 7 confirms that high risk of long delays does not always followed by large normalized passenger delays.

Figure 8 gives another example of the divergence of passenger experience and flight experience on different routes. The top 20 most crowed routes are ranked by total emplanements in 2004, and the top 20 busiest routes are ranked by total number of departures in 2004. Large aircrafts are flied on routes with most enplanements. Conversely, routes with high flight frequency prefer smaller aircrafts than large aircrafts, especially on those short haul routes with commuter flights, such as SAN-LAX, LGA-BOS and DCA-BOS.

TOP 20 MOST CROWDED ROUTES				TOP 20 BUSIEST ROUTES			
Rank	Route	Enplanements	Avg Aircraft Size (Seats)	Rank	Route	# of Departures	Avg Aircraft Size (Seats)
1	LAX-ORD	1,310,749	167	1	SAN-LAX	14,525	34
2	ORD-LAX	1,301,040	163	2	LAX-SAN	14,507	36
3	LAS-LAX	1,235,815	139	3	BOS-LGA	13,450	100
4	ATL-MCO	1,225,073	199	4	LGA-BOS	13,077	100
5	MCO-ATL	1,220,497	199	5	DCA-LGA	13,040	100
6	LAX-LAS	1,210,615	138	6	LGA-DCA	13,011	100
7	ATL-DFW	1,178,334	140	7	LAX-LAS	12,486	138
8	DFW-ATL	1,143,173	141	8	LAS-LAX	12,482	139
9	LAX-HNL	1,105,922	164	9	ORD-MSP	11,922	108
0	ATL-LGA	1,098,481	126	0	ATL-DFW	11,763	140
11	HNL-LAX	1,096,112	189	11	MSP-ORD	11,717	108
12	LGA-ORD	1,092,498	163	12	DFW-ATL	11,387	141
13	TPA-ATL	1,087,063	128	13	LGA-ORD	11,285	126
14	LAX-JFK	1,085,796	163	14	ORD-LGA	11,219	128
15	ORD-LGA	1,083,400	188	15	PHX-LAS	10,804	140
16	LGA-ATL	1,075,956	139	16	LAS-PHX	10,757	139
17	ATL-TPA	1,065,572	163	17	DCA-BOS	10,500	78
18	LAS-PHX	1,062,651	140	18	LAX-PHX	10,466	122
19	JFK-LAX	1,051,082	206	19	BOS-DCA	10,423	78
20	PHX-LAS	1,049,438	208	20	PHX-LAX	10,419	122
...	...	...	...	...	...	...	...

Fig. 8 Sample rankings of routes in terms of enplanements and departures

After normalizing the passenger delays with total enplanements on routes, we observed that flights on some of the short haul routes in Eastern area produced the largest normalized passenger delays (minutes per enplanement). Figure 9 shows the histogram of normalized passenger delays on 1044 routes formed by OEP35 airports in 2004. Notably, routes on the tail with normalized passenger delays more than an hour are all short haul flights in Eastern area.

In addition, those routes not only produced the highest normalized passenger delays, but also have highest risk of passenger being delayed more than an hour. The top 3 routes with highest probability of passenger being delayed more than an hour are EWR-BOS with probability 0.31, BWI-BOS with probability 0.23, and BWI-MEM with probability 0.22. Passenger delays propagated disproportionally from those

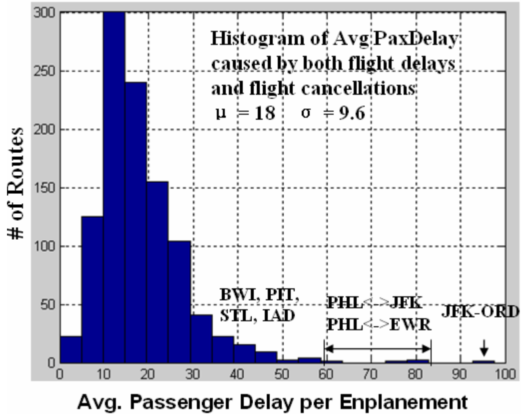


Fig. 9 Histogram of normalized passenger delays (minutes per enplanement)

eastern airports to other connected airports. Based on our estimation, 26% of the total passengers arrived at OEP35 airports in 2004 were transported from airports in eastern area[3]. But these 20% passengers propagated 35% of total passenger delays to their destinations.

V. CONCLUSIONS

This paper shows the divergence of flight on-time experience and passenger on-time experience. In order to achieve a more accurate and complete measurement of the air transportation system performance, a passenger-based metric, "estimated passenger delay", is proposed in the paper.

Passenger delay measures the time penalties on passenger trip time due to both flight delays and flight cancellations. It captures the large delays caused by small amount of cancellations, and well reflects the actual delays to passengers. Compared with passenger delay analysis conducted by Bratu and Barnhart, algorithms in this paper founded on public accessible databases, and extend the analysis from restricted level (single airline, short time period) to system-level.

Future work is to apply the passenger relevant properties on different routes to build a simulation model that tracks and predicts passenger flow though the air transportation network.

Advent of internet and web-based ticketing system has enabled passengers to place a greater emphasis on ticket value. We hope to adjust passengers' behavior by providing them with disruption-risk analysis of the routes and airlines they chose. Helping passengers, the direct buyers and users of the service, to differentiate themselves by providing both ticket value and disruption-risk prediction, could prompt the airlines to differentiate their services.

[3] Airports in eastern area include: PHL, LGA, EWR, JFK, IAD, DCA, BWI, CVG, PIT and ORD

VI. REFERENCES

[1] Stephane Bratu, Cynthia Barnhart, "An Analysis of Passenger Delays Using Flight Operations and Passenger Booking Data", *Journal of Transportation and Statistics*, Vol 1, 2005. .

[2] Stephane Bratu, "Airline Passenger On-Time Schedule Reliability: Analysis, Algorithms and Optimizaiton Decision Models", *Ph.D Thesis*, Massachusetts Institute of Technology, Cambridge, MA.

[3] Oster and Strong, "Evolution of U.S. Domestic Airline Route Networks Since 1990", *TRB conf*, 2006.

[4] Cheih Yu Hsiao and Mark Hansen, "An Econometric Analysis of U.S. Airline Flight Delays with Time-of-Day Effects", *TRB conf*, 2006.

[5] W. Post, "Metrics and Performance Management", *6th USA/Europe ATM R&D Seminar*, 2005

[6] Dispasis Bhadra and Pamela Texter, "Airline Networks An Econometric Framework to Analyze Domestic US Air Travel.", *Journal of Transportation and Statistics*, Vol. 7 Number 1, 2005.

[7] R. Guimera and L.A.N.Amaral , "Modeling the world-wide airport network", *The European Physical Journal*, B38 pp.381-385, 2004

[8] Paul T. R. Wang, Lisa A. Schaefer, and Leonard A. Wojcik, "Flight Connections and Their Impacts on Delay Propagation", *MITRE Technical Paper*, 2003.

[9] Yoshinori Suzuki, "The Relationship Between On-Time Performance and Airline Market Share: A New Approach", *Transportation Research* Part E 36, pp. 139-154, 2000

[10] Zalman A. Shavell, "The Effects of Schedule Disruptions on the Economics of Airline Operations", 3rd USA/Europe Air Traffic Management R&D Seminar Napoli, 2000.

[11] Stephen Earl Foreman and Dennis G. Shea, "Publication of Information and Market Response: The Case of Airline on Time Performance Reports", *Review of Industrial Organization 14*: 147-162, 1999.

[12] Catherine N. Bolczak and Jonathan H. Hoffman, "National Airspace System Performance Measurement", MITRE Technical Report, 1997.

[13] JD Power and Association (2004), "Airport Passenger Satisfaction Study", Westlake Village, CA.

[14] Eurocontrol, "Performance Review Report", 2004

[15] Bureau of Transportation and Statistics, On-Time Performance - Flight Delays at a Glance. Available: http://www.transtats.bts.gov/HomeDrillChart.asp?URL_SelectYear=2005&URL_SelectMonth=4&URL_Time=1&URL_Selection=1

[16] Bureau of Transportation and Statistics, Airline On-Time Performance Data. Available: http://www.transtats.bts.gov/Tables.asp?DB_ID=120&DB_Name=Airline%20On-Time%20Performance%20Data&DB_Short_Name=On-Time

[17] Bureau of Transportation and Statistics, Form 41 Traffic T-100 Domestic Segment Data. Available: http://www.transtats.bts.gov/Tables.asp?DB_ID=110&DB_Name=Air%20Carrier%20Statistics%20%28Form%2041%20Traffic%29&DB_Short_Name=Air%20Carriers

[18] Department of Transportation, Air Travel Consumer Report. Available: http://airconsumer.ost.dot.gov/reports/index.htm

[19] Federal Aviation Administration (2004), Airport Capacity Benchmark Report 2004.

Distance between aircraft trajectories related to functional data analysis

Daniel Delahaye
Applied Math Laboratory
Air Navigation Research Center
7 Avenue Edouard Belin
31055 Toulouse, FRANCE
delahaye@recherche.enac.fr

Stephane Puechmorel
Applied Math Laboratory
French Civil Aviation School
7 Avenue Edouard Belin
31055 Toulouse, FRANCE
puechmor@recherche.enac.fr

Abstract— **In this paper the problem of aircraft trajectories representation and analysis is addressed. In many operational situations, there is a need to have a value expressing how trajectories are close to each other. Some measures has been previously defined, mainly for trajectory prediction applications, all of them being based on distance computations at given positions in space and time. The approach presented here is to consider the trajectory as a whole object belonging to a functional space and to perform all computations in this space. In a second part of the paper, the way of representing those infinite dimensional objects by finite dimensional approximation is discussed.**

I. INTRODUCTION

Future Air Traffic Management relies, in part, on the use of decision support tools (DST) to provide improved service to the user community under increasing traffic demand. Furthermore, this improvement has to be validated by the mean of system performance metrics such as complexity, robustness, capacity. As aircraft fly 4D trajectories, there is a strong need to quantify the associated trajectory accuracy in order to validate aircraft models and trackers. Such validation is usually based on a comparison between the actual trajectory and a reference by the mean of a trajectory distance. This last point is the key element of the whole process. Such trajectory distance is still needed for ATM applications and the goal of this paper is to present a new trajectory distance based on rigourous mathematical concepts. Although trajectories are well understood and studied, relatively little investigation on the precise comparison of trajectories is presented in the literature. A key issue in performance evaluation of ATM decision support tools (DST) is the distance metric that determines the similarity of trajectories. Most existing measures [6], [17] compute a mean distance of the corresponding positions of two equal duration trajectories. Supplementary statistics such as variance, median, minimum, and maximum distances are also suggested to extend the description of similarity. In recent work, Needman [12] proposed an alignment based distance metric that reveals the spatial transition and temporal shift between the given trajectories, and introduced an area based metric that calculates the total enclosed area between trajectories using trajectory intersection.

One main disadvantage of the existing approaches is that they are all limited to the equal duration (lifetime) trajectories. By duration we refer to the number of coordinate points that constitute the trajectory. These coordinates are sampled at different instances. Since the existing measures depend on the mutual coordinate correspondances, they cannot be applied to trajectories that have different durations. Conventional distance measures assume that the temporal sampling rates of trajectories are equal. They do not cope with the uneven sampling instances, i.e. varying temporal distance between the coordinates. Therefore, there is a need to develop other alternatives that can effectively measure the difference between unrestricted trajectories.

The first part of this paper presents a list of ATM applications for which an exact trajectory distance is needed. The second part presents some current trajectory distance metrics and shows their limitations. The third part gives a detailed mathematical description of our new trajectory distance and finally, the fourth part presents the associated algorithm implementation of such a distance.

II. CURRENT TRAJECTORY DISTANCE

A aircraft trajectory is a time sequence of coordinates representing the aircraft path over a period of time and may be represented by a N-uple :$T = \{(x_1, y_1, z_1, t_1), (x_2, y_2, z_2, t_2), ..., (x_N, y_N, z_N, t_N)\}$ where N is the duration.

The simplest metric used for computing the distance between a pair of trajectories is the mean of coordinate distance, which is given as

$$m_1(T^a, T^b) = \frac{1}{N}\sum_{n=1}^{N} d_n$$

where the displacement between the positions is calculated using the Cartesian distance

$$d_n = [(x_n^a - x_n^b)^2 + (y_n^a - y_n^b)^2 + (z_n^a - z_n^b)^2]^{\frac{1}{2}}$$

Note that, the mean distance metric makes three critical assumptions :

1) the durations of both trajectories are the same :$N^a = N^b = N$

2) the coordinates are synchronized $t_n^a = t_n^b$
3) the time sampling rate is constant $t_{n+1}^a - t_n^a = t_{m+1}^a - t_m^a$

It is evident that the mean of distance is very sensitive to the partial mismatches and cannot deal with the distortions in time.

To provide more descriptive information, the second order statistics such as median, variance, minimum and maximum diatance may be incorporated. For instance variance trajectory distance is defined as

$$m_2(T^a, T^b) = \frac{1}{N} \sum_{n=1}^{N} (d_n - m_1(T^a, T^b))^2$$

Although these statistics supply extra information, they inherit (even amplify) the shortcomings of the ordinary mean of distance metric m_1. Besides, none of the above metrics is sufficient enough by itself to make an accurate assessment of the similarity.

An area based distance metric is proposed in [12]. The crossing point $q : T^a(p_i) = T^b(p_j)$ of two paths are used to define regions $Q_j, j = 1, .., J$ between trajectories. For each region, a polygon model is generated and the enclosed area is found by the parameterized shape. The resulting distance is given by :

$$m_3(T^a, T^b) = sum_{n=1}^{N} area(Q_j)$$

This metric can handle more complex trajectories, however it is sensitive to entanglements of the trajectory, it discards the time continuity, and fails to distinguish two trajectories in opposite directions. Furthermore, it is not adapted to 3D trajectories.

The next section present a new trajectory distance which can handle 4D aircraft trajectories in a rigorous manner.

III. Computations on trajectories

Since objects of interest are aircraft trajectories, we need to find an adapted framework in which computations may be made on trajectories as a whole. There are basically two ways of understanding what a trajectory is :

- The time/position approach. In this case, a trajectory can be represented as a mapping from a bounded interval of $\mathbb{R}$ (the life time of the trajectory) to $\mathbb{R}^3$ or $\mathbb{R}^6$ depending on whether speed is part of the data or not. Since there is an explicit dependance on time, there is a need to calibrate trajectories with time shifts for all applications involving trajectory comparison (supervised or unsupervised classification, mean trajectory estimation, bounding tubes ...). We will see in the following that there is nevertheless a mean of reducing the problem so that origin of time is automatically calibrated. For simulation applications or for trajectory prediction/interpolation, this will be the natural framework.
- The shape approach. Here, trajectories are understood as paths and time is not directly relevant (from a more formal point of view, we take the quotient of the trajectories understood as mappings by the group of diffeomorphisms acting on time), so that we may assume that the underlying life time of trajectories is always the interval $[0, 1]$. This is the right framework for dealing with optimal sectorization problems, alternate routes generation or major flows estimation.

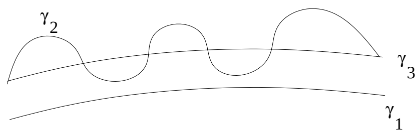


Fig. 1. Different trajectories with same sup distance

A. *Trajectories as mappings*

We will assume in the following that trajectories are given as mappings from a compact interval of $\mathbb{R}$ to $\mathbb{R}^3$. The case of mappings from $\mathbb{R}$ to $\mathbb{R}^6$ (that is with explicit speed, for example as given by radar tracking filter) can be derived with minor changes and thus will not be adressed here. Since physical trajectories are smooth unless there is a perturbating noise, we made the choice to take all trajectories as smooth mappings from a compact interval of $\mathbb{R}$ to $\mathbb{R}^3$. The first point to deal with is the necessary calibration of the origin of time for trajectories comparison. Remembering that there is an explicit dependance on time, one cannot just time shift one trajectory in time in order to make it coincident with another in order to compare them : this will result in forgotting distortions in time, that is trajectories with the same range (as mappings) but different positions at different times may become equal. A possible candidate for the distance between two trajectories γ_1 and γ_2 will simply be to take the supremum norm, that is :

$$\sup_{s \in \mathbb{R}} \|\gamma_1(s) - \gamma_2(s)\|$$

Since γ_1 and γ_2 are constant outside bounded intervals of $\mathbb{R}$, the supremum is well defined. However, this metric is not sensitive to global properties of curves. In the figure III-A, the curves γ_1 and γ_2 are at the same distance from γ_3 but have very different shapes. From an operational point of view, γ_1 is just a shifted copy of γ_3 while γ_2 will probably not be realistic.

Since we choose to compare trajectories as mappings, a good candidate for computing the distance will be to integrate over time and set :

$$d(\gamma_1, \gamma_2)^2 = \int_{\mathbb{R}} \|\gamma_1(t) - \gamma_2(t)\|^2 dt$$

as a distance from trajectory γ_1 to γ_2. Since in our formalism γ_1 and γ_2 are only defined on compact intervals of $\mathbb{R}$, say $[a_1, b_1]$ (resp. $[a_2, b_2]$), we assume that we complete the mappings on $\mathbb{R}$ by letting $\gamma_i(t) = \gamma_i(a_i)$ for $t < a_i$ and $\gamma_i(t) = \gamma_i(b_i)$ for $t > b_i$. Unfortunately, unless trajectories to be compared originate from the same point and have the same end point (which would be the case when comparing trajectories on the same origin-destination), distance is infinite as soon as $\gamma_1(a_1) \neq \gamma_2(a_2)$ or $\gamma_1(b_1) \neq \gamma_2(b_2)$. Of course, this is not what we want, so the naive approach of summing square errors must be improved. A very intuitive way of coping with that is to evaluate a mean error instead of the raw sum of squares :

$$d_T(\gamma_1, \gamma_2)^2 = \frac{1}{2T} \int_{-T}^{T} \|\gamma_1(t) - \gamma_2(t)\|^2 dt$$

with $T > 0$. Or, if we allow the mean to be weighted :

$$d_T(\gamma_1, \gamma_2)^2 = \frac{1}{T} \int_{\mathbb{R}} h(t/T) \|\gamma_1(t) - \gamma_2(t)\|^2 dt$$

and h a positive summable function such that :

$$\int_{\mathbb{R}} h(u) du = 1$$

Now, since the derivative of $\|\gamma_1(t) - \gamma_2(t)\|^2$ is :

$$2(\gamma_1(t) - \gamma_2(t), \gamma_1'(t) - \gamma_2'(t))$$

($(.,.)$ denotes the usual scalar product) we have :

$$\|\gamma_1(t) - \gamma_2(t)\|^2 = \Delta^- + \int_{-\infty}^{t} 2(\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) ds$$

with $\Delta^- = \lim_{t \to -\infty} \|\gamma_1(t) - \gamma_2(t)\|^2$ (this limit always exists since we have assumed at the beginning that trajectories are mappings from compact intervals). The expression of the mean square error becomes :

$$d_T(\gamma_1, \gamma_2)^2 = \Delta^- + \frac{1}{T} \int_{\mathbb{R}} h(t/T) \left(\int_{-\infty}^{t} (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) ds \right) dt$$

By fubini theorem, we have :

$$d_T(\gamma_1, \gamma_2)^2 = \Delta^- + \frac{1}{T} \int_{-\infty}^{-T} (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) ds + \int_{-T}^{T} \left(\int_{s}^{T} h(t) (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) dt \right) ds$$

and finally :

$$d(\gamma_1, \gamma_2)^2 = \Delta^- + \int_{-\infty}^{-T} (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) ds + \int_{-T}^{T} (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) (\int_{s/T}^{1} h(u) du) ds$$

Limits as $T \to 0, T > 0$ and $T \to +\infty$ can be obtained. If $T \to O, T > 0$, since $\int_{s/T}^{1} h(u) du$ is bounded, we have :

$$\lim_{T \to O^+} d_T(\gamma_1, \gamma_2)^2 = \Delta^- + \int_{-\infty}^{0} (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) ds$$

This is in fact $\|\gamma_1(0) - \gamma_2(0)\|^2$.

γ_1, γ_2 are constant mappings outside a compact interval of $\mathbb{R}$, so both γ_1' and γ_2' are compactly supported. Furthermore the support of the mapping $s \to (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s))$ is included in the union of the supports of γ_1' and γ_2', thus $A > 0$ exists such that this mapping vanishes outside $[-A, +A]$. This means that for $T > A$ the first integral in the previous expression vanishes. The second has value :

$$\int_{-T}^{T} \left((\int_{s/T}^{1} h(u) du) (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) \right) ds$$

and here again, as soon as $T > A$ this reduced to :

$$\int_{-A}^{A} \left((\int_{s/T}^{1} h(u) du) (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) \right) ds$$

so :

$$\lim_{T \to +\infty} d_T(\gamma_1, \gamma_2)^2 = \Delta^- + \lim_{T \to +\infty} \int_{-A}^{A} \left((\int_{s/T}^{1} h(u) du) (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) \right) ds$$

by dominated convergence theorem, this gives :

$$\lim_{T \to +\infty} d_T(\gamma_1, \gamma_2)^2 = \Delta^- + K \int_{-A}^{A} (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) ds$$

with $K = \int_0^1 h(u) du$. and finally, using again the fact that γ_1', γ_2' are compactly supported, this is :

$$\lim_{T \to +\infty} d_T(\gamma_1, \gamma_2)^2 = \Delta^- + K \int_{\mathbb{R}} (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) ds$$

Now, using the fact that :

$$\int_{\mathbb{R}} (\gamma_1(s) - \gamma_2(s), \gamma_1'(s) - \gamma_2'(s)) ds = \Delta^+ - \Delta^-$$

with $\Delta^+ = \lim_{s \to +\infty} \|\gamma_1(s) - \gamma_2(s)\|^2$, we have :

$$\lim_{T \to +\infty} d_T(\gamma_1, \gamma_2)^2 = (1 - K) \Delta^- + K \Delta^+$$

So the limit case $T \to +\infty$ is a convex combination of the initial and final differences (this will be 0 if trajectories are on the same origin destination, thus the limit of the d_T is not a distance). Letting the weighting function h depend on time shift τ yields the final definition of a family of metrics :

$$d_{T,\tau}(\gamma_1, \gamma_2) = \sup_{\tau \in \mathbb{R}} \frac{1}{T} \int_{\mathbb{R}} h((t - \tau)/T) \|\gamma_1(t) - \gamma_2(t)\|^2 dt$$

with the property that the limit case $T \to O, T > 0$ reduces to the supremum distance. This can be seen as a scale base distance, with T parameter being the scaling factor.

The previous family of distances has nice features because of the scaling ability, but since it is not a single metric, it is difficult to use standard algorithms based on distances (for example, classification algorithms). There is thus a need for another definition of proximity between trajectories that will yield a single value while capturing interesting global characteristics. First of all, we need to go into more detail for the space of trajectories. Since all trajectories are assumed to be constant outside of a compact interval of $\mathbb{R}$ and are smooth, an alternative way of representing them is as a pair (a, γ') with $a = \lim_{t \to -\infty} \gamma(t)$. We have :

$$\gamma(t) = a + \int_{-\infty}^{t} \gamma'(s) ds$$

so that it is equivalent to give γ or the pair (a, γ'). From that, the space of trajectories can be defined as the product $W = \mathbb{R}^3 \times C_c^\infty(\mathbb{R}, \mathbb{R}^3)$ with $C_c^\infty(\mathbb{R}, \mathbb{R}^3)$ the space of smooth compactly supported function from $\mathbb{R}$ to $\mathbb{R}^3$ (this space will be abbreviated by C_c^∞ in the following). C_c^∞ will be endowed with its standard topology that induces the following notion of convergence : a sequence $(\gamma_n)_{n \in \mathbb{N}}$ has limit γ as $n \to +\infty$ if it exists a fixed compact $K \subset \mathbb{R}$ such that the support of all the γ_n are in K and all derivatives of γ_n uniformly converge to those of γ. The space of trajectories W will be given the product topology. From now, all trajectories will be represented as points (a, γ) of W.

An homotopy between smooth trajectories $(a, \gamma_1), (b, \gamma_2)$ is a continuous path $\Phi \colon [0, 1] \to W$ such that :

$$\Phi(0) = (a, \gamma_1),\ \Phi(1) = (b, \gamma_2)$$

An homotopy is said to be smooth if it is smooth as a mapping. Smooth homotopies always exist since :

$$\Phi_0(t, s) = ((1-t)a + tb, (1-t)\gamma_1(s) + t\gamma_2(s))$$

is such an homotopy. Given a trajectory (a, γ), a tangent vector (mapping) with base point (a, γ) will be a 4-uple $(a, \gamma, v, \tilde{\gamma})$ with $\tilde{\gamma}$ an element of C_c^∞, $v \in \mathbb{R}^3$ and such that the mapping $s \to (\gamma(s), \tilde{\gamma(s)})$ is smooth and compactly supported. A picture of that is the following, with $\tilde{\gamma}$ represented as a field of infinitesimal displacements from γ and v a displacement of the origin a :

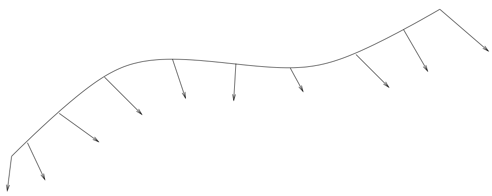


Fig. 2. A trajectory with a tangent vector

As it is easily seen, a smooth homotopy Φ defines for any $t \in [0, 1]$ a tangent vector $(\Phi, \frac{\partial \Phi}{\partial t}(t, .))$.

Let be defined for any trajectory (a, γ) a mapping $F_{(a,\gamma)}$ associating to each tangent vector $(a, \gamma, v, \tilde{\gamma})$ with base point (a, γ) a positive real number and that is homogeneous of degree 2 in $v, \tilde{\gamma}$ (that is, for any $\lambda > 0$, $F_{(a,\gamma)}(\lambda v, \lambda \tilde{\gamma}) = \lambda^2 F_{(a,\gamma)}(v, \tilde{\gamma})$). We further require some kind of smoothness of F, namely that F is at least C^1 on the space $\mathbb{R}^3 \times C_c^\infty(\mathbb{R}, \mathbb{R}^3) \times \mathbb{R}^3 \times C_c^\infty(\mathbb{R}, \mathbb{R}^3)$ of tangent vectors.

Now, define the energy relative to F of an homotopy Φ from γ_1 to γ_2 as :

$$\int_0^1 F_{\Phi(t,.)}(\frac{\partial \Phi}{\partial t}(t, .)) dt$$

A possible choice for $F_{(a,\gamma)}$ is :

$$F_{(a,\gamma)}(v, \tilde{\gamma}) = \|v\|^2 + \int_{\mathbb{R}} \|\tilde{\gamma}(s)\|^2 ds$$

If we take as a smooth homotopy between γ_1, γ_2 the linear homotopy Φ_0, we have :

$$\frac{\partial \Phi_0}{\partial t}(t) = (b - a, \gamma_2 - \gamma_1)$$

and so :

$$E(\Phi_0) = \|b - a\|^2 + \int_{\mathbb{R}} \|\gamma_2(s) - \gamma_1(s)\|^2 ds$$

The interpretation of that energy is easy by reverting to the initial trajectories : $\|b - a\|^2$ is the square distance between the origins of the trajectories, while the integral expresses the square L^2 norms between speeds. If γ_2 is a translation of γ_1 by a vector v, then the energy of the homotopy reduces to $\|v\|^2$. Many other choices for F are possible, yielding different energies. Among them, a very interesting possibility is to introduce a Sobolev energy by taking a differential operator L acting on vectors of $\mathbb{R}^3$ and letting :

$$F_\gamma(v, \tilde{\gamma}) = \|v\|^2 + \int_{\mathbb{R}} \|\tilde{\gamma}\|^2 + \|L\tilde{\gamma}\|^2 ds$$

This will enforce smoothness conditions on the homotopy, since an extra penalty term depending on the regularity of $\tilde{\gamma}$ is added.

Once the energy of an homotopy is defined, a distance on the space W can easily be deduced as :

$$d((a, \gamma_1), (b, \gamma_2))^2 = inf_\Phi E(\Phi)$$

where the infimum is taken over all possible smooth homotopies between γ_1 and γ_2.

B. Trajectories as embeddings

The previous definition of distance assume that trajectories are considered as mappings. In some cases however, the time dependance is irrelevant, as well as the speed along the trajectory. This is for example the case of major flows classification, since only the path followed is of interest. For that purpose, a kind of invariance with respect to time must be introduced. Altough quite different from a conceptual point of view, there is little differences with the previously introduced distances. Trajectories will now be considered as smooth embeddings of the unit interval $[0, 1]$ into $\mathbb{R}^3$, so that self-crossings will not be allowed (this is not really a restriction even if some trajectories failed to fulfill this condition : in all operational cases, self-crossings are not relevant for defining the overall shape of the trajectory and can be safely suppressed). Without going into technicalities, homotopies between embeddings are defined quite the same way as for mappings, but the energy must now be invariant under the action of the increasing diffeomorphisms of $\mathbb{R}$. One way of having such a property is to define F as :

$$F_{(a,\gamma)}(v, \tilde{\gamma}) = \|v\|^2 + \int_{\mathbb{R}} \|\tilde{\gamma}(s)\|^2 \|\gamma(s)\| ds$$

The Sobolev energy can be made invariant using the same trick.

C. Computations

If we except the case of the first family of metrics based on sums of weighted squares, distances between trajectories are obtained as a functional optimization problem (namely, we want to optimize over a set of smooth homotopies). It is generally not possible to find a closed form solution of this problem, so that numerical approximation of the distance is necessary. However, since the original optimization problem is infinite dimensional, it is not directly possible to code it on a computer. For that purpose, some kind of finite dimensional approximation must be done. An investigation of what can be done in order to select a good finite dimensional approximation is the purpose of the second part of the paper, but we will just show now how to use a given finite dimensional expansion for solving the distance problem.

Assume that smooth homotopies are given by a finite expansion :

$$\Phi(t,s) = (\sum_{i=1}^{M} a_i \theta_i(t), \sum_{k=1}^{N} b_k(t)\psi_k(s))$$

with the boundary conditions :

$$\Phi(0,s) = (a, \gamma_1(s)), \Phi(1,s) = (b, \gamma_2(s))$$

and θ_i, ψ_k fixed functions. The derivative with respect to t gives :

$$(\sum_{i=1}^{M} a_i \theta_i'(t), \sum_{k=1}^{N} b_k'(t)\psi_k(s))$$

and the associated energy (for the previous simple case) is :

$$\int_{[0,1]} \| \sum_{i=1}^{M} a_i \theta_i'(t) \|^2 dt +$$
$$\int_{[0,1]} \int_{\mathbb{R}} \| \sum_{k=1}^{N} b_k'(t)\psi_k(s) \|^2$$

This expression can be expanded as :

$$\sum_{i,j=1}^{M} a_i a_j \int_{[0,1]} (\theta_i'(t), \theta_j'(t)) dt +$$
$$\sum_{k,l=1}^{N} b_k(t) b_l(t) \left(\int_{[0,1]} \int_{\mathbb{R}} (\psi_k(s), \psi_l(s)) ds \right) dt$$

Letting :

$$G_{kl} = \int_{[0,1]} \int_{\mathbb{R}} (\psi_k(s), \psi_l(s)) ds$$

and :

$$g_{ij} = \int_{[0,1]} (\theta_i'(t), \theta_j'(t))$$

the energy has the form :

$$\sum_{i,j=1}^{M} a_i a_j g_{ij} +$$
$$\sum_{k,l=1}^{N} G_{kl} \int_{[0,1]} b_k(t) b_l(t) dt$$

This expression is still dependant on functional data, namely the coefficients b_k. Take again a finite expansion as :

$$b_k(t) = \sum_{p=1}^{L} c_{kp} \eta_p(t)$$

with η_p fixed function and putting :

$$H_{p,q} = \int_{[0,1]} \eta_p(t) \eta_q(t) dt$$

the final form of the energy is :

$$\sum_{i,j=1}^{M} a_i a_j g_{ij} +$$
$$\sum_{k,l=1}^{N} G_{kl} \sum_{p,q=1}^{L} c_{kp} c_{lq} H_{p,q}$$

which depends only on a finite set of real numbers. The optimization problem is thus reduced to a quadratic program with constraints, for which ready-made routines exist. The only remaining problem is how to choose the right expansion, and this will be the purpose of the following.

IV. Functional data representation

A. Computing requirements

An ideal functional data processing algorithm should work directly on the functional space in which relevant data lie. However, from a practical point of view, data representation on a computer needs to be finite dimensional, preventing the use of functional algorithm to full extent. The classical way of dealing with this problem is to choose a basis of the functional space [13], then expand functional data on this basis up to a finite number of terms. From the computational point of view, getting the decomposition coefficients must be as fast as possible, and this generally limits the kind of basis functions that may be used. In the hilbert space case, computing decomposition coefficients reduces to evaluate a scalar product, for which fast algorithms exist.

The problem of finding a finite dimensional representation of functional data is not the only one that must be adressed in our case. Since functional data are generally known only by samples at given points, an interpolation problem occurs. Fortunately, it is quite simple to deal with interpolation and decomposition altogether by writing down a least mean square problem.

In the following, we assume that functional data belongs to a (infinite dimensional) real hilbert space H and that $f \in H$ is an application from $\mathbb{R}$ to $\mathbb{R}^p$. For almost any ATM/ATC application, we will have $p = 3$ or $p = 6$ depending on if the speed is relevant or not. It is important to note that it makes a difference for the interpolation problem if speed is taken into account since it doubles the size of the representing space and adds a corresponding number of constraints. Generally speaking, if speed is available as a data, it must be taken into account (this helps producing smooth enough approximation of the true trajectory and improves the accuracy of the finite dimensional approximation) ; otherwise, speed will be a by-product of the interpolation process (like in the case of Kalman filter) but with a lower accuracy.

B. Truncated expansions

Let $(\phi_i)_{i\geq 1}$ be a sequence of linearly independent vectors from the functional space H. Let $V_N = span(\phi_1, \dots \phi_N)$ be the subspace of H spanned by the vectors ϕ_i up to the index N. Any function in V_N admits an expansion of the form :

$$\sum_{i=1}^{N} \lambda_i \phi_i$$

Assuming that functional data f is given as a sequence $(x_i, y_i)_{i=1\dots M}$, so that $f(x_i) = y_i$, we seek for coefficients λ_i^f minimizing the interpolation error :

$$\sum_{i=1}^{N} \|y_i - \sum_{j=1}^{N} \lambda_j^f \phi_j(x_i)\|^2$$

It is well known that the solution of this equation must satisfy :

$$\forall k = 1 \dots N,\ \sum_{i=1}^{M} \langle \phi_k, y_i \rangle = \sum_{j=1}^{N} \lambda_j^f \sum_{i=1}^{M} \langle \phi_k(x_i), \phi_j(x_i) \rangle$$

or, in matrix form :

$$D = G\lambda$$

with :

$$D = (\sum_{i=1}^{M} \langle \phi_j(x_i), y_i \rangle)_{j=1\dots N}$$

$$G = (\sum_{i=1}^{M} \langle \phi_k(x_i), \phi_j(x_i) \rangle)_{k,j=1\dots N}$$

It is important to note that increasing the number of samples for the function f (i.e. increasing M) does not change the dimension of the matrix G. Furthermore, it is possible to compute incrementaly G by using the Sherman-Woodbury-Morisson formula relating the inverse of a sum of matrices to inverses of matrices themselves [5] :

$$(A + UV^t)^{-1} = A^{-1} - A^{-1}U(Id + V^t A^{-1} U)^{-1} V^t A^{-1}$$

with A a $n \times n$ matrix, U, V $n \times k$ matrices (of course one needs to assume that A and $(Id + V^t A^{-1} U)$ are non singular for the formula to be valid). Let us denote by G_M the matrix obtained with M samples. We have :

$$G_{M+1} = G_M + \tilde{G}$$

with :

$$\tilde{G} = (\langle \phi_i(x_{M+1}), \phi_j(x_{M+1}) \rangle)_{i,j=1\dots N}$$

Since $\tilde{G}$ is a Gram matrix, it admits a cholesky decomposition $\tilde{G} = LL^t$. Applying the Sherman-Woodbury-Morrison formula yields the following expression :

$$G_{M+1}^{-1} = G_M^{-1} - G_M^{-1} L (Id + L^t G_M^{-1} L)^{-1} L^t G_M^{-1} \quad (1)$$

from which we can derive :

$$L^t G_{M+1}^{-1} L = (Id + L^t G_N^{-1} L)^{-1} L^t G_M^{-1} L$$

and finally :

$$G_M^{-1} L = G_{M+1}^{-1} L (Id - L^t G_{M+1}^{-1} L)^{-1}$$

Substituting this into 1 gives :

$$G_M^{-1} = G_{M+1}^{-1} + G_{M+1}^{-1} L (Id - L^t G_M^{-1} L)^{-1} L^t G_{M+1}^{-1}$$

On the other hand, incremental formulas when a new function ϕ_{N+1} is added can be otained by using the well known expression for the inverse of a partitionned matrix. If matrix G is written as :

$$G = \begin{pmatrix} G_1 & G_2 \\ G_3 & G_4 \end{pmatrix}$$

the inverse of G can be expressed as :

$$G^{-1} = \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix}$$

with :

$$Q_1 = (G_1 - G_2 G_4^{-1} G_3)^{-1}$$
$$Q_2 = G_1^{-1} G_2 (G_3 G_1^{-1} G_2 - G_4)^{-1}$$
$$Q_3 = (G_3 G_1^{-1} G_2 - G_4)^{-1} G_2 G_1^{-1}$$
$$Q_4 = (G_4 - G_3 G_1^{-1} G_2)^{-1}$$

Truncated expansions are thus easy to use from a computational point of view and allow for incremental update either when the number of samples changes or when the number of subspace basis functions changes. An exposition of truncated expansion for functional data processing in the case of neural networks can be found in [15], [14]

C. Working with truncated expansions

Distance on the subspace spanned by the $(\phi_i)_{i=1\dots p}$ is given by the ambient distance in H. However, if we compute distance between two sampled trajectories c_1, c_2 we find the classical euclidean distance in $\mathbb{R}^p$. The discrepancy between the two may be computed since :

$$\|c_1 - c_2\|_H^2 = \sum_{i,j=1}^{p} (c_{1,i} - c_{2,i})(c_{1,j} - c_{2,j}) \langle \phi_i, \phi_j \rangle$$

or in a matrix form :

$$\|c_1 - c_2\|_h^2 = (c_1 - c_2)^t G (c_1 - c_2)$$

with :

$$G = (\langle \phi_i, \phi_j \rangle)_{i,j=1\dots p}$$

a Gram matrix. As previously quoted, G admits a Cholesky factorization $G = LL^t$, from which can be deduced :

$$\|c_1 - c_2\|_h^2 = \|L(c_1 - c_2)\|_{\mathbb{R}^p}^2$$

So that distance in the subspace is the euclidean distance for coefficients vectors scaled by the matrix L. Equality between the two distances occurs only when basis $(\phi_i)_{i=1\ldots p}$ is orthonormal.

In order to avoid scaling troubles, specially in the case of an ill conditioned gram matrix, it is better to work directly with scaled coefficients.

Scaling matrix is subject to efficient update in case of new basis function addition. Assuming that ϕ_{p+1} is linearly independant from $(\phi_i)_{i=1\ldots p}$, the cholesky decomposition of the new Gram matrix $\tilde{G} = \tilde{L}\tilde{L}^t$ can be expressed as :

$$\tilde{L} = \begin{pmatrix} L & 0 \\ (L^{-1}u)^t & \sqrt{\|\phi_{p+1}\|^2 - u^t G^{-1} u} \end{pmatrix}$$

with $\tilde{G}$ written in block form :

$$\tilde{G} = \begin{pmatrix} G & u \\ u^t & \|\phi_{p+1}\|^2 \end{pmatrix}$$

This update form is in fact a rewriting of the Gram-Schmidt orthonormalisation procedure.

D. Finite subspaces selection

Working with a single approximation subspace amounts to no more than discretizing the initial functional problem. For the distance computation application, this will reduce the initial variation calculus problem to a finite dimensional minimum seeking problem. In many cases however, the optimal subspace for representing data with a fixed number of basis functions is not known. A standard procedure for choosing a suitable basis relies on the leave-one-out score.

Assuming that we want to compare two truncated expansions :

$$\sum_{i=1}^{p} a_i \phi_i(x)$$

and

$$\sum_{j=1}^{q} b_j \psi_j(x)$$

of the sampled functional data $(x_k, y_k)_{k=1\ldots N}$, a crude procedure will be to simply compute interpolation errors :

$$\sum_{k=1\ldots N} \|y_k - \sum_{i=1}^{p} a_i \phi_i(x_k)\|^2$$

and

$$\sum_{k=1\ldots N} \|y_k - \sum_{i=1}^{p} b_i \psi_i(x_k)\|^2$$

This will not give the desired result since basis with a high number of basis functions will be selected, yielding most of the time over-fitted models. Since we must insure robustness for the solution of the interpolation problem, it make sense to compare on the leave-one-out score instead of interpolation error. This value is computed for the subspace $span(\phi_1, \ldots \phi_p)$ as follows :

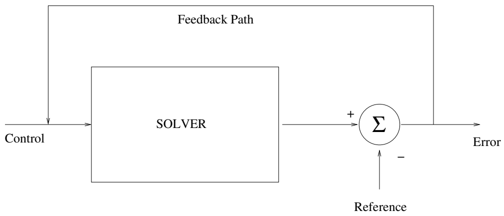


Fig. 3. General simuation model

- For each $l = 1 \ldots N$, compute the coefficients (a_{il}) minimizing the interpolation error :

$$\sum_{k=1\ldots N, k \neq l} \|y_k - \sum_{i=1}^{p} a_{il} \phi_i(x_k)\|^2$$

  . Let e_l be the value of the minimal interpolation error.
- Evaluate the mean value of the interpolation errors to get the leave-one-out score :

$$LOOS(\phi_1, \ldots \phi_p) = \frac{1}{N} \sum_{l=1\ldots N} e_l$$

In the case of multiple interpolation data (that is several sequences (x_i, y_i) describing samples of functional data), the overall $LOOS$ will be obtained by summing up all individual $LOOS$. Subspaces with lowest $LOOS$ will be selected.

Computation of $LOOS$ by brute force techniques is expensive since N different least mean square problems must be solved in order to get the leave-one-out score for a single functional sample. However, recalling that there is an incremental way of computing the inverse of the Gram matrix involved in the solution of the problem, a computationally efficient algorithm may be designed.

V. TRAJECTORY DISTANCE ATM APPLICATIONS

This section presents a list of ATM applications for which an exact trajectory distance metric is needed.

A. Aircraft model Inference

All aircraft models are based on ODEs (Ordinary Differental Equation), including tabular ones V-A.

Control input includes condition and model parameters. The model refinement (and computational complexity) range from tabular to many degrees of freedom. The aircraft model inference consists in answering the following question :**Given a parametrized model and a goal trajectory, can we infer the best parameter values?**

A model can be viewed as a mapping from the control space into the trajectory space. The way to answer the previous question is then given by the closest model to the goal trajectory (see figure V-A).

In order to find the closest model in this trajectory space, a reliable trajectory distance is needed. The model inference

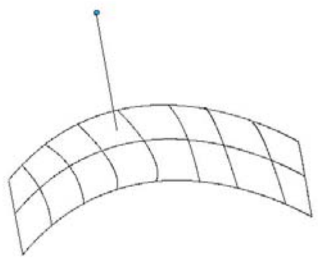


Fig. 4. Finding best model from a given class

problem has to solve the accuracy-smoothness dilemma :**Over-fitted models are generally poor predictors**.

The previous construction gives the shortest path (and thus the distance) between the goal trajectory and the trajectory set which can be synthesized by the model. All deformations from the goal trajectory to the trajectory set are taken into account but many of those transformations are not feasible from a physical point of view. The solution consists in defining a smooth path between trajectories by an infinitesimal displacement field and by adding a field smoothness requirement to the energy functional. On the other hand, if we decide to control the maximum distance between the goal trajectory and a given parametrized model, it is possible to adjust the most economical model to the curent phase of flight. As a matter of fact, computation effort is proportional to the degrees of freedom used and to the evolution equation used. Then, the relevant question is the following :**Can we switch from different level of accuracy depending on flight phases and computational power available?**. The solution consists in switching from one model of given accuracy to another by following a shortest path in the trajectory space.

B. Trajectory prediction

Air traffic management research and development has provided a substantial collection of decision support tools that provide automated conflict detection and resolution [3], [1], [23],trial planning [9], controller advisories for metering and sequencing [21], [2], traffic load forecasting [10], [8], weather impact assessment [7], [20], [4]. The ability to properly forecast future aircraft trajectories is central in many of those decision support tools. As a result, trajectory prediction (TP) and the treatment of trajectory prediction uncertainty continue as active areas of research and development (eg [18], [22], [11], [16], [19].

Accuracy of TP is generally defined as point spatial accuracy (goal attainment) or as trajectory following accuracy. The last one can be rigorously defined by the mean of trajectory space. The first one is a limit case of the second by adding a weight function in the energy functional. Since we may prescibe smoothness accuracy of a simplified model relative to a finer one, may be computed. Finally such trajectory distance allows comparison between TP models.

When we refer to trajectory prediction errors for a specific DST, we are typically comparing the predicted trajectory for a specific DST to the actual trajectory to be experienced by an aircraft. Discrepancies between these two types of trajectories typically affect the performance of the DST.

C. Radar tracker evaluation

The goal of a radar tracker is to eliminate the residual noise coming from the radars. It is a key element of the ATM system and its accuracy is one of the factors which determines the separation norm. In order to validate such trackers, an exact reference trajectory is generated and perturbated by a white gaussian noise. This perturbated trajectory is then used as input of the tested tracker. The tracker generates an estimated trajectory which is compared to the reference trajectory. In order to do such comparison, a reliable trajectory distance is needed.

D. Alternative route synthesis

Airspace congestion is related to aircraft located in the same area during the same period of time. Then, when congestion has to be minimized, algorithms have to separate aircraft in time (slot allocation), in space (route allocation) or both (bi-allocation). When route allocation is investigated, associated algorithms need alternative routes set in order to spread the traffic on them. A route is said to be alternative to another if it is different enough based on a trajectory distance.

E. Major flows definition

When radar tracks are oberved over a long period of time in a dense area, it is very easy to identify major flows connecting major airports. The expression ”major flows” is often used but never rigorously defined. Based on an exact trajectory distance and a learning classifier, it is possible to answer the following questions :**Given a set of observed trajectories, can we spit it into ”similar” trajectory classes? If yes, classes with highest number of elements will rigourously define the major flows. Given those classes and a new trajectory, can we tell if it belongs to a major flow and which one?** The principle of the major flows definition is to use shape space to represent trajectory shapes as points and to use a shape distance. (the shape of a trajectory is the path followed by an aircraft, that is the projection in the 3D space of its 4D trajectory. The speed on the path has no impact).

VI. Conclusion

This paper has presented several ways of computing distances on the space of trajectories and has introduced concepts originating from functional analysis in order to work directly on trajectories as a whole. The first family of metrics, scale based, is mainly useful for descriptive purpose and to quickly analyse a set of trajectories (for example, as a tool complementary to standard descriptive statistics). The scale dependence bears some similarity with continuous wavelets analysis of signal, and in fact if the weighting function is kept constant, then the resulting family of metrics is the continuous wavelet transform of the point distances along tranjectories. The limit as scale goes to $O, T > 0$ is the supremum distance, while the limit as scale goes to infinity depends only on the distances between origins and destinations (and is thus not a true distance).

For more in depth analysis of trajectories, another kind of distance has been introduced that is based on the infimum of the energies of homotopies joining pairs of trajectories. This yield to a variational problem that cannot be solved directly, but may be reduced to a quadratic optimization problem with constraints by expanding on a finite family of functions. This kind of distance allows computations to be done on trajectories understood as mappings, that is with time dependence, or on trajectories understood as shapes (or embeddings).

The last part of the paper has studied the finite dimensional representation of functional data so that finite expansions used in computing distances are optimaly chosen.

REFERENCES

[1] D. Brudnicki and A. McFarland. User request evaluation tool (uret) conflict probe performance and benefits assessment. In *FAA/Eurocontrol ATM Seminar Saclay France*, 1997.

[2] R. Coppenbarger, R. Lanier, D. Sweet, and S. Dorsky. Design and development of the enroute descent advisor (eda) for conflict-free arrival metering. In *AIAA-2004-4875 AIAA GNC Conference Providence RI*, 2004.

[3] H. Erzberger, R. Paielli, D. Isaacson, and M. Eshowl. Conflict detection in the presence of prediction error. In *FAA/Eurocontrol ATM Seminar Saclay France*, 1997.

[4] J. Evans and al. Reducing severe weather delays in congested airspace with weather support for tactical air traffic management. In *FAA/Eurocontrol ATM Seminar BUdapest Hungary*, 2003.

[5] G. Golub and C. V. Loan. *Matrix Comoutations*. The John Hopkins University Press, 1989.

[6] C. Jaynes, S. Webb, R. Steele, and Q. Xiong. n open development environment for evaluation of video surveillance system. In *Proceedings of PETS, Copenhagen*, June 2002.

[7] D. Kirk and al. Problem analysis resolution and ranking (parr) development and assessment. In *FAA/Eurocontrol ATM Seminar Santa Fe NM*, 2004.

[8] A. Masalonis and al. Using probabilistic demand prediction for traffic flow management decision support. In *AIAA-2004-5231 AIAA GNC Conference Providence RI*, 2004.

[9] B. McNally, R. Bach, and W. Chan. Field test evaluation of the ctas conflict prediction and trial planning capability. In *AIAA-1998-4480 AIAA GNC Conference Boston MA*, 1998.

[10] C. Meckiff, R. Chone, and J. Nicolaon. The tactical load smoother for multi-sector planning. In *FAA/Eurocontrol ATM Seminar Orlando FL*, 1998.

[11] S. Mondoloni, S. Paglone, and S. Green. rajectory modeling accuracy for air traffic management decision support tools. In *ICAS Congress Toronto ON*, 2002.

[12] C. Needham and R. Boyle. Performance evaluation metrics and statistics for positional tracker evaluation. In *Proceedings of ICVS, Graz, Austria*, pages 278–289, April 2003.

[13] J. Ramsay and B. Silverman. *Functional Data Analysis*. Springer-Verlag, 1997.

[14] F. Rossi and B. Conan-Guez. Functional multilayer perceptrons : a nonlinear tool for functional data analysis. *Neural Network*, 2004.

[15] F. Rossi, B. Conan-Guez, and F. Fleuret. Functional data analysis with multilayer perceptrons. In *Proceedings of IJCNN 2002*, pages 2843–2848. IEEE/NNS/INNS, 2002.

[16] H. Ryan, M. Paglione, and S. Green. Review of trajectory accuracy methodology and comparison of error measure metrics. In *AIAA-2004-4787 AIAA GNC Conference Providence RI*, 2004.

[17] A. Senior, A. Hampapur, Y. Tian, L. Brown, S. Pankanti, and R. Bolle. Appearance models for occlusion handling. In *Proceedings of PETS, Hawai, Kauai*, December 2001.

[18] R. Slattery and Y. Zhao. Trajectory synthesis for air traffic automation. In *AIAA Journal Guidance Control and Dynamics*, volume 20, pages 232–238, 1997.

[19] S.Mondoloni and I. Bayraktutar. mpact of factors, conditions and metrics on trajectory prediction accuracy.

[20] V. Sud and al. Air traffic flow management collaborative routing coordination tools. In *AIAA-2001-4112 AIAA GNC Conference Montreal PQ*, 2001.

[21] H. Swensen and al. Design and operational evaluation of the traffic management advisor at the forth worth air route traffic control center. In *FAA/Eurocontrol ATM Seminar Saclay France*, 1997.

[22] S. Swierstra and S. Green. Common trajectory prediction capability for decision support tools. In *FAA/Eurocontrol ATM Seminar BUdapest Hungary*, 2003.

[23] A. Vink. Eatchip medium term conflict detection: Part 1 eatchip context. In *FAA/Eurocontrol ATM Seminar Saclay France*, 1997.

Possible Model to Highlight Interdependencies between Flow Regulations

Brankica Pešić Le Foll, Claus Gwiggner

Abstract— **This paper addresses the problem of detecting interdependencies between flow management measures called regulations. The phenomenon is known as *network effect*. Our approach is based on the analysis of data obtained from systematic simulations of air traffic. We use a tree-based method to highlight interactions between regulations.**

We compare different indicators for dependencies in order to obtain results that can be generalized. We demonstrate the approach on a hypothetical example where we identify expected dependencies. We also identify a dependency that is not expected. This is an exploratory study showing an alternative to purely modeling/simulation based approaches to the network effect.

Index Terms—**air traffic flow management, network effect, classification trees, data models**

I. INTRODUCTION

A. General Background

FORECASTS of air traffic growth are not likely to resolve already congested airspace in Europe. In order to continue to balance traffic demand with the available capacity, the air traffic community is searching for a better understanding of the mechanisms that govern the airspace system. Based on this, new organisation of airspace and of traffic flows, new methodologies in control or new technology can be proposed.

B. Some definitions

In order to understand better the problem we focus on, the following definitions are given:

--Regulation: A regulation is a flow management measure defined in order to regulate traffic demand in case of congestion. It can be applied on reference locations (sector, airport, etc.) It is defined by a flow rate per time period. A flow rate is the maximal number of aircraft that can enter a reference location.

Manuscript received February 25, 2006. This research was supported and sponsored by the EUROCONTROL Experimental Center. The work is done under supervision of M Serge MANCHON.

B. Pešić Le Foll, is student at the Faculty of Transport and Traffic Engineering, University of Belgrade, Serbia and Montenegro (e-mail: brankica.le-foll@lwanadoo.fr).

C. Gwiggner is PhD student with the Polytechnique school, France (e-mail: claus-peter.gwiggner@eurocontrol.int)

--Regulation plan: A regulation plan is a set of regulations that aims at protecting the entire European airspace during one day. A regulation plan also includes other information not relevant for our study.

--Network effect: Regulations might have overlapping effects since they are linked by traffic flows. Thus, a regulation in one sector may have an impact on the number of aircraft that wish to enter another, not necessarily neighbouring sector. This propagation is called *network effect*.

--**Generated delay**: The slot allocation procedure works out ground delays to flights in order to satisfy the regulations. The sum of ground delays due to a given regulation is called its *generated delay*.

The aim of ATFCM (Air Traffic Flow and Capacity Management) is to balance demand and available capacity of airspace.

Improving ATFCM can be achieved through:

--a better alignment of the ATM (Air Traffic Management) capacity towards the traffic demand,

--by improving efficiency of the ATFCM process in the balancing of demand and capacity,

--by developing network management which will consider the overall impact of restrictions and measures to utilise available capacity.

Network management is one of the major concepts of ATFCM [1]. Thus, the need to predict the impact of those solutions to resolve capacity shortfalls could all have upon elements of the system (airports and sectors) is obvious.

The paper is organized as follows. The problem formulation, related work and our approach are given in the next part. Classification tree analysis is presented in the third part. In the fourth part we give the data description and result of the classification tree application. Conclusion and the future work are in the last part.

II. PROBLEM FORMULATION

The subject of this research is to find dependencies between regulations. For a given day of traffic, regulation plan and corresponding generated delay, we would like to find interactions between regulations applied.

Knowledge about such interactions could be useful during all ATFCM phases, from the strategic and pre-

tactical ATFCM phases in order to avoid situations that generate important delay (demand/capacity imbalance), but also, during the tactical phase when changes in regulation rate have to be decided with short notice.
The effect of a regulation can be measured with the generated delay (performance of the system). But beside the sector where applied, the regulation can influence the generated delay of neighbouring sectors. By consequence, it can have an impact on geographically far lying sectors that are linked by the traffic.

This is a difficult problem to formulate analytically.

A. *Related Work*

Simulation studies are done in: MITRE/CAASD - CRCT[2] tool; and in Boeing - the Boeing National Flow Model. The network effect is quantified using event-based simulations. In those studies, inter-dependences are not analysed.

The NEVAC[3] tool offers the possibility to analyse generated delay and to trace dependencies but, as to our knowledge, there are no published results. To sum up, analysis and modelling of the network effect is still an open question.

An alternative to simulation is to use descriptive methods in answering: "Where are the interactions?" This is a first step in the analysis of the problem.

In the following we shall give our approach to answer this question.

B. *Proposed approach*

We formulate the problem of finding interactions as a classification problem. Different classes of delay will be defined (e.g. low, high, etc) and rules will be identified to describe these classes. Our approach is entirely based on the analysis of data obtained from traffic simulation.

We use a method that derives from earlier methodology called "automatic interaction detection"; t*ree-based method*[4]. The application of this method is convenient for two types of problems, classification problems and regression ones. The result of this method is a tree which, in our case, is a classifier for the indicator delay. Analysing the tree, it is possible to identify interactions between regulations. In this research study we used a popular implementation for the tree-based regression and classification called CART [4].

III. CLASSIFICATION TREE ANALYSIS

Tree based methods belong to the class of 'supervised learning' methods in data-mining [2], [5].

The basic idea of tree based methods is to partition the input space into regions that correspond well to the classes. In classification, CART does this with the following recursive principle: split the data into the most homogeneous groups until all classes are identified or the size of a group is too small to further split. The homogeneity of groups is measured by an 'impurity index', and the splitting rule is called 'binary recursive splitting'. The result of the algorithm is a tree with low impurity, that means with as homogenous groups as possible. Following paths from the root until a leaf of the tree gives a rule which predicts the class of the indicator (please see below for examples). The validation of a tree is based on the expected prediction error. Trees with low errors are searched. Classical hypothesis tests are not available for this technique. Two details are:

- finding the least impurity is a difficult combinatorial optimization problem. The CART strategy is a heuristic to find this optimum because it optimizes the impurity function locally, that is, at every splitting point of the tree. Its result is in general a local optimum.
- experience has shown that tree building based on the above principle alone leads to unstable results. The common strategy to improve this is by pruning the tree to a more stable size. This is done by cross validation, Hastie et al. in [5].

The key advantages of the tree-based methods are in tree interpretability and in the possibility to find nonlinear and complex interactions. Drawbacks of this method are a high variance, local optimality and the non availability of formal inference techniques.

This implies that results from this analysis alone are not recommended for decision making. This study is exploratory and results should be critically examined and validated with complementary techniques.

IV. RESULTS

A. *Data description*

We created a simple example which permits to demonstrate the usefulness of our approach. We have chosen three regulations to apply in the ACC Reims in France (sector LFEUN and collapsed sectors LFEESE and LFEUEXE, see Fig. 1). The indices of each regulation are presented in Table I.

TABLE 1
INDICES OF REGULATIONS

Indice	R1	R2	R3
Regulation	LFEUEXE	LFEUN	LFEESE

Knowing the traffic flow orientation, it can be expected that interaction between regulations LFEESE and LFEUEXE exist, but not between the others (figure 1) .

The input data for the CART application are results of systematic simulations in COOSAC (Common Simulator to

[2] Collaborative Routing Coordination Tool
[3] Network Evaluation Validation & Analysis of Capacities, simulation tool developed by Network Capacity & Demand Management in EUROCONTROL
[4] for more details about tree-based method we refer to [2] and [3]

Assess ATM Concepts) tool.

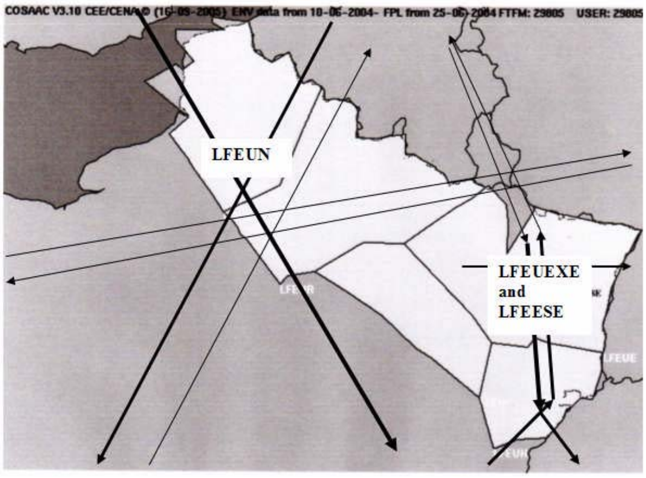


Fig. 1. ACC Reims: sectors where regulations are applied and a traffic flow orientation

One simulation gives results for one day of traffic and one combination of flow rates for the three regulations applied. As a range for the flow rates we choose a 10% interval around the initial flow rate. This results in seven possible values for each of the three regulations. That means 343 (7^3) simulations were done in order to cover all possible combinations of flow rates for the three regulations applied for one day.

Simulations are done for real traffic data from eight Fridays (11th June, 18th June, 25th June, 2nd July, 9th , 16th ,23rd and 30th July 2004), resulting in an input data set with 2744 cases. Figure 2 shows the elements of the analysis.

CART is implemented in different statistical tools. In this study we use the language R [6] and [7].

B. Classification tree construction and analysis

As an indicator for interaction detection, the generated delay is used. The classification tree aims in explaining the indicator delay D with different regulation rates R_i. It permits to detect a functional relation, even non-linear. This is a typical classification problem where D is the dependent and the R_i are the independent variables.

We test three indicators. First, we select the indicator 'absolute delay'. The result is a classification tree dependent on a specific day. Thus, generalization is not possible. We have the same situation for the indicator 'relative delay'. We can conclude that the indicators absolute and relative delay are not efficient in highlighting interactions. Therefore we test a third indicator – 'delta'. The delta presents difference in delay. In order to calculate this difference, we define a baseline scenario. This scenario consists in the application of three regulations with flow rates that correspond to the mean values for each regulation. The baseline delay presents a generated delay for a given baseline scenario. We chose to define the indicator delta with three classes $\{-1, 0, 1\}$ as in (1). The interval $[a,b]$ is defined in order to test two definition for the class 0.

$$Delta = \begin{cases} -1, \ if \left(Generated\ delay - Baseline\ delay\right) < a \\ 1, \ if \left(Generated\ delay - Baseline\ delay\right) > b \\ 0, if \left(Generated\ delay - Baseline\ delay\right) \in \left[a, b\right] \end{cases} \quad (1)$$

The delta belongs to class -1 if we have a delay reduction comparing to the baseline delay. The delta belongs to the class 1 if the generated delay is greater than the baseline delay.

Two tested cases for class 0 are the following ones:

First case: we define class 0 with a single value 0; class 0 is attributed to the delta only in the case when the generated delay is equal to the baseline delay. Thus the boundaries of the interval $[a,b]$ are a=0 and b=0.

Second case: we define class 0 with an interval; class 0 is assigned to the difference in delay between -200 and 500 minutes comparing to the baseline delay. For this case we have following boundaries: a= - 200 and b=500. Those values for a and b divide the input data set in three sets with approximately the same number of cases in each.

A classification tree for this case is given on Fig. 3. It is constructed for the regulation in LFEESE which is coded by R3 (see Table I).

We can identify two things: firstly, a relation between the regulations R3 and R1 exists. Secondly, the relation is not dependent on a specific day, because the variable 'day' is not used in tree construction. For analyzing the tree we proceed as follows.

Starting from the root, we obtain for example:

$$\begin{aligned} & if\ (30.5 \le R3 < 31.5)\ and\ \ if\ (R1 < 35.5) \\ & then\ \ CLASS = -1 \end{aligned} \quad (2)$$

This classification rule is presented with a bold line in Fig. 3. The rule (2) can be interpreted as follows: the delta is in the class -1, meaning that delay is reduced for more than 200 minutes comparing to the baseline. This also means that the intensity of interaction between R1 and R3 is important. Highlighted case is if the regulation R1 flow rate is less than 35 aircraft per hour and flow rate 31 for regulation R3.

Classification trees for the two other sectors, LFEUN and LFEUEXE, are also constructed. As expected we do not find that there is interaction between regulation applied in LFEUN and LFEESE. We also find that the interaction between regulation in LFEESE and LFEUEXE is not symmetric, meaning that LFEUEXE has more influence to the system performance in LFEESE than in the opposite case.

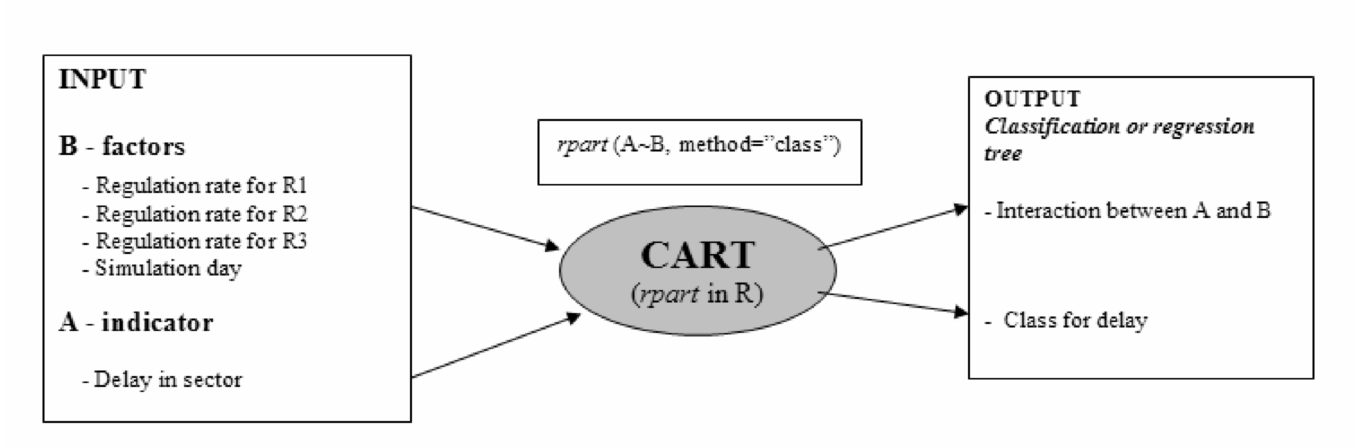


Fig. 2. Application of the tree based method

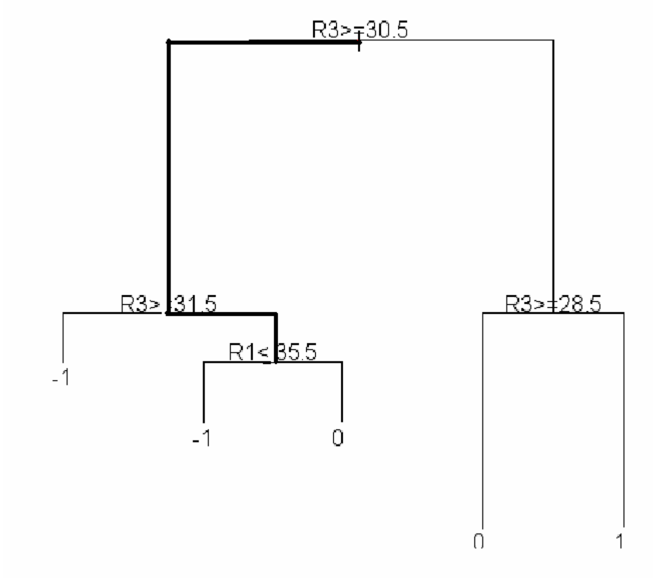


Fig. 3 Classification tree for LFEESE where difference in delay is coded by 0 if difference in delay is superior of -200 minutes and inferior of 500 minutes. Interaction between regulations LFEESE (R3) and LFEUEXE (R1)

C. Discussion

In total, the application of tree-based methods shows that all of the expected interactions are identified.

It has to be said that an unexpected interaction is detected too. This has to be analyzed in more detail. Either there is an indirect interaction or the indicator still has to be improved.

This is an example of the limitations of the study, which is exploratory. We use a descriptive method without a specific hypothesis or model in mind to analyze a new, rather complex dataset. There are no formal hypothesis tests to do at this stage. We use the trees as descriptions of the data which lead, when critically used, to more precise, substantial research hypotheses.

V. CONCLUSION AND FUTUR WORK

The network effect is a complex phenomenon for which a mathematical formulation is difficult to give. This study is an exploratory one in order to propose a new approach in network effect modelling. We use a descriptive method, a tree-based method to detect simple, but non-linear functional relations between regulations.

We compare several indicators with the aim of producing results that can be generalized.

This study is showing that tree-based methods permit to detect interactions between flow management measures. But it has to be highlighted that it is too early to use classification tree in decision making.

In order to confirm the findings, these interactions should be analysed with classical classification methods like linear and non-linear logistic regression. Also, stability tests would have to be done if further operational use were envisaged. To reduce instability of trees, several ways are already developed [5].

A main objective of this modelling is of course to build a tree for an ECAC wide scenario which models the network effect in ECAC. A future work should consist in improving a simulation scenario. This means to create trees for different periods of the day where regulation plans with 80-140 regulations are taken into account. Here, exhaustive simulation is no more feasible; thus a sampling strategy has to be developed.

One important question remains open: Which other indicators might be relevant in interaction detection?

REFERENCES

[1] EUROCONTROL, 2004, Air Traffic Flow and Capacity Management, Evolution Plan for the ECAC States, Edition1.0
[2] D. Hand, H. Mannila, P. Smyth, *Principles of Data Mining*, Massachusetts Institute of Technology, 2001.
[3] J. Maindonald, J. Braun, *Data Analysis and Graphics Using R - An Example-Based Approach*, Cambridge University Press, 2003, ISBN 0521813360
[4] L. Breiman, J. Friedman, Ohlsen, Ch. Stone, *Classification and Regression Trees, Chapman & Hall/CRC*, 1984, ISBN 0412048418.
[5] Hastie, Tibshirani, Friedman, *The Elements of Statistical Learning*, Springer, 2003, ISBN 038795284.
[6] W. N. Venables, D. M. Smith and the R Development Core Team, *An Introduction to R*, 2002, Version 1.5.0
[7] T. M. Therneau, B. Atkinson, R Development Core Team, *The rpart Package*, 2005, Version 3.1-23

Generalised Intrinsic Characteristics as a Forecasting Tool: A dynamic perspective

Radosav B. Jovanović

Abstract— **The trip generation model presented combines sequentially cross-sectional and time-series data analyses, utilising only basic statistical tools. It exemplifies one possible way of tackling one of the frequent forecasting-related issues: restricted sample size – either spatially (small number of observational units) or/and longitudinally (insufficiently long time-series). The model developed makes use of the observed autocorrelation of the regression residuals over time, and by including a dynamic element into the formerly known theory of intrinsic characteristics, takes account of the gross effect of extraneous factors on air passenger traffic volume. The modification applied significantly improves the forecasting accuracy of the original cross-sectional model.**

Index Terms—**Autocorrelation, cross-sectional analysis, intrinsic characteristics, regression analysis, time-series, trip generation model.**

I. Introduction

Modelling transport, in general, is a context-dependent process, and no universal technique exists that might be successfully applied to every problem, it is rather that every single model should be tailored to its particular context. Naturally, theoretical soundness of the model proposed is a necessary condition, still, one should bear in mind that accuracy is only one aspect of forecasting, and that cost, ease of use, utility of output and ease of interpretation are almost as important [1], [2]. Therefore, balance ought to be sought between theoretical consistency and expedience in the forecasting process [3]. In many cases, relatively simple (inexpensive and readily applicable) modelling techniques may provide satisfactory forecasts with respect to most of the criteria previously mentioned.

This paper presents the results of such an application, blending two related, but so far separate concepts — to obtain an air trip generation forecast for a set of European countries. The first concept is purely statistical - autocorrelation of the error terms in a time-series regression, whereas the second, a theory of intrinsic characteristics, but with a dynamic element incorporated, has a reasonable practical interpretation. Cross-sectional and time-series approaches were combined, with only basic statistical tools utilised for parameter estimation.

The trip generation model developed is inevitably concerned with the rest of the modelling process, being an integral part, thus a brief overview of the broader modelling context is thought to be helpful in recognizing the framework in which the model has been built.

Manuscript received February 26, 2006.

Radosav B. Jovanović is a part-time research assistant at the Air Transport Department, Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade, Serbia and Montenegro (phone: +381 64 1256 829, +381 11 3091 352; fax: +381 11 2496 476; e-mail: rgrozni@eunet.yu).

II. Context

The model described in this paper has been developed as a part of the *"Serbia and Montenegro (SCG) Air Traffic Analysis and Forecast"* study. The study ordered by the Serbia and Montenegro Air Traffic Services Agency (SMATSA) was undertaken by the Faculty of Transport and Traffic Engineering, University of Belgrade, and supported by EUROCONTROL, in the period December 2004 – June 2005. Its objective was the forecast of annual and peak day SCG air traffic volume for years 2007, 2010 and 2015 [4].

A classical sequential four-phase transportation demand model structure has been used for that purpose. Due to the absence of serious intermodal competition in the area of interest a mode-specific approach has been employed. The trip generation model was used to estimate the air trip production of a set of European zones identified to be the principal generators of the SCG transit air traffic (Figure 1). Its output has served as one of the inputs for the trip distribution model, the output of which represented itself one of the inputs for the final, route-choice phase.

Four scenarios concerning the design years' levels of explanatory variables in the route-choice phase were provided by the study customer, thus any branching in earlier stages would have led to an impracticably large number of variants at the very end. Therefore, the point forecast of the trip production level of each of the zones was needed.

A. Data Availability

It is well known that data availability and modelling approach constrain and condition each other [5]. Whereas it is sometimes difficult to decide on the predominant direction of this two-way relationship, in our case the model structure was clearly heavily influenced by the data

Fig. 1. Main air traffic flows crossing SCG airspace
Source: [4]

availability. Apart from the problems of obtaining a long enough historical data series for each of the zones considered, as a ten-year forecast of the zonal passenger traffic was the required model output, the equivalent length of a credible forecast of any would-be explanatory variable was subsequently needed too. This turned out to be a highly restrictive prerequisite, with only a few variables having the desired properties.

Historical data on the volume of international air passenger traffic by reporting country were available for the 1995-2003 period. Data on GDP levels, a few tourism indicators and on the volume of external trade by country were collected. The importance of the air fare level as a measure of travel impedance was recognized too, however, both lack of systematic historical data on ticket prices, and non-existence of reliable forecasts on their future level made the inclusion of this factor impracticable.

B. *Variable Choice*

Having conducted preliminary analyses based on available data sets, being in principle inclined towards developing a causal rather than a univariate model, and bearing in mind the required form of the model output, it was finally decided to model the *zonal volume of international air passenger traffic production* (IP) using zonal GDP as the only explanatory variable. It should be mentioned that the other exogenous variables tested have also showed statistical significance with respect to the zonal traffic production, but were also highly correlated with the GDP level of the zones considered, which posed a problem of multicollinearity.

A literature review also showed that GDP is most often considered a main driver for passenger growth [6] - [8]. Moreover, such an approach is in line with the practice followed by ICAO [9] and EUROSTAT [10], which correlate passengers to GDP only.

III. Trip Generation Model Concept

It should be stressed at the very beginning of this section, that no strict trip generation model concept had been set in advance of the data analysis process. The only point fixed was the goal of the analysis, that is, the required form of the model output, thus freedom existed in that sense to make different methodological choices at different points of the process, in order to approach the desired goal as close as possible. To keep some cohesion in the discussion below, the comments on emerging theoretical (and practical) issues at the points where alternative options were possible will be given right after a particular choice is made, rather than in a separate section.

A. *Cross-sectional vs. Time-series Approach*

Having a nine-year long time series of both IP and GDP values for 14 zones, and opting for the regression analysis as a forecasting tool, it was possible either to try to set the common relationship for all zones, or, alternatively, to model an IP versus GDP temporal relationship (multivariate model) for each of the zones considered. Pros and cons for each of the alternatives have been widely discussed in the referent literature.

It is generally argued that the kind of behaviour measured from cross-sectional data is typically long-run in nature, whereas time-series data tend to yield short-run responses [11], [12]. It is certain that by establishing a common pattern, cross-sectional models have advantages in terms of generality and ease of interpretation. However, the gain in generality is usually offset by a corresponding loss of predictive accuracy, since it is fairly unlikely that the static cross-sectional structure can satisfactorily reflect the changing nature of the underlying process for each of the observational units. On the other hand, the impact of sample size on the choice of the modelling approach has to be recognized too, and a time-series approach typically requires relatively large numbers of temporal observations to draw statistically valid conclusions [13].

Both approaches are, however, concerned with certain theoretical problems when using linear regression for model parameter estimation. Cross-sectional observations are frequently *heteroscedastic*, that is, the assumption of equal variance over observations is often violated, usually because of differing factors related to the size of the different cross-sectional entities or varying background conditions [14]. On the other hand, time-series models typically violate the assumption of independence of the error terms, i.e. they suffer from *autocorrelation* [15], [16]. The primary cause of this problem is considered to be the failure to include one or more important regressors in the model when there happens to be a high degree of temporal correlation in their cumulative effect [2]. The violation of the assumption of *normally distributed disturbances* can generally also pose a problem, yet in cases of point forecasts needed "ignoring normality will not hinder the ability to make predictions" [17]. The presence of these and some related problems might affect the efficiency and unbiasedness of parameter estimates and hence lead to invalid conclusions, unless they are appropriately tackled [14], [18].

B. Original Model Concept

Cross-analysing the EUROSTAT air passenger data and the corresponding data on the zonal GDP, it turned out that the number of international passengers flying to/from the zones (IP) is highly correlated with their GDP level. The power curve

$$\mathbf{IP = \alpha \times GDP^{\beta}} \qquad (1)$$

proved to facilitate this relationship most appropriately, with a multiple determination coefficient (R^2) ranging from 0.86 to 0.88 for every year from 1997 to 2003.

Having decided on the form of the relationship, the following step was to select the base year for the model parameter estimation. For more than one reason our choice was 1997. Firstly, it was a relatively calm (free of external disturbances) year concerning the air transport industry. The economic recession of the early 1990s was already a few years away. In addition, the corresponding statistical database turned out to be richer relative to other years. Finally, it was considered an advantage to be able to test the model on annual data both preceding and succeeding the base modelling year. This was thought to be helpful in both detecting its possible shortcomings and at the same time testing its temporal (un)stability.

After linearising, equation (1) becomes:

$$\mathbf{IP' = \alpha' + \beta \times GDP'} \qquad (2)$$

where:

$$IP' = \ln(IP),$$
$$\alpha' = \ln\alpha, \text{ and}$$
$$GDP' = \ln(GDP).$$

The summary of the model (2) parameter estimation for 1997 is shown in Table I.

TABLE I
TRIP GENERATION MODEL PARAMETER ESTIMATION, 1997

Expl. Var.	Coeff.	Est. Coeff.	t-stat.
Constant	α’	10.2425	14.061
GDP’	β	0.5358	9.337
R^2_{adj}	0.869		

Taking anti-logarithm of (2) the original trip generation model finally becomes:

$$\mathbf{IP = 28072 \times GDP^{0.5358}} \qquad (3)$$

where:

IP stands for the volume of international air passengers flying to/from the zone considered, and

GDP stands for the gross domestic product (in millions of 2003 US$) of the corresponding zone.

Figure 2 shows the fitted regression curve (3) for the 1997 data. Although GDP itself explained 87% of the total variance, which can be considered a fairly good result, it can be seen that in a few cases a difference between estimated and observed traffic was beyond tolerable, especially having in mind the forecast to be made. Traffic was significantly underestimated for two zones (UKI and SPA, see Table III for the list of zones and labels used), while Italian trip production was noticeably overestimated.

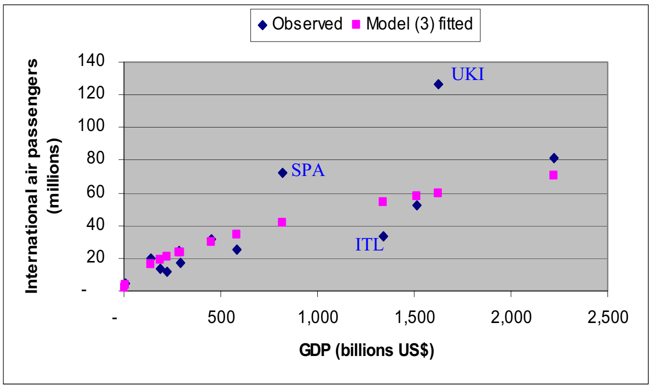


Fig. 2. Trip generation model estimation, 1997 data.

C. Checking Regression Assumptions

A plot of *model fitted values vs. residuals* was inspected first to assess homoscedasticity [17]. There seemed to be no severe violation of the homoscedasticity assumption, i.e. no consistent pattern of the disturbances magnitude with respect to the magnitude of the fitted dependent variable was observed.

The normality, on the other hand, appeared to be somewhat violated, which was ascertained through the analyses of disturbances’ summary statistics and normal probability Q-Q plots, as well as by conducting the Kolmogorov-Smirnov test. However, it has to be stressed that the sample size (only 14 observations in this case) directly influences the degree to which disturbances exhibit “normal“ behaviour, hence normally distributed small samples can exhibit “apparently non-normal behaviour as viewed through diagnostic tools” [17]. Anyway, as already mentioned, in cases when prediction of the point estimates is the only modelling purpose, the impact of non-normality on accuracy is practically negligible, but it could partly explain the presence of the regression outliers [1].

The regression outliers could be dealt with in different ways. They could simply be removed from the regression and modelled separately whereas the original model would be re-estimated, with a likely improvement of fit, but also a certain loss of coherence. Alternatively, the outliers could be left in, and some other way sought to resolve the problem and make the model fit the data [17]. However, prior to opting for any of these choices, there still remained a temporal dimension of the problem to be examined.

D.Residual Analysis

Testing (3) on a number of annual datasets, a certain regularity has been observed. For most of the zones, there turned out to be a high degree of consistency in the behaviour of the residuals (conveniently transformed to relative terms - ratios of observed/estimated IP) over time, i.e. a clear monotony has been observed in the way that the estimated number of passengers differs from the observed one for the zone considered (Figure 3). On the other hand, for BLX and SWI zones, the variations of this ratio have coincided with the crises and eventual bankruptcy of Sabena and Swissair in 2001 and 2002.

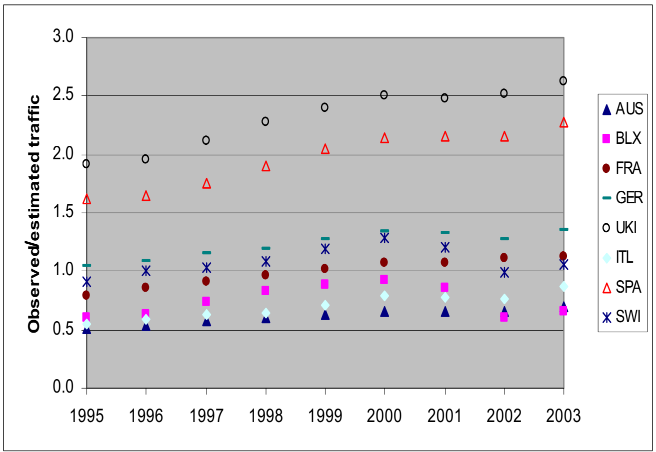


Fig. 3. Zonal residual time-series.

It has already been mentioned that using regression for time series analysis could potentially pose the problem of the linear association of successive residuals, referred to as autocorrelation. But it is not necessarily bad news. Put another way, use could be made of the strong dependence observed. On the one hand, it has been recognized that the error terms contain the information on the influence of the factors not explicitly included in the model [13]. A number of methods have been developed attempting to exploit this opportunity, e.g. “cross - sectionally heteroscedastic and timewise autoregressive model” [18], “error components model” [19], “regression models with a time-series error component” [20], etc. These generally fall under the *panel (or pooled) data analysis* category, as they deal with cross-sections of observational units observed over time. Most of them are, however, associated with relatively complex statistical procedures for parameter estimation, yet little is known on their performance with small samples [13].

On the other hand, a simple transformation of original error terms into *ratios of observed/fitted values* (see y-axis, Figure 3) implies an analogy with what is known as the theory of intrinsic characteristics.

E. The Theory of Intrinsic Characteristics

The idea has originally been employed in a model estimating the trip distribution pattern for a set of North-American cities. It points out the inability of a general transportation demand model to account for all the differences in the trip production and attraction characteristics of zones. Realizing that the error term may include the effect of all omitted variables it has been suggested that this might be used for forecasting purposes. For that reason, the error term has been decomposed into two independent components: “*intrinsic characteristic*” (reflecting the specificity of the corresponding city-pair) and the *random error*. Both Generalised Least Squares (GLS) and Ordinary Least Squares (OLS) methods were used to estimate the intrinsic characteristics (IC), with GLS performing somewhat better. Finally, the original application of this theory assumed that the intrinsic characteristics do not change substantially over time [21].

F. Critique, Generalisation and Synthesis

More than a few authors, however, oppose the constant specification models. Chatfield, for instance, argues that “a local model, which changes through time, may be preferred to a global model, that has constant parameters” [1]. Kanafani suggests that the model specifications can be altered to permit variable elasticities and “to allow for a levelling off of air traffic growth as related to the growth in the demand variables” [6].

With all of this put together, the possibility of generalisation of the original concept of intrinsic characteristics to the trip generation model seemed rather straightforward. The basic hypothesis, that the effects of the conventional demand variables on intercity travel vary with the economic and social characteristics of the city, applies equally well to the traffic generating potential of different zones. It is both intuitively clear and in accordance with empirical evidence that traffic growth is driven by a different (both qualitatively and quantitatively) blend of factors for different geographical entities. The extent to which factors other than GDP influence the trip production volume of different zones could therefore be represented by zonal intrinsic characteristics. In addition, it is to be expected that the intensity of this aggregate effect changes with time, reflecting the temporal changes of the factors constituting it (e.g. transport supply, tourism, external trade, etc.).

In summary, the following approach was adopted: to exploit the strong temporal correlation between the successive zonal relative residuals, referred to as *generalised zonal intrinsic characteristics* (GIC), and derive a separate trend for each zone’s GIC series, which would finally be integrated into the original model. Hence, a dynamic element has been incorporated into the authentic IC theory, in order to reflect the changing nature of the underlying process with time.

It turned out that in most of the cases a simple linear approximation of the GIC time-series provided well-fitted sub-models (Figure 4). Although it might be more reasonable to assume that the regression coefficients in the sub-models evolve stochastically with time, giving rise to what is called a “stochastic trend” [1], the practical issues (i.e. point forecast needed) were again overwhelming.

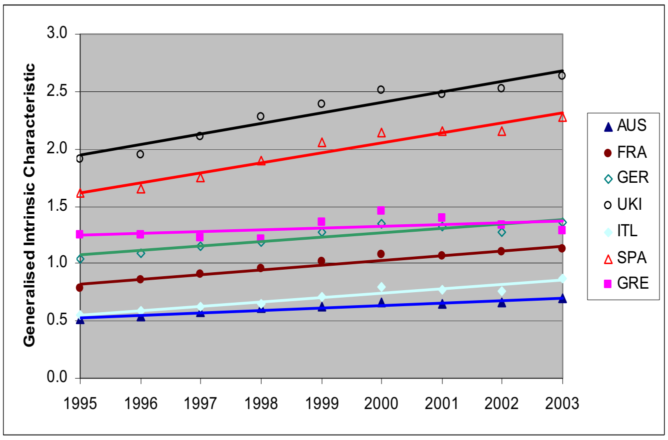


Fig. 4. Zonal generalised intrinsic characteristics, linear approximation.

Thus the deterministic GIC trends were finally incorporated into the original model, so that the modified trip generation model became:

$$\mathbf{IP_{zt} = 28072 \times GDP_{zt}^{\;0.5358} \times GIC_{zt}} \qquad (4)$$

where:

$\mathbf{IP_{zt}}$ is the number of international air passengers flying to/from zone **z** in a year **t**;

$\mathbf{GDP_{zt}}$ is the gross domestic product (in millions of US$) of a zone **z** in a year **t**, while

$\mathbf{GIC_{zt}}$ stands for the generalised intrinsic characteristic of a zone **z** in a year **t**.

IV. Model Validation

Model (4) validation was originally conducted using 2003 figures on air passenger traffic and the GDP levels, whereas zonal GICs where derived based on 1995-2002 time-series. The series has been shortened one year so that the model prediction could be compared with information not used during the process of the model estimation [3].

The forecasting output for 2003 is presented in Figure 5, along with the observed values and the model (3) predictions. The gain in accuracy is more than obvious, with *Mean Average Percentage Error* (MAPE) being 8.3% for model (4) estimates, while the overall 2003 traffic volume for 14 zones considered was overestimated by 3.6%, Table II.

Table III presents the comparison of the model (4) estimates and in the interim released observed traffic volumes in 2004.

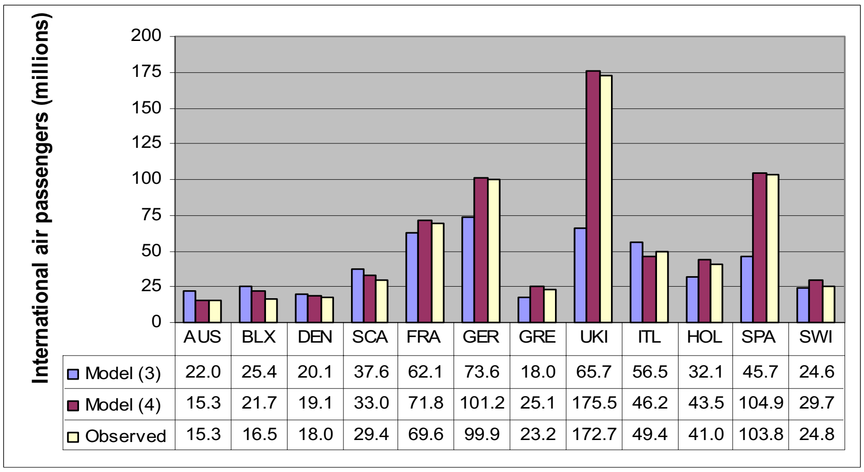


Fig. 5. Model validation, 2003 data.

TABLE II
MODEL VALIDATION – SUMMARY STATISTICS

Year	2003			2004		
	MAPE (%)	Total traffic volume prediction error (%)	R^2_{adj}	MAPE (%)	Total traffic volume prediction error (%)	R^2_{adj}
Model (3)	28.2	-27.1	0.646	-	-	-
Model (4)	8.3	+3.6	0.998	4.95	-1.6	0.997

TABLE III
MODEL (4) VALIDATION – 2004 DATA
(millions of international air passengers)

Zone label	AUS	BLX	DEN	SCA	FRA	GER	GRE	UKI	ITL	HOL	SPA	SWI	Total
Constituting countries	Austria	Belgium & Luxemb.	Denmark	Sweden, Finland & Norway	France	Germany	Greece	UK & Ireland	Italy	Holland	Spain & Portugal	Switzerland	
Observed	17.7	19.0	19.4	32.3	76.0	114.5	24.0	186.7	57.0	44.4	111.7	25.8	728.6
Estimated	16.0	20.2	19.9	32.5	75.4	105.5	25.6	185.3	50.6	45.6	111.1	28.5	716.1

V. SUMMARY AND CONCLUSIONS

To obtain a forecast of the international air passenger traffic production for a set of European zones, a sequential combination of cross-sectional and time-series data analysis has been employed, resulting in a common trip generation model established which also allows for the specificity of the different geographical entities considered. This has been accomplished by incorporating a dynamic element into the theory of intrinsic characteristics, accounting that way for the time-varying gross effect of the causal variables not explicitly specified in the model.

The modification applied significantly improved the explanatory power and predictive accuracy of the original cross-sectional model, producing fairly accurate forecasts as a result. Two out-of-sample validation tests performed showed the reduction of MAPE from 28% to 5-8%, with the corresponding R^2_{adj} value rise from 0.65 to 0.99.

The model specification allows for a straightforward re-calibration once any new historical data are released, with a consequent likely improvement in its forecasting accuracy. This feature is in accordance with theoretical recommendations on the use of macro-models with constant elasticities for transport demand forecasting [6].

However, the model presented should primarily be seen in the light of its particular context, i.e. as a convenient practical tool to obtain short-to-medium run forecasts in situations with small samples and/or relatively short historical data time-series.

Further generalisations of the concept are, nevertheless, possible, and more advanced methods could be employed to examine in more detail the behaviour of zonal generalised intrinsic characteristics over time. That would generally enable a plausible interpretation of the different impact of extraneous factors on the predicted variable, with a likely improvement of the overall effectiveness of the forecasting system. Examining the effects of the integration of such methodology into the concept presented might be an interesting problem for further research.

REFERENCES

[1] C. Chatfield, *The Analysis of Time Series: An Introduction.* 6th ed. Boca Raton: Chapman & Hall/CRC Press, 2004.

[2] D. C. Montgomery, L. A. Johnson, and J. S. Gardiner, *Forecasting and Time Series Analysis.* 2nd ed. New York: McGraw-Hill, 1990.

[3] J. de D. Ortuzar and L. G. Willumsen, *Modelling Transport*. 2nd ed. Chichester: John Wiley & Sons, 1994.

[4] Faculty of Transport and Traffic Engineering, University of Belgrade, *Serbia and Montenegro (SCG) Air Traffic Analysis and Forecast.* Final report. Belgrade: FTTE, 2005.

[5] D. Hensher and K. Button (eds.), *Handbook of Transport Modelling.* Amsterdam: Pergamon, 2000.

[6] A. Kanafani, *Transportation Demand Analysis*, New York: McGraw-Hill, 1983.

[7] N. J. Ashford and P. H. Wright, *Airport Engineering*. 3rd ed. New York: John Wiley & Sons, 1992.

[8] L. Castelli, S. Schiratti, and W. Ukovich, "Analysis of passenger demand for air transportation", Work Package 1 of the EUROCONTROL *CARE Innovative Action Project: "Innovative Route Charging Schemes"*. Final Report, 2002.

[9] International Civil Aviation Organization, *Annual Review of Civil Aviation*, 2002.

[10] EUROSTAT, "Statistics in Focus, Highlights of the Panorama of Transport 1970 – 1999", in *Transport* 3/2002, 2002.

[11] E. Kuh and J. Meyer, "How extraneous are extraneous estimates?", in *The Review of Economics and Statistics.* Vol. 39, No. 4, 1957, pp. 380-393.

[12] B. Baltagi and M. Griffin, "Short and long run effects in pooled models", in *International Economic Review*. Vol. 25, No. 3, 1984, pp. 631-645.

[13] T. E. Dielman, *Pooled Cross-Sectional and Time Series Data Analysis.* New York: Marcel Dekker, Inc, 1989.

[14] D. L. Harnett and J. L. Murphy, *Statistical Analysis for Business and Economics.* 3rd ed, Reading, Massachusetts: Addison-Wesley Publishing Company, 1985.

[15] T. H. Wonnacott and R. J. Wonnacott, *Introductory Statistics for Business and Economics*. New York: John Wiley & Sons, 1972.

[16] C. Chatfield, *Time-Series Forecasting*. Boca Raton: Chapman & Hall/CRC Press, 2001.

[17] S. P. Washington, M. G. Karlaftis, and F. L. Mannering, *Statistical and Econometric Methods for Transportation Data Analysis.* Boca Raton: Chapman & Hall/CRC, 2003.

[18] J. Kmenta, *Elements of Econometrics.* 2nd ed. New York: Macmillan Publishing Company, 1986.

[19] T. Wallace and A. Hussain, "The use of error components models in combining cross-section with time-series data", in *Econometrica.* Vol. 37, No. 1, 1969, pp. 55-72.

[20] A. H. Choudhury, R. Hubata, and R. D. St. Louis, "Understanding time-series regression estimators," in *The American Statistician.* Vol. 53, No. 4, 1999, pp. 342-348.

[21] R. E. Quandt and K. H. Young, "Cross-sectional travel demand models: estimates and tests," in *Journal of Regional Science.* Vol. 9, No. 2, 1969, pp. 201-214.

Radosav B. Jovanović (1978) is at the moment a full-time student at the MSc Transport Planning and Management programme, Transport Studies Group, University of Westminster, London. He graduated from the Air Transport Department, Faculty of Transport and Traffic Engineering, University of Belgrade, Serbia and Montenegro, in 2004. The principal fields of his interest include transport demand analysis, transport planning and transport economics, in particular air transport.

Session 5

Future Concepts

Sculpting the sound of aircraft: a novel MDO approach for noise annoyance alleviation

Umberto Iemma, Matteo Diez, & Vincenzo Marchese
Department of Mechanical and Industrial Engineering
University *Roma Tre*
via della vasca navale 79, 00146, Roma, Italy
Email: u.iemma@uniroma3.it

***Abstract*— The paper presents a novel approach to include community-noise considerations in a MDO (Multidisciplinary Design Optimization) formulation for the conceptual design of commercial transportation aircraft. The basic idea deals with the introduction of the *sound quality* concept into the design of aircraft, as an alternative to the classical sound–level–based approach, currently used for the environmental certification of airplanes. In the present work, the noise produced by the configuration under analysis is compared to a *target sound*, and the design concept is marked as a good one if the matching between the two sounds is satisfactory. This research is part of the European project SEFA (Sound Engineering For Aircraft), whose aim is to provide indications to the designers on how to reduce the annoyance of aircraft noise by improving the *quality* of the acoustic emissions. The contribution of the authors to SEFA deals with the development of a methodology for including this concepts into an MDO framework. It is evident that such an approach requires an optimization algorithm capable to modify the *shape* of the noise spectrum, rather than only reduce its level, thus requiring the use of simulation modules dedicated to each noise source. In this respect, the use of the word *sculpting* in the title of this work is twofold: on one hand, it effectively describes the practical difficulty of *shaping* the noise spectrum in the whole audible range; at the same time, it evocates, in a metaphorical sense, the aesthetic implication in the struggle for making *beauty* emerge from a rough block of rock. In order to define a robust method to *quantify* the matching between the current sound and the target, the *Sound Similarity Index* (SSI) is introduced and defined as the L^p–norm of the difference of the two sounds spectra. Preliminary results show that the SSI represents a good *measure of the distance* between the two sounds, and that is capable to effectively drive the optimization process towards those configurations satisfying sound quality requirements.**

I. Introduction

The formulation presented in this paper represents a novel approach to include community-noise considerations in a Multidisciplinary Design Optimization (MDO) framework for the conceptual design of civil transportation aircraft. In the present work, the attention is focused on the improvement of the quality of aircraft sound, as an alternative to the classical approach based on the reduction of the EPNL (Effective Perceived Noise Level). To accomplish this, the noise emissions of the aircraft concept under analysis is compared to a *target sound*, whose characteristics will be defined in the following; if the matching of the two sounds is satisfactory (according to a suitable similarity criterion), then the design concept can be considered as a “good one” in terms of acoustic annoyance. The present work is part of the European project SEFA, Sound Engineering for Aircraft. The aim of SEFA is to provide the aircraft designers with indications and recommendations for improving the acceptance of aircraft sound by the population. To this aim, the most annoying features of the aviation acoustic emissions are identified through an extensive campaign of psychometric tests, and modern techniques of sound engineering are used to synthesize a *target sound* having all the characteristics yielding a well-accepted (or, at least, weakly–annoying) noise. The problem is addressed in a highly multidisciplinary fashion, and twenty research groups from eight European countries are involved in a wide-spreading research activity. The need of such an innovative approach (at least, in the aeronautical community) stems from several concurring factors. Indeed, the air transportation system is experiencing, nowadays, a tremendous growth-rate, especially in the European countries. The commercial success of the low-cost philosophy, and the increasing demand of air cargo for goods transportation are causing the intensification of the air traffic. As a consequence of this rise in the market demand, a large number of secondary airports are experiencing a commercial rebirth, due to lower rates charged to airlines for operating. For all these reasons, the number of citizens exposed every day and every night to a heavy acoustic pollution due to aircraft operations, is enormously increased during the last decade. In this scenario, the community noise is becoming a crucial issue in the development of the air transportation system. Authorities and governments are introducing much stricter regulations in noise emissions for next aircraft generation, and, at the same time, are charging the airline operators with increasing noise taxes, to make the management of noisy fleets more expensive. In addition, the noise emission of currently–flying aircraft can not be significantly further reduced in the next future, without the introduction of highly innovative concepts, whose development is, typically, very expensive and time-consuming. In this respect, the approach presented in this work may give an effective alternative to the designers, introducing the quality of the sound as an additional ”degree of freedom” in the reduction of the aircraft annoyance.

The main difficulty in the present methodology resides in the fact that the “tracking” of a reference spectrum during the optimization process requires a dedicated description of each noise source present onboard. In addition, the matching level of the two sounds must be somehow *quantified* by introducing a carefully defined similarity index. In the present work, we propose to evaluate the matching of the two sounds by means of a Sound Similarity Index (SSI), defined as the L^p-norm of the difference between the two spectra. The properties of this parameter appear to suit well the requirements of the problem at hand, and preliminary results reveal that the optimization process is effectively driven towards those aircraft configurations whose emissions satisfy the matching with the target.

In the next Section, a brief overview of the fundamentals simulation modules (*i.e.*, structural dynamics, aerodynamics, aeroelasticity, and flight mechanics) implemented in the MDO formulation is given. In Section III, the approach for the inclusion of the aeroacoustic simulation modules is presented, whereas, in Section IV, the SSI is introduced, and its use as a part of the objective function is outlined. Finally, Section V presents preliminary numerical results obtained considering only the simulation modules related to the airframe noise sources.

II. The main MDO framework

The MDO framework on the basis of the present work has been developed by the authors and their collaborators during the last decade (see Refs. [3]- [13]), and is extended here to include community noise considerations based on sound matching. Before starting the description of the additional variables and features introduced in the present work, we would like to outline the structure of the pre-existing modules, in order to give to the reader an idea about the level of detail and accuracy of the modeling used. In the following subsections, the structural dynamics, aerodynamics, aeroelasticity, and flight mechanics modules are briefly described.

A. Structural Analysis

The model utilized for the structural analysis of the wing is that of a six degree of freedom bending-torsional beam in three-dimensional space, with geometric and structural parameters varying in the spanwise direction. These include structural element geometric dimensions (rib area, spar and skin panel thickness, etc.), wing twist, mass properties plus bending and torsional moments of inertia. A simple linear root-to-tip variation law has been used for the geometric parameters for both wing and tail; furthermore clamped boundary conditions have been considered at root in order to take into account the wing-fuselage and the tail-fuselage junctures.

Set

$$\mathbf{u}(\mathbf{x},t)=\sum_{j=1}^{N}u_j(t)\boldsymbol{\psi}_j(\mathbf{x}) \tag{1}$$

where $\mathbf{x}$ are the undisturbed spacial Cartesian coordinates ($u_j(t)$ are the parameters - state variables - chosen to define the function $\mathbf{u}(\mathbf{x},t)$) and $\boldsymbol{\psi}_j(\mathbf{x})$ are the base functions of the FEM approach (three displacements, two slopes and the twist angle at each node). Using the Galerking method, *i.e.*, projecting in the direction of the base function $\boldsymbol{\psi}_j(\mathbf{x})$, one obtains

$$\mathsf{M}\ddot{\mathsf{u}}+\mathsf{K}\mathsf{u}=\mathsf{f} \tag{2}$$

where $\mathsf{u}=\{u_i\}$ is the state-space vector, whereas $\mathsf{f}=\{f_j\}$ is the external load vector. In the present applications Eq. 2 is used for both statics (*i.e.*, stress analysis) and dynamics (*i.e.*, aeroelastic analysis; divergence and flutter are dealt with using the same algorithm). In the first case, the inertial terms vanish, and Eq. 2 is used to determine the displacement vector u, and from this the maximum stress. In the second one, Eq. 2 is used to determine the natural modes. In this case, the forcing term vanishes and the associated eigenproblem is given by

$$-\omega_n^2\mathsf{M}\mathsf{z}_n+\mathsf{K}\mathsf{z}_n=\mathbf{0} \tag{3}$$

where ω_n ($n=1,...,N$) denotes the n-th natural frequency, corresponding to the n-th eigenvector, $\mathsf{z}_n=\{z_{n_i}\}$, here assumed to be normalized so that the so-called generalized masses equal one. The solution of the above eigenproblem is used to model the dynamics of the wing using a modal approach. The displacement field is expressed in terms of the approximate modes of vibration, $\boldsymbol{\Phi}_m(\mathbf{x})$, as

$$\mathbf{u}(\mathbf{x},t)=\sum_{m=1}^{M}q_m(t)\boldsymbol{\Phi}_m(\mathbf{x}). \tag{4}$$

Note that typically $M\ll N$ (*i.e.*, the modes required are much less than the finite-element degrees of freedom necessary to evaluate the modes). This is the main reason to prefer an expansion in terms of approximate modes, instead of the finite-element shape functions, $\boldsymbol{\psi}_j(\mathbf{x})$ (for a deeper analysis of this point, see Morino and Bernardini [21]). The corresponding Lagrange equations of motion are

$$\ddot{\mathsf{q}}+\Omega^2\mathsf{q}=\mathsf{e} \tag{5}$$

where q denotes the Lagrangian-coordinate vector, Ω the diagonal matrix of the wing natural frequencies, and $\mathsf{e}=\{e_n\}$ the vector of the generalized forces

$$e_n=-\oint_{S_B}p\mathbf{n}\cdot\boldsymbol{\Phi}_n d\mathcal{S}. \tag{6}$$

B. Aerodynamic Modeling

As mentioned above, the basic physical model used for aerodynamics is that of quasi-potential flows, *i.e.*, flows that are potential everywhere except for the wake surface, S_W, which is the locus of the points emanating from the trailing edge. The effects of viscosity are included through a boundary-layer model, which gives an adequate estimate for the viscous drag.

For the sake of simplicity, the quasi-potential formulation is outlined here only for incompressible flows (for details, see Morino [22] and Morino and Bernardini [23]). The extension of the formulation which includes the effects of compressibility is available [14], [22], and is included in the code.

Kelvin's theorem implies that an inviscid, incompressible, initially-irrotational flow remains, at all times, quasi-potential, as defined above. In this case, the velocity field, $\mathbf{v}$, may be expressed as $\mathbf{v} = \nabla\varphi$ (where φ is the velocity potential). Combining with the continuity equation for incompressible flows, $\nabla \cdot \mathbf{v} = 0$, yields

$$\nabla^2\varphi = 0. \tag{7}$$

The boundary conditions for this equation are as follows. The surface of the body (aircraft), S_B, is assumed to be impermeable. This implies $(\mathbf{v} - \mathbf{v}_B) \cdot \mathbf{n} = 0$, *i.e.*,

$$\frac{\partial\varphi}{\partial n} =: \chi = \chi_B := \mathbf{v}_B \cdot \mathbf{n} \qquad (\mathbf{x} \in S_B) \tag{8}$$

where $\partial/\partial n = \mathbf{n} \cdot \nabla$, whereas $\mathbf{v}_B$ is the velocity of a point $\mathbf{x} \in S_B$, and $\mathbf{n}$ is the outward unit normal to S_B. At infinity, in a frame of reference fixed with the unperturbed air, we set $\varphi = 0$. The boundary conditions on the wake surface, S_W, are obtained from the conservation of mass and momentum across a surface of discontinuity, and are: (i) the wake surface is impermeable, and (ii) the pressure, p, is continuous across it. These imply: (i) $\Delta(\partial\varphi/\partial n) = 0$ (where Δ denotes discontinuity across S_W), and (ii) $\Delta\varphi$ = constant in time following a wake point $\mathbf{x}_W$, the velocity of which is, by definition of $\mathbf{x}_W$, the average of the fluid velocity on the two sides of the wake. Thus, $\Delta\varphi(\mathbf{x}_W, t)$ equals the value it had when $\mathbf{x}_W$ left the trailing edge:

$$\Delta\varphi(\mathbf{x}_W, t) = \Delta\varphi(\mathbf{x}_{TE}, t - \tau) \tag{9}$$

where τ is the convection time from $\mathbf{x}_{TE}$ to $\mathbf{x}_W$. The value of $\Delta\varphi$ at the trailing edge is obtained by imposing the trailing-edge condition that, at the trailing edge, $\Delta\varphi$ on the wake equals $\varphi_2 - \varphi_1$ on the body, where the subscripts 1 and 2 denote the two sides of the wing surface (for a detailed analysis of this issue, see Morino and Bernardini [23]).

In the methodology used in the MDO code, the above problem for the velocity potential is solved by a boundary integral equation method. In formulating unsteady aerodynamics (used for flutter analysis), the problem is linearized: this implies that the wake surface is considered as fixed in the body frame of reference (specifically, composed of vortex lines parallel to the undisturbed velocity; accordingly $\tau = (x_W - x_{TE})/U_\infty$, where U_∞ is the undisturbed flow velocity). Then, the boundary integral representation for the above problem in the Laplace domain is given by

$$\begin{aligned}\tilde{\varphi}(\mathbf{x}) &= \int_{S_B} \left(G\tilde{\chi} - \tilde{\varphi}\frac{\partial G}{\partial n} \right) dS(\mathbf{y}) \\ &- \int_{S_W} \Delta\tilde{\varphi}_{TE} e^{-s\tau} \frac{\partial G}{\partial n} dS(\mathbf{y})\end{aligned} \tag{10}$$

where $\tilde{\chi}$ denotes the Laplace transform of the unsteady portion of $\mathbf{v}_B \cdot \mathbf{n}$ (so as to have $\chi = 0$ for $t < 0$), whereas $G = -1/4\pi\|\mathbf{y} - \mathbf{x}\|$. Note that, by applying the trailing-edge condition, $\Delta\tilde{\varphi}_{TE}$ may be expressed in terms of $\tilde{\varphi}$ over the body. Thus, Eq. 10, in the limit as $\mathbf{x}$ tends to S_B, represents a boundary integral equation for $\tilde{\varphi}$ on S_B, with $\tilde{\chi}$ on S_B known from Eq. 8. Once $\tilde{\varphi}$ on S_B is known, $\tilde{\varphi}$ (and hence $\mathbf{v}$ and, by using Bernoulli's theorem, p) may be evaluated everywhere in the field. This boundary integral equation is solved numerically by discretizing the body and wake surfaces with quadrilateral elements, assuming $\tilde{\varphi}$, $\tilde{\chi}$, and $\Delta\tilde{\varphi}$ to be constant within each element, and imposing that the integral equation be satisfied at the center of each surface element (zeroth-order boundary-element collocation method). This yields

$$\tilde{\mathsf{f}}_\varphi = \mathsf{E}_{IE}\tilde{\mathsf{f}}_\chi \tag{11}$$

where $\tilde{\mathsf{f}}_\varphi = \{\tilde{\varphi}_j\}$ and $\tilde{\mathsf{f}}_\chi = \{\tilde{\chi}_j\}$ are the vectors of the values of $\tilde{\varphi}$ and $\tilde{\chi}$ at the centers of the elements (see Morino [22] for details).

Next, consider the modeling for viscous flows. In view of the stated applications (*i.e.*, civil aviation), the analysis is limited to attached high-Reynolds-number flows, where the vortical region (*i.e.*, boundary layer and wake) has a small thickness. Also, we consider only steady flows, since in aeroelastic applications the viscous effects are typically neglected, unless one deals with control surfaces and/or stall, not examined here. Outside boundary layer and wake, the flow is quasi-potential and is solved by using the model described above. The viscous flows model is based upon the transpiration velocity concept, *i.e.*, the Lighthill [24] equivalent-source approach. This consists of modifying the boundary condition for $\partial\varphi/\partial n$, Eq. 8, into

$$\chi = \chi_B + \chi_V \tag{12}$$

where the transpiration velocity χ_V is given by

$$\chi_V = \frac{\partial}{\partial s_1}\int_0^\delta (u_e - u)\, d\eta + \frac{\partial}{\partial s_2}\int_0^\delta (v_e - v)\, d\eta \tag{13}$$

where s_1 and s_2 are local orthogonal arclengths over the wing surface, η is the arclength along the normal, δ is the boundary-layer thickness, and u_e and v_e the velocities at the external edge of the boundary layer respectively in s_1 and s_2 directions. A similar correction is used on the wake surface, with $\Delta\left(\partial\varphi/\partial n\right) = (\chi_V)_2 + (\chi_V)_1$ (for an in-depth analysis of this point, see Morino *et al.*[25], where an exact extension of Lighthill [24] is presented). An integral two-dimensional boundary-layer strip-theory formulation is used for the boundary layer (for attached flows, this approach yields results as accurate as those obtained by differential boudary-layer methods with considerably reduced computational effort). Indeed, three dimensional effects within the boundary layer may be neglected with a minor loss of accuracy of the predictions for applications to wings with large aspect ratio and reduced sweep angle.

C. Aeroelastic Analysis

It is known that in linear unsteady aerodynamics, the frequency ω always appears together with the air speed U_∞, through the dimensionless parameter $k = \omega\ell/U_\infty$, known as the reduced frequency (ℓ is a reference length). Accordingly, in our case, we formulate the problem in terms of the dimensionless Laplace parameter $\check{s} = s\ell/U_\infty$, known as the complex reduced frequency; for, if $\omega = Imag(s)$, then $k = Imag(\check{s})$

(here, the symbol p typically used for the complex reduced frequency is avoided, since here it is used to denote pressure). Then, we have

$$\tilde{\mathsf{e}} = q_{_D}\ \mathsf{E}(\check{s})\ \tilde{\mathsf{q}} \tag{14}$$

where $q_{_D}$ is the dynamic pressure. The specific expression for the matrix of the generalized aerodynamic forces, $\mathsf{E}(\check{s})$, in Eq. 14 may be obtained as

$$\mathsf{E}(\check{s}) = \mathsf{E}_{_{GF}} \mathsf{E}_{_{BT}}(\check{s}) \mathsf{E}_{_{IE}}(\check{s}) \mathsf{E}_{_{BC}}(\check{s}) \tag{15}$$

where:

(i) the matrix $\mathsf{E}_{_{BC}}(\check{s})$ is obtained from the boundary condition (Eq. 8, *i.e.*, $\chi = (U_\infty \mathbf{i} + \sum_n \dot{q}_n \mathbf{\Phi}_n) \cdot (\mathbf{n}_0 + \sum_m \Delta\mathbf{n}_m q_m)$, where $\mathbf{n}_0$ is the unit normal to the undeformed surface and $\Delta\mathbf{n}_m$ is the variation of $\mathbf{n}$ due to q_m), and relates the Laplace-transformed vector $\tilde{\mathsf{f}}_\chi = \{\tilde{\chi}_n / U_\infty\}$ of the dimensionless linear unsteady portion of the normalwash, evaluated at the element centers, to the Laplace-transformed Lagrangian-coordinate vector $\tilde{\mathsf{q}} = \{\tilde{q}_n\}$, as $\tilde{\mathsf{f}}_\chi = \mathsf{E}_{_{BC}}\ \tilde{\mathsf{q}}$;

(ii) the matrix $\mathsf{E}_{_{IE}}(\check{s})$ is obtained from the discretization of the integral equation, and relates the Laplace-transformed vector of the dimensionless velocity potential, $\tilde{\mathsf{f}}_\varphi = \{\tilde{\varphi}_n / U_\infty \ell\}$, evaluated at the element centers, to $\tilde{\mathsf{f}}_\chi$, as $\tilde{\mathsf{f}}_\varphi = \mathsf{E}_{_{IE}} \tilde{\mathsf{f}}_\chi$ (see Eq. 10);

(iii) the matrix $\mathsf{E}_{_{BT}}(\check{s})$ is obtained from the discretization of the linearized Bernoulli theorem, $c_p = -2(\dot{\varphi} + U_\infty \partial\varphi/\partial x)/U_\infty^2$, and relates the Laplace-transformed vector of the pressure coefficient evaluated at the element centers, $\tilde{\mathsf{f}}_{c_p}$, to $\tilde{\mathsf{f}}_\varphi$, as $\tilde{\mathsf{f}}_{c_p} = \mathsf{E}_{_{BT}} \tilde{\mathsf{f}}_\varphi$;

(iv) the matrix $\mathsf{E}_{_{GF}}$ is obtained from the discretization of Eq. 6, and relates the Laplace-transformed vector of the generalized aerodynamic forces, $\tilde{\mathsf{e}}$, to $\tilde{\mathsf{f}}_{c_p}$, as $\tilde{\mathsf{e}} = q_{_D} \mathsf{E}_{_{GF}} \tilde{\mathsf{f}}_{c_p}$.

In order to perform the aeroelastic analysis in the framework of the optimization procedure, a finite-state approximation (reduced order model, ROM) for the aerodynamic matrix $\mathsf{E}(\check{s})$ is considered. Specifically, following Ref. [9], the transcendental function $\mathsf{E}(\check{s})$ is approximated as

$$\mathsf{E}(\check{s}) \simeq \mathsf{E}_2 \check{s}^2 + \mathsf{E}_1 \check{s} + \mathsf{E}_0 + (\check{s}\mathsf{I} - \mathsf{P})^{-1} \mathsf{R} \tag{16}$$

where all the matrices on the right hand side are evaluated by a least-square procedure starting from $\mathsf{E}(\check{s}_i)$, where $\check{s}_i$ is a suitable set of complex reduced frequencies. Substituting Eq. 16 into Eq. 5, and back-transforming into the time domain, the system can be reduced to the standard state-variable format, $\dot{\mathsf{x}} = \mathsf{A}(U_\infty)\mathsf{x}$, where the parametric dependence of the matrix A upon the air speed has been emphasized. This approach allows one to reduce the aeroelastic stability analysis to a root locus for the matrix $\mathsf{A}(U_\infty)$, thereby avoiding standard methods (*e.g.*, k and p-k method), which are cumbersome and would unnecessarily complicate the optimization process.

D. Flight Mechanics

The static longitudinal stability, an essential issue for aircraft, is satisfied by imposing that the derivative with respect to the angle of attack of pitch moment coefficient (evaluated with respect to the center of mass G) be less than zero: $C_{M_\alpha} < 0$ (static stability). This is performed by evaluating in the MDO process the global aerodynamic loads acting on the aircraft (evaluated via BEM) and the total mass distribution (so as to determine the location of G).

In order to evaluate fuel consumption, the mission profile considered in this work consists of: (i) take-off, (ii) climb, (iii) cruise, (iv) descent, and (v) landing. The range is computed according to the Breguet equation $R = (V_c E/c)\ln(W_i/W_f)$, where V_c is the cruise speed, c is the specific fuel consumption, $E = L/D$ is the aerodynamic efficiency (lift to drag ratio), and W_i and W_f the initial and final weights of the cruise segment, respectively. Finally, expressing the fuel consumptions for the mission segments before and during the cruise segment as fractions of the usable mission fuel weight W_{uf} (indicated as k_1 and k_2, respectively), W_i and W_f can be written as: $W_i = W - k_1 W_{uf}$ and $W_f = W - (k_1 + k_2) W_{uf}$.

III. Aeroacoustics

The aeroacoustic simulation deserves a careful discussion. Indeed, the prediction of the noise spectrum perceived at a specified location requires an accurate modeling of several physical phenomena. On the other hand, during a complete optimization process, each module can be called hundreds of times, and thus, a compromise between accuracy and computational efficiency is mandatory. To accomplish this, we divided the aeroacoustic simulator into three sub-modules: first, suitable prediction models are implemented for all the on-board sources of noise (the modules are implemented so as to activate only those of them which are relevant in the flight conditions under analysis); second, a scattering model of the aircraft, based on a Boundary Integral Formulation (BIE) for moving boundaries (see *e.g.*, Morino et al. [15], Iemma and Gennaretti [16], and Gennaretti and Iemma [17]), is introduced into the MDO framework to take into account the scattering effects of the fuselage, the wing and the tail; finally, a propagation predictor, based on the same BIE, is used to evaluate the spectrum that reaches an observer at rest, including Doppler and directional effects.

At the present stage of the work, the scattering and propagation modules are in course of development, and will be included soon. Nevertheless, in the present paper, our attention is focused on the effectiveness of the sound-matching criterion used to estimate how much similar the two sounds are. From this point of view, using of a simplified formulation in noise spectrum evaluation is not a limitation. In this paper the noise emissions are computed using a simplified predictor which includes only the airframe noise and implements the Fink model (for details, see Ref. [18]). Table I shows the relevant parameters used in this work for the airframe noise evaluation.

TABLE I
FINK MODEL RELEVANT PARAMETERS

wing area
wing span
horizontal tail area
horizontal tail span
vertical tail height
vertical tail area
flap span
flap area
number of slots
flap deflection angle
slats status
landing gear geometry (not used in this work)
airspeed

IV. THE SOUND SIMILARITY INDEX

A crucial point in the present approach, is represented by the criterion used in the MDO to evaluate the similarity between the *current sound* (*i.e.*, the sound produced by the configuration under analysis), and the target. An estimate of this similarity has to be quantified in some way, and included in the objective function during the optimization process. In order to do this, we focus our analysis on the difference of the two sound spectra, $g_c(\xi)-g_t(\xi)$, where ξ is a non-dimensional parameter depending on the frequency f, defined later. Then, we take advantage of the mathematical tools provided by the Functional Analysis, by considering the difference between the two spectra as belonging to a normed space $L^p(\mathcal{D})$.

Lets recall that, if a function $h(\xi)$ belongs to the normed space $L^p(\mathcal{D})$, then its L^p-norm is by definition[1]

$$\|h(\xi)\|_p := \left[\int_{\mathcal{D}} |h(\xi)|^p \mathrm{d}\xi\right]^{\frac{1}{p}}, \tag{17}$$

and the following property holds:

$$\|h(\xi)\|_\infty := \lim_{p\to\infty}\left[\int_{\mathcal{D}} |h(\xi)|^p \mathrm{d}\xi\right]^{\frac{1}{p}} = \max_{\xi\in\mathcal{D}}\{h(\xi)\}. \tag{18}$$

The desired *quantification* of the similarity of the two sounds is obtained by introducing the Sound Similarity Index, $\mathcal{I}_{\mathsf{ss}}$, defined as the L^p-norm of the difference between the two spectra, *i.e.*,

$$\mathcal{I}_{\mathsf{ss}} := \|g_c(\xi)-g_t(\xi)\|_p = \left[\int_0^1 |g_c(\xi)-g_t(\xi)|^p \mathrm{d}\xi\right]^{\frac{1}{p}}, \tag{19}$$

where g_c and g_t are respectively the current and the target spectra, suitably normalized so as to have the same L^1-norm (in this work we are more interested in the *shape* of the spectra, rather than in the noise level). Furthermore,

$$\xi := \frac{\log f - \log f_{\min}}{\log f_{\max} - \log f_{\min}}, \tag{20}$$

[1]We refer here to a simply normed space, which is, in general, a Banach space. If an inner product is defined, then the space becomes a Hilbert space. Note that every Hilbert space is a Banach space, but the converse is not always true (see [27], [28]).

is the non-dimensional, normalized parameter cited above, and introduced to make $\mathcal{I}_{\mathsf{ss}}$ independent on the measure of the domain, $\mu(\mathcal{D})$. Note that $\mathcal{I}_{\mathsf{ss}}$ can be considered as the *distance* of the two sounds in the L^p vector space the two spectra belong to. Note also that the $\mathcal{I}_{\mathsf{ss}}$ is a (very complex) function of the design/procedural variables, through the noise models introduced in Section III. Nevertheless, the nature of this dependance does not allow us to consider the SSI as a measure of the *distance* between the current aircraft configuration and the (unknown) target aircraft in the design/procedural space. In other words, two aircraft can be very similar, and produce two completely different sounds.

Recalling Equations 17 and 18, it follows that:

- for $p = 1$, a small $\mathcal{I}_{\mathsf{ss}}$ indicates a small difference in global shape of the two spectra
- for $p \to \infty$, a small $\mathcal{I}_{\mathsf{ss}}$ indicates small local differences between the two spectra.

These peculiarities make the SSI a useful parameter in the evaluation of the *distance* between two sounds. Indeed, the noise emission of an aircraft contains both tonal contributions (fan and compressor noise, buzz-saw, etc.), as well as broadband components (*e.g.*, the airframe noise). A proper choice of p makes it possible to have a measure of the concentrated differences (tonal components missing or misplaced), as well as the distributed ones. Hence, indicating with x the N-dimensional column vector containing the design and procedural variables, the objective function can be written as

$$\mathcal{J}(\mathsf{x}) = \sum_{k=1}^{M} \alpha_k\, \mathcal{F}_k(\mathsf{x}) + \alpha_{\mathsf{ss}}\, \mathcal{I}_{\mathsf{ss}}(\mathsf{x}). \tag{21}$$

In the previous equation, the $\mathcal{F}_k(\mathsf{x})$ are arbitrary functions of the variables vector x, included in the optimization (*e.g.*, gross weight, range, fuel consumption, life-cycle costs, etc.) with weights α_k, whereas α_{ss} represents the weight given in the optimization process to the target-sound matching.

V. NUMERICAL RESULTS

In this section, preliminary numerical results in including the Sound Similarity Index in the MDO process are presented. The following results have to be considered as an early example on the use of the present approach in multidisciplinary design optimization. Having in mind this, let us consider a New Large Aircraft of the A380 category (see Tab. II and III) in approach condition and flying over at an altitude of 3200 ft. As a noise source, we consider only the airframe noise due to the lifting system and to the vertical tail (see Fink [18]) and we move our analysis in the subspace of the design and procedural variables related to flight conditions and high-lift devices (only plain flaps are considered). Table IV shows the variables space under consideration.

The SSI between the spectrum of the configuration under analysis and a given target is included in the optimization process as the objective (see Section IV). The static equilibrium of the aircraft is used as a constraint. A Sequential Quadratic

TABLE II
DESIGN CONFIGURATION

number of seats	555
payload, kg	64,000
take-off weight, kg	339,000
empty weight, kg	167,000
design range, nm	5,000
cruise Mach number	0.8
cruise altitude, ft	30,000
overall length, m	73
fuselage diameter, m	7.16
number of engines	4
max thrust per engine, lb	70,000

TABLE III
WING AND TAIL CONFIGURATION

span, m	80.00
root chord, m	13.05
tip chord, m	2.00
root built-in angle of attack, deg	5.8
tip built-in angle of attack, deg	6.6
sweep angle, deg	26.4
wing area, m^2	583.19
horizontal tail span, m	21.51
horizontal tail area, m^2	238.56
vertical tail height, m	13.97
vertical tail area, m^2	114.83

TABLE IV
VARIABLE SPACE

Desing/procedural variables		lower b.	upper b.
flap span (as a ratio of wing span)	l_f/l_w	0.33	0.66
flap chord (as a ratio of wing chord)	c_f/c_w	0.10	0.30
flap deflection angle, deg	δ	0	60
trim angle of attack, deg	α	-5.0	20.0
speed, m/s	U	40	140

Programming (SQP) algorithm (see Refs. [19], [20]) in used to solve the constrained minimization problem.

As already mentioned, the noise emissions are computed using a simplified formulation which includes only the airframe noise and implements the Fink model [18]. The steady aerodynamic is calculated for cruise condition using the boundary integral formulation described in Subsection II-B (the grid used for computations is shown in Fig. 1) and the high-lift devices are taken into account by means of the section flap lift coefficient, c_{lf} (see Ref. [26])

$$c_{lf} := \frac{l_f}{q_{_D}\, c_f} = y_1(c_f/c_w)\, c_l + y_2(c_f/c_w)\, \delta; \qquad (22)$$

l_f is the section flap lift, $q_{_D}$ is the dynamic pressure, c_l is the unflapped section lift cofficient, δ is the flap deflection angle and y_1 and y_2 are known function of the ratio between flap chord c_f and wing chord c_w (see Ref. [26], pag.193).

First, we take as target sound an airframe noise as generated by the Fink model for a Boeing 747 in approach condition, and we evaluate the SSI in the optimization process using respectively $p = 1$ and $p \to \infty$. The final spectra are similar for the two cases and converge to the target airframe noise (see Fig. 2 and Tab. V). Moreover, we observe that the convergence to the final solution is faster when using the L^∞-norm.

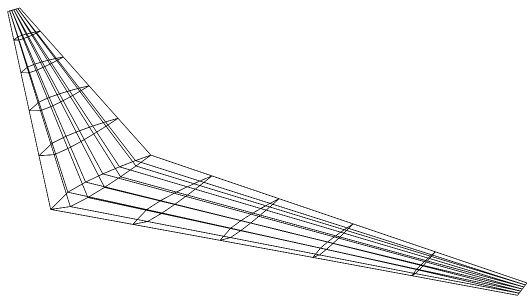


Fig. 1. The grid used for aerodynamic computation in cruise condition (wing)

TABLE V
MATCHING A TARGET AIRFRAME NOISE

design/procedural variables	initial conf.	final conf. $p = 1$	final conf. $p \to \infty$
flap span / wing span	0.66	0.51	0.65
flap chord / wing chord	0.30	0.12	0.13
flap deflection angle, deg	60.00	49.25	60.00
trim angle of attack, deg	20.00	-3.30	-5.00
speed, m/s	60.00	117.40	127.75
SSI in L^1, dB	2.83	0.06	-
SSI in L^∞, dB	9.08	-	0.22

We repeat the optimization procedure for a more complex target sound. Specifically, a spectrum generated within the SEFA project and including all the major noise sources (airframe, fan and compressor, jet, buzz-saw, and atmospheric attenuation) is taken as target. The optimization is repeated for $p = 1$ and $p \to \infty$ and the results are shown in Fig. 3 (see also Tab. VI). It may be noted that performing the minimization of the SSI in the L^1-space yields the best achievable spectrum in global shape or broad-band matching (the average of the absolute difference between the two spectra is minimized); the use of the L^∞-space yields the best solution with respect to local differences between the two spectra. In the present example, the two solutions give very close spectral shape (the spectra coincide within plotting accuracy). Figure 4 shows the convergence history of the SQP algorithm for $p = 1$ and $p \to \infty$. Again, we observe that using of the L^∞-norm makes the convergence to the final solution faster.

TABLE VI
MATCHING A TARGET SOUND

design/procedural variables	initial conf.	final conf. $p = 1$	final conf. $p \to \infty$
flap span / wing span	0.66	0.66	0.66
flap chord / wing chord	0.30	0.10	0.10
flap deflection angle, deg	60.00	57.84	42.28
trim angle of attack, deg	20,00	-4.55	-3.79
speed, m/s	60.00	140.00	140.00
SSI in L^1, dB	8.25	5.06	-
SSI in L^∞, dB	24.08	-	14.87

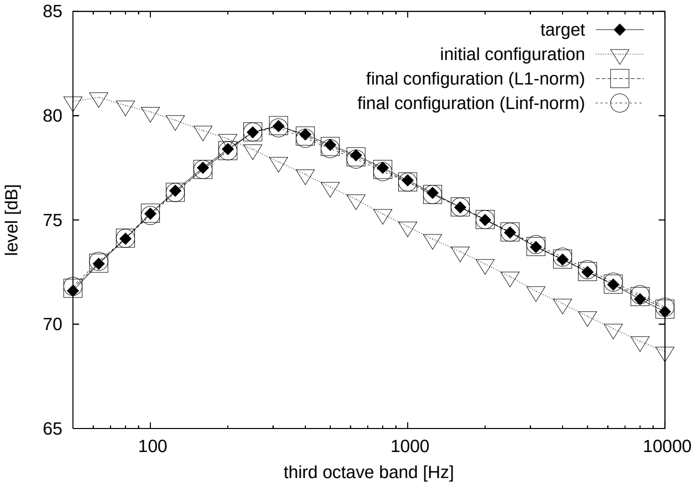


Fig. 2. Matching a target airframe noise

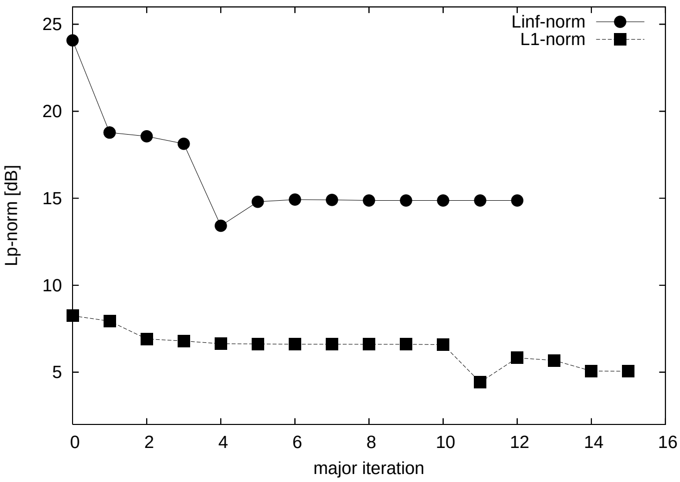


Fig. 4. Sequential Quadratic Programming. Convergence to the final solution

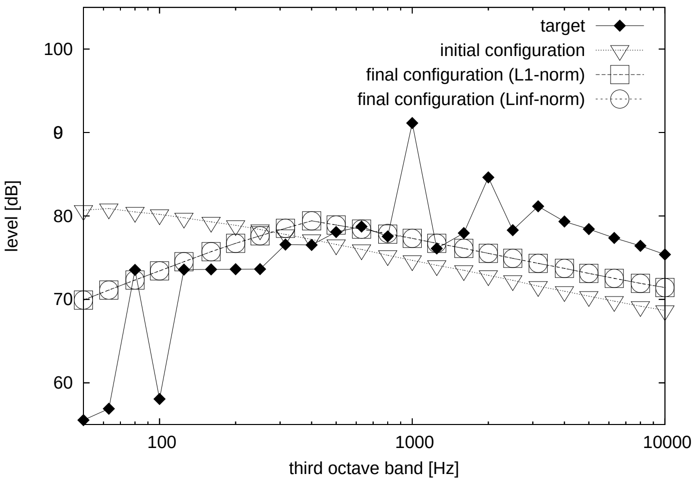


Fig. 3. Matching a target sound

VI. Conclusion

A novel approach for including community-noise considerations into an MDO framework has been presented. The attention is focused on the improvement of the quality of aircraft acoustic emissions. The objective is achieved by introducing, into the MDO procedure, a methodology to evaluate the similarity of the noise produced by the aircraft under analysis with a *target sound* properly defined. To do this, the Sound Similarity Index, $\mathcal{I}_{ss}$, is defined as the L^p-norm of the spectral difference between the two sounds, and introduced into the objective function to be minimized. Preliminary numerical results, obtained with a simplified formulation which includes only the airframe noise sources, show that $\mathcal{I}_{ss}$ represents an effective measure of the distance between the two sounds, and that efficiently leads the MDO procedure towards the aircraft configurations producing an acoustic emission satisfying the required matching to the target.

Acknowledgments

Part of this work has been funded by the European Commission within the project "Sound Engineering for Aircraft" (SEFA).

References

[1] Balling, R.J., Sobieszczanski-Sobieski, J., "Optimization of Coupled Systems: A Critical Overview of Approach," AIAA Journal, Vol. 34, No. 1, Jan. 1996.

[2] Balling, R.J., and Wilkinson, C.A., Execution of Multidisciplinary Design Optimization Approaches on Common Test Problems, AIAA Journal, Vol. 35, No. 1, 1997.

[3] Mastroddi, F., Vitale, A., Arsuffi, G., Morino, L., "Integrated Aerodynamic-Structural Constrained Optimization of a Wing Structure," proceedings of XIV Congresso Nazionale AIDAA, Napoli, Oct. 20-24, 1997, Vol. 3, pp. 1157-1168.

[4] Mastroddi, F., Ciancaleoni, E., Morino, L., "Aeroelastic Constraints in MDO," Proceedings of CEAS Internation Forum on Aeroelasticity and Structural Dynamics 1999, Williamsburg, VA, NASA/CP-1999-209136/PT1, pp. 33-42.

[5] Mastroddi, F., Morino, L., Bernardini, G., Simonetti, R., "MDO for Preliminary Design of a Highly Innovative Configuration," CEAS International Forum on Aeroelasticity and Structural Dynamics 2001, Madrid, 2001.

[6] Morino, L., Bernardini, G., Da Riz, W., Del Rio, V., "Aerodynamic Issues in MDO for Preliminary Design of an Innovative Configuration," Proceedings of IABEM 2002 (International Association for Boundary Element Methods), UT Austin, TX, USA, May 28-30, 2002.

[7] Mastroddi, F., Bonelli, C., Morino, L., Bernardini, G., "Multidisciplinary Design and Optimization for Fluid-Structure Interactions," Proceedings of 5th ASME Fluid-Structure Interaction, Aeroelasticity, Flow-Induced Vibration & Noise Symposium , New Orleans, Luisiana, USA, 2002.

[8] Morino, L., Bernardini, G., and Mastroddi, F., "First-principle based dynamic aircraft modeling for MDO," keynote lecture at ICCES 03 (International Conference on Computational and Experimental Engineering and Sciences), Corfu, Greece, July 24-29, 2003.

[9] Morino, M., Mastroddi, F., De Troia, R., Ghiringhelli, G. L., Mantegazza, P., "Matrix Fraction Approach for Finite-State Aerodynamic Modeling," AIAA Journal, Vol. 33. No. 4, April 1995, pp. 703-711.

[10] Iemma U., Diez M., "Optimal Life-cycle-costs Design of New Large Aircraft Including the Cost of Community Noise," International Conference on Computational & Experimental Engineering and Sciences, ICCES'05, invited lecture, Chennai, India, 2005.

[11] Iemma U., Diez M., Morino L., "Community Noise Impact on the Conceptual Design of Innovative Aircraft Configurations," AIAA Paper 2005-2982, Eleventh AIAA/CEAS Aeroacoustics Conference, Monterey, US, 2005.

[12] Morino, L., Iemma, U., Bernardini, G., Diez, M., "Community Noise Considerations in Multidisciplinary Optimization for Preliminary Design of Innovative Configurations," AIAA 2004-2809, 10th AIAA/CEAS Aeroacoustics Conference, Manchester, UK, 10-12 May 2004.

[13] Morino, L., Bernardini, G., De Gregorio, D., Willcox, K. E., Harris, W. L., "Life-Cycle Cost Multidisciplinary Optimization for Prandtl-Plane," Proceedings of the 2004 International Conference on Computational & Experimental Engineering & Sciences (ICCES'04), Madeira, Portugal, July 26-29, 2004.

[14] Morino, L., "A general theory of unsteady compressible potential aerodynamics," NASA CR-2464, 1974.

[15] Morino L, Gennaretti M., Iemma U., Salvatore F., "Aerodynamics and aeroacoustics of wings and rotors via BEM - unsteady, transonic, and viscous effect," Computational Mechanics, Vol. 21, pp. 265-275, 1998.

[16] Iemma U., Gennaretti M., "CHIEF Regularization Approach for Aeroacoustoelastic Modeling in State-space Format," AIAA Paper 2001-2276, 2001.

[17] Gennaretti M., Iemma U., "Aeroacoustoelasticity in State-Space Format using CHIEF Regularization," Journal of Fluids and Structures, vol. 17, n. 7, pp. 983-999, 2003.

[18] Fink, M. R., "Airframe Noise Prediction," Federal Aviation Administration, FAA-RD 77-29, 1977.

[19] Gill, P.E., Murray W., and Wright M. H., *Practical Optimization*, London, Academic Press, 1981.

[20] Gill, P. E., Murray W., Saunders M. A., "SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization," SIAM Journal on Optimization, Volume 12, Number 4, pp. 979-1006, 2002.

[21] Morino, L., Bernardini, G., "On Modeling for Innovative Configuration MDO with Recent Aerodynamics Developments," in A. Frediani and A. Miele (Eds.), Advanced Design Problems in Aerospace Engineering, Plenum Press, New York, NY, 2003.

[22] Morino, L., "Boundary Integral Equations in Aerodynamics," Appl. Mech. Rev., Vol. 46, 1993, pp. 445-466.

[23] Morino, L., Bernardini, G., "Singularities in BIE's for the Laplace Equation; Joukowski Trailing-Edge Conjecture Revisited," Journal of Engineering Analysis with Boundary Elements, Vol. 25, pp. 805-818, 2001.

[24] Lighthill, M.J. "On Displacement Thickness," J. Fluid Mech., Vol. 4, pp. 383-392, 1958.

[25] Morino, L., Salvatore, F., Gennaretti, M., "A Velocity Decomposition for Viscous Flows: Lighthill Equivalent-Source Method Revisited," in: Morino, L., and Wendland, W.L., (eds.): Boundary Integral Methods for Nonlinear Problems, pp. 161-166. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[26] Abbott, H. I., Von Doenhoff, E. A., *Theory of Wing Sections*, Dover Publications, New York, Ny, 1959.

[27] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, New York, USA: Dover Publications, INC., 1990.

[28] B. Dayanand Reddy, *Functional Analysis and Boundary-value Problems: an Introductory Treatment*, Harlow, England: Longman Scientific & Technical, 1986.

A Human in the Loop Experiment to Assess the Dual Airspace Concept

L. Guichard, S. Guibert, D. Dohy, JY. Grau

***Abstract* — The entire ATM community is highlighting the fact that the current air traffic management system cannot cope with the challenges of future air transport system. This paper proposes an original air traffic management system which will be able to cope with the peaks in traffic demand expected in the future. The *Dual Airspace* paradigm aims at increasing the en-route traffic by introducing a functional splitting of the traffic. This concept is based both on traffic pattern partitioning and the creation of two kinds of control unit: highways and sectors. This paper presents the first results obtained after a proof of concept experiment, assessing for an en-route sector the impact of a highway located inside its airspace. It represents an initial concept assessment, in terms of capacity and efficiency, while still maintaining a high safety level, allowing its operational applicability.**

***Index Terms* — Capacity, Dual Airspace, Highways**

I. Introduction

By taking into account the current and future features of ATM, the Paradigm SHIFT project [1], initiated in January 2004, proposes two major concepts as the skeletal structure for the shift of control paradigm: *Contract of Objectives* and *Dual Airspace*. The two concepts can be combined or considered independently as there is no contradiction in their mode of operations.

The *Dual Airspace* concept aims at increasing en-route traffic by introducing a functional splitting of the traffic, based on its features (i.e. climbing, descending, cruising). The objective is the cohabitation and sharing of the same geographical airspace area by two traffic management operational systems in order to cope with the demand peaks expected in the future. It is an innovative proposal for task sharing in ATM.

The *Dual Airspace* concept introduces a small number of continental highways conveying long haul cruise traffic complementing today's sector-based traffic (district). The purpose is to release pressure on local air navigation services by separating long-haul routes from the current routes.

L Guichard, S. Guibert, *EUROCONTROL Experimental Centre, rue des bois des Bordes, 91222 Brétigny sur Orge, France, Tel:+33 1 69 88 75 00, email:laurent.guichard@eurocontrol.int ; sandrine.guibert@eurocontrol.int*

D Dohy, J.Y. Grau, NeoSys, 7 rue du théâtre, 91300 Massy, France, Tel: +33 1 60 11 80 03, email: didier.dohy@neosys.fr, grau@synrjy.com

II. Concept

The *Dual Airspace* principle is founded on a traffic division not based on geographical adjustment (sector segmentation) but on functional separation of flows. The functional separation of flows in a common airspace is defined by:

- Independent and autonomous traffic flows, except in precise and well-defined exchange areas.
- Potentially different operational control modes for each flow.
- Total "compartmentalization" for air traffic controllers, between various flow management modes.

Dual airspace as defined in Paradigm SHIFT is characterized by:

- A limited number of highways on a continental scale, crossing the core area. Highways are located in the upper airspace and are defined in order to respond to the major traffic flows over-flying Europe and the core area. They do not link directly airport terminal areas (TMAs) to other TMAs. They receive aircraft from and hand them over to other airspace volumes. The aircraft using the highways are steady at the flight levels allocated to the highways. These levels are defined in accordance with a flight level allocation system (FLAS) in order to minimize interactions between highways and enhance safety.
- A route network similar to that currently in existence, the objectives of which will be to ensure the transfer of aircraft from TMAs to highways and vice versa, and to safeguard cruising traffic with no access to a highway. The route network will be divided into sectors, like today's airspace. Within the sectors, working methods will be similar to today's. Outside the highway cohabitation areas, sectors will be managed as they are today or as they might be in the future once new operational control methods are introduced.

III. Human in the Loop Experiment Objectives & Hypotheses

The aim of the "Human in the Loop" experiment which took place at the EUROCONTROL Experimental Centre on 21-28 October 2005 with four en-route controllers was to carry out an operational assessment of the concepts developed within the Paradigm SHIFT study on *Dual Airspace*. This study was limited to assess the impact on an en-route control sector of a highway passing through it. It was a local approach to the *Dual Airspace* concept, focusing on the gains to be made from the concept in terms of capacity and efficiency, while still maintaining a high level of safety. It represents an initial assessment of the concept.

The hypotheses of the experiment are as follows:

- Potential gains in capacity within a high-density control area.
- The preservation of a safety level compatible with air traffic operational targets.
- The efficiency of the control method in conjunction with the changes introduced by the presence of the highway within a control sector.

These will be achieved through the assessment of:

- The impact of the presence or not of a highway crossing the measured sector.
- The impact of the highway configuration in relation to the volume it occupies in the sector; this volume depends on the airspace occupied both horizontally and vertically.
- The impact, for a given traffic load, of the solutions adopted for the highway configuration levels in order to deal with meteorological turbulence (requiring aircraft within the highway to change flight level).
- The support tools and aids to facilitate the work of controllers in the sector;

IV. HUMAN IN THE LOOP EXPERIMENT VARIABLES

On the basis of the hypotheses listed above, the following variables were set or measured.

A. *INDEPENDENT VARIABLES*

1) *Existence or absence of a highway*

The highway represents an obstacle to traffic management in the sector it crosses. So the first independent variable will be the existence or not of a highway in the measured sector.

For this reason a reference scenario, without any highway, was created.

2) *Configuration of the highway*

In the vertical dimension, two configurations were assessed (Figure 1):

- Highway 1 (H1), where the highway occupies three distinct and separate flight levels, namely FL 370, FL 330 and FL 310.
- Highway 2 (H2), where the highway also occupies three flight levels, two of which are adjacent, namely FL 370, FL 330 and FL 320.

For the lateral dimension, a constant width of 22 NM (based on two tracks in each direction) was proposed for both configurations.

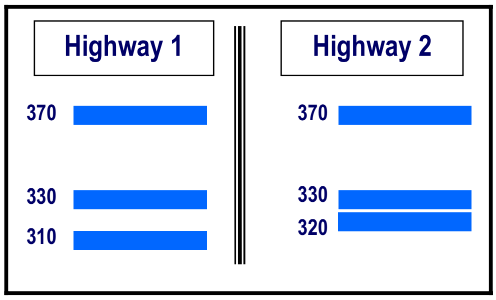


Figure 1: Vertical configurations of the highway

3) *Traffic density*

Two density levels were set: a medium traffic load and a heavy traffic load. These levels were determined according to the declared capacities of the operational sector chosen as a reference for the definition of the experimental airspace. They had to be adapted to the limitations on controllers' tasks caused by the experimental platform, i.e. on one hand the lack of coordination with adjacent sectors and on the other hand the fact that the interactions with aircraft were carried out solely via the radar interface with no radio communication. Moreover, instructions given to aircraft were carried out immediately and flawlessly by the latter. Consequently controller workload was reduced, necessitating an increase in traffic density to ensure equivalence with what happens in the operations room for medium and high traffic loads. Traffic adjustments were made and validated with two operational controllers.

Traffic load can be defined by two parameters: the number of aircraft in the measured sector per hour and the instantaneous number of aircraft in the sector. These two parameters reflect the difficulty involved in defining traffic, because of its variability. This is why the traffic used in the experiment was defined and rated on the basis of an average instantaneous traffic level (over a 15 minute period). The two traffic levels were therefore as follows:

- Medium traffic level: average instantaneous number of aircraft in the sector of 9.5 – 9.65.
- High traffic level: average instantaneous number of aircraft in the sector of 11.2 – 11.35.

One should note from these figures that the perception of traffic level is very sensitive to the number of aircraft: there is a threshold above which one quickly passes from a comfortable level (medium density) to a level requiring considerable cognitive resource mobilization (high density).

4) *Response to disruption due to turbulence*

The proposed system should be resistant to disruptive factors. Rather than entering into an exhaustive assessment of all the possible disruptive factors (refer to Paradigm SHIFT Operational Concept [1]), only the "meteorological turbulence" disruptive factor was assessed. It refers to meteorological phenomena which oblige all aircraft to request a change in flight level. This differs from stormy weather, which mostly requests a horizontal change to the flight path.

Since highway levels are pre-determined flight levels, any request from aircraft navigating in the highway system to change flight level because of turbulence must be feasible within the proposed system.

Two adjustment mechanisms for the vertical highway configuration were defined, each corresponding to a vertical hypothesis for the highway configuration (Figure 2):

- For Highway 1, it includes in the highway an extra flight level between FL 330 and FL 310. The highway thus occupies FL 370, FL 330, FL 320 and FL 310.
- For Highway 2, it involves bringing the two adjacent flight levels down two levels. The highway thus occupies FL 370 and FL 330 to FL 300.

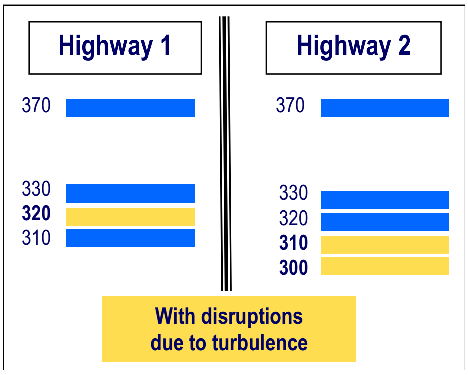


Figure 2: Highway turbulence-adjustment mechanisms

B. DEPENDENT VARIABLES

1) Dependent variables relating to controller activities

These variables quantify controller traffic management activities on the basis of independent variables monitored in the simulation exercises. The following dependent variables were measured: workload, traffic situational awareness or "picture", instructions given to aircraft by controllers, and lastly the controllers' opinion after they had carried out the simulation exercises.

a) Workload

The workload was assessed by means of a self-assessment questionnaire using a scale of 1 to 6. The questionnaire was based on NASA-TLX and SWAT questionnaires.

b) Situation awareness

Situational awareness was assessed by means of a self-assessment questionnaire using a scale of 1 to 6. The questionnaire was based on the concepts of situational awareness developed by Endsley [6], Leroux, Amalberti and Grau [7].

c) Control orders

The instructions given to aircraft are an objective reflection of the controllers' cognitive activity when managing the traffic. They do not represent all the aspects of this cognitive activity since a controller's understanding, reasoning and decisions are not necessarily all translated into by an instruction given to an aircraft. However, they make it possible to assess the decisions taken to:

- Organize the traffic and make it compatible with the cognitive resources available in line with the constraints of the current and expected situation (both flight plan and sector constraints).
- Resolve conflicts.
- Manage avoidance of the highway.

To quantify the given orders, an "instructions given" parameter was defined which corresponded to the average number of instructions given by the controller to each aircraft for each scenario phase corresponding to the setting of monitored variables.

d) Post-simulation interviews

The post-simulation interviews allowed us to find out the controllers' views on their activities. They were recorded and a content analysis was carried out. Lasting around one hour per controller, the interviews were carried out using open and closed questions and a semi-structured interview format.

2) Dependent variables relating to safety

Safety became an issue in the experiment where there was a loss of separation between aircraft and/or where an aircraft entered highway airspace. To avoid this and warn the controller before one or more aircraft became a safety risk, the control position was equipped with two alarms, a short term conflict alert (STCA), and an area infringement warning (AIW) which identifies the imminent intrusion of an aircraft into the highway envelope.

The activation of these alarms was considered a safety issue even though there remained sufficient time for the controller to act immediately to avoid loss of separation. In all cases, the alarms reflected a lack of anticipation on the controller's part in managing aircraft flight paths in relation to other aircraft or the obstacle presented by the highway. This is why all the alarms (STCAs and AIWs) which were triggered were accounted for and recorded for each scenario phase corresponding to the setting of monitored variables. However, the triggering of these alarms is the result of an algorithmic calculation which does not take controllers' intentions into account. This is why the alarm may have been triggered in certain circumstances even though the controller was aware of the situation and was capably handling the separation between aircraft and/or with the highway. To deal with this limitation, all the alarms were reclassified according to the context (controller's opinion and observations of the person running the experiment) and only "relevant" alarms were taken into account.

There are two types of safety data:

- The number of times the STCA alarm was triggered.
- The number of times the AIW alarm was triggered.

3) Dependent variables relating to performance

Although safety was the dependent variable of primary interest, air traffic control performance was also assessed through capacity, stability and efficiency criteria.

a) Capacity

Capacity assessment is based on the ability to manage the different traffic used for the experiment (Traffic density A.3). Since the traffic level calculation takes into account the limitations of the controllers' tasks, the experiment's conclusions on capacity are expressed in relative and not absolute terms.

b) Stability

Stability is characterized by the number of instructions given by control to aircraft in the sector. This dimension is the same as that defined to quantify controller activities (Given orders B.1.c).

c) Efficiency

The term "efficiency" denotes the traffic variability induced by the sector, its route network and the working methods used by controllers to manage the traffic. Knowledge of this variability makes it possible to quantify the impact on traffic of the highway passing through the sector. This is important in order to determine how variability may be acceptably allocated within the network, especially in the context of the Contract of Objectives as it is envisaged in the Paradigm SHIFT project.

One way of quantifying this variability is compliance with aircrafts' scheduled flight plan parameters on leaving the sector. The efficiency index takes account of divergences from predictions made when aircraft enter the sector concerning the duration and the distance needed to travel through it. These divergences are calculated for each aircraft and expressed in the form of an average time and distance per aircraft.

In order to assess control efficiency, the following data were gathered for each aircraft:

- Average time difference between the flight duration and the predicted flight duration in the sector.
- Average distance difference between the real distance flown and the planned distance flown in the sector.

V. CONDUCT OF THE EXPERIMENT

The experiment took place over six days with four en-route controllers from three air navigation operational centers: one Czech controller, one Slovenian controller and two French controllers.

The airspace and traffic scenarios were designed on the basis of a real-life situation. However, so that all the controllers began on an equal footing in terms of any prior knowledge of the circumstances, the airspace and traffic were changed to make them anonymous.

For each controller, the experiment was conducted as follows:

- One day to present the concept, the experiment, the platform and training with scenarios of the same difficulty as those to be measured.
- One day to carry out three 70-minutes scenarios, each followed by a post-simulation interview.

A. Structure of the scenarios

The scenarios carried out by the controllers were built around the idea of independent variables testing the experiment's hypotheses. Each scenario was built around five phases:

- Phase 1 – warm up, to start the scenario and put the controller in situation (15 minutes).
- Phases 2, 3 and 4 – the measured phases (15 minutes each).
- Phase 5 – finishing the exercise (10 minutes).

Phases 1 and 5 were not taken into account in the results analysis.

In all scenarios 70% of traffic was cruising and 30% climbing and descending.

B. The three experimental scenarios

The three scenarios were as follows: Reference scenario (R), Highway 1 scenario (H1) and Highway 2 scenario (H2).

Phases 2 and 3 have a common purpose for the three scenarios and were designed the same way:

- Phase 2 with a high-density traffic load.
- Phase 3 with a medium-density traffic load.

1) Reference Scenario (R)

The R scenario has no highway. The last measured phase (i.e. Phase 4) consisted of medium-density traffic load and turbulence.

2) Highway 1 Scenario (H1)

The H1 scenario introduces Highway 1. The last measured phase consisted of medium-density traffic and turbulence necessitating the application of the H1 adjustment (vertical enlargement of the highway at FL 320).

3) Highway 2 Scenario (H2)

The H2 scenario introduces Highway 2. The last measured phase consisted of medium-density traffic and turbulence necessitating the application of the H2 adjustment (vertical enlargement of the highway at FL 300 and FL 310).

C. Tools to facilitate the work of controllers in managing vertical constraints

Even if it’s not the goal of this experiment, as for others projects interested in highways’ displays [8][9], an appropriate assessment of the impact of an highway on a sector couldn’t be made without giving specific tools to the ATCOs to deal with and to be able to perform their tasks.

So, four specific tools were developed for this experiment:

- A highway level filter to display aircraft on the same level as the highway flight levels.
- An alarm to signal imminent incursion of an aircraft into highway airspace.
- A dynamic vertical representation of the overall traffic

situation in the sector.

- A static representation of the vertical flight path of a selected aircraft and its potential vertical maneuver shape based on its performance.

D. Automated management of vertical constraints linked to penetration of the highway

The vertical constraints imposed by highway crossings generate additional workload for the controller. Since these constraints are fixed and permanent, it has been suggested that they could be managed automatically and systematically in the flight plan. It would therefore be left to the aircraft to respect the constraints automatically where programmed into the aircraft's flight management system. In the experiment, constraints were declared and respected in the flight plan and the process was therefore automatic. The integration of the vertical constraints into the flight plan meant that the controller did not need to take any action with regard to aircraft flight plans; no aircraft could penetrate the airspace at the entry point of the highway.

The corollary of this choice is that aircraft spontaneously climb or descend in the vertical plane without receiving control instructions but in conformity with the flight plan. This means that the controller must memorize the planned aircraft movements to avoid being surprised when they happen. Similarly, he or she must integrate these movements in the event of an aircraft flight path change for traffic management or conflict resolution purposes, since these flight plans are not conflict-free.

VI. RESULTS

A. Criteria for analyzing the data collected

1) Average values

The results obtained from the four controllers enabled trends and approaches to be identified on the basis of the confirmation or rejection of the hypotheses used in the experiment. The data obtained cannot be processed statistically and at this stage of the investigations, to do so would be meaningless.

This paper will present only the average values for all controllers (Figure 3)

		Phase 2	Phase 3	Phase 4
R	Average instant capacity	11.25	9.65	9.60
	Workload	2.75	2.00	2.00
	Picture	3.00	1.50	1.75
	STCA	0.75	0.00	0,50
	AIW	0.00	0.00	0.00
	ATCO's orders	1.29	1.32	1.82
	Delta time(s)	-2.75	-1.50	1.00
	Delta dist. (1/10 NM)	-2.50	-2.00	1.50
H1	Average instant capacity	11.20	9.50	9.70
	Workload	3.00	2.25	2.50
	Picture	3.25	1.75	1.50
	STCA	0.50	0.25	0.50
	AIW	0.00	0.00	0.00
	ATCO's orders	1.76	1.76	2.06
	Delta time(s)	-3.75	0.75	-0.25
	Delta dist. (1/10 NM)	-2.75	0.00	0.00
H2	Average instant capacity	9.60	11.35	9.85
	Workload	1.75	3.00	3.25
	Picture	1.50	2.75	3.50
	STCA	0.00	0.75	1.00
	AIW	0.00	0.00	0.50
	ATCO's orders	1.28	1.85	2.16
	Delta time(s)	-3.00	-3.50	-2.50
	Delta dist. (1/10 NM)	-4.50	-5.75	-3.75

High traffic load

Medium traffic load

Medium traffic load with

Figure 3: Summary of the average values for the data gathered in the experiment

2) Subjective assessments (workload and picture)

The small number of controllers meant that significance tests could not be carried out on the results obtained. It was therefore considered that a difference existed between the analyzed results when it exceeds 0.75 (15% for the workload and picture value). This figure may seem high but it takes into account the low number of controllers and the consequent variability of the results while also making it easier to judge whether there is a similarity or discrepancy between the data obtained for the "outliers". We are aware that using this method meant that the analysis criteria were fairly crude but they were nevertheless compatible with the experiment objectives, i.e. to identify strong trends in order to continue developing the concept.

3) Given orders (ATCO's orders)

The unit defined for instructions given is the average number of instructions given per aircraft. Since we had no reference by which to qualify controller activity using this value in relation to traffic density or complexity, we arbitrarily decided that a difference of 1 was significant.

4) Safety criteria (STCA and AIW)

The selected safety criteria were the triggering of alarms. These alarms leave time to react in order not to infringe the minimum separations between aircraft or the minimum distances from the highway. This means that alarms are not exceptional events, even if they represent a lack of anticipation on the controller's part. Moreover, we acknowledge that there is a "simulation" effect on the number of alarms observed in the sense that the alarms are triggered more regularly owing to reduced involvement of the controllers in traffic management because a simulation can never reproduce the real context [3].

This is why a minimal differential value of 0.50 was retained as a safety significance criterion for all controllers. It corresponds to a difference of two alarms for all four controllers.

5) Traffic management efficiency criteria (delta time and distance)

As with the instructions given, it was proposed to quantify traffic efficiency using units defined for this experiment, namely the average difference in sector-crossing time and distance for each aircraft between what was scheduled in planning and the actual time/distance flown during the scenario. Since we had no experience on how to use or value this data, we deliberately refrained from defining significance criteria.

B. Airspace and air traffic density

1) Medium traffic density

With medium traffic density (Figure 4), there is no difference between the three scenarios for the activity carried out by the controllers (workload, number of orders given), the safety level achieved or the traffic management performance.

This means that the highway has no impact on traffic management in the sector, whichever highway configuration is envisaged (H1 or H2).

	Ref	Highway1	Highway2
Workload	2.00	2.25	1.75
ATCO instructions	1.32	1.76	1.28
Picture	1.50	1.75	1.50
STCA	0.00	0.25	0.00
AIW	---	0.00	0.00
Delta time(s)	-1.50	0.75	-3.00
Delta dist. (1/10 NM)	-2.00	0.00	-4.50

Figure 4: Comparison of medium-density airspace

2) High traffic density

For high traffic density (Figure 5), there was a change in the variables in all scenarios, reflecting controller activity in relation to medium traffic density. This change took the form of an increase in perceived workload and reduced traffic situational awareness. The latter did not, however, affect safety. The change in situational awareness seems more significant than that in workload. However, there was no change in the average number of orders given to aircraft.

As regards safety, more STCA events were triggered in the three scenarios. This increase is completely consistent with the perceived increase in workload and perceived decrease in situational awareness. The consistency between these three variables shows that the traffic levels in the scenarios were such that controllers had to apply safety strategies. Above all, the controllers no longer had adequate cognitive resources to apply "elegant" solutions [4]. However, these results also show that they never reached the point of breakdown, entailing loss of control of the situation and thus loss of safety.

It should also be noted that the changes observed were similar for the three scenarios, from which we can conclude that there is no difference between the reference scenario and either of the highway scenarios.

As regards safety, the only alarms observed were those concerning aircraft separation, regardless of the scenario and the traffic density. No alarm was triggered signaling loss of separation with the highway. This result is promising within the framework of the concept. However, this shows that the controllers were able to manage conflict avoidance between aircraft in the context of the constraints imposed by the highway.

	Ref	Highway1	Highway2
Workload	2.75	3.00	3.00
ATCO instructions	1.29	1.76	1.85
Picture	3.00	3.25	2.75
STCA	0.75	0.50	0.75
AIW	---	0.00	0.00
Delta time(s)	-2.75	-3.75	-3.50
Delta dist. (1/10 NM)	-2.50	-2.75	-5.75

Figure 5: Comparison of high-density airspace

3) Traffic management strategy

The “average number of orders given” variable leads us to conclude that no difference can be observed between the scenarios, even with different traffic densities. In fact, observing traffic management strategies shows that the controllers were using different strategies from one another and also from one scenario to another. The terms “strategy” and “traffic management” are used to mean interventionism or a certain degree of anticipation in the instructions given. We can thus observe highly anticipatory and interventionist strategies where the controller structures the traffic very early on as he/she sees fit on the basis of future constraints or wait-and-see strategies where the controller prefers to wait until the last moment before giving instructions only when they are essential. This variability in behavior shows that the "average number of instructions given" variable depends more on controller behavior than on traffic density or airspace structure factors (with or without a highway).

This variability in traffic management strategies affects the analysis of the traffic efficiency criteria which we defined. In fact, the aircraft sector-crossing times and distances are also linked to the traffic management strategies used by the controllers. Thus, as shown by the negative times in the reference scenario, the controllers instinctively use direct routes which reduce the average crossing time and therefore the distances traveled. The tendency to use direct routes further intensifies in periods of high traffic density, as shown in the results. Moreover, in high traffic density the crossing times and distances are lower in the two highway scenarios than in the reference scenario.

In order to explain these results we can suggest that in medium traffic density the controllers have the cognitive resources allowing them to apply strategies without seeking a genuine saving or to optimize these resources. In high traffic density, the need for savings is more urgent because we are approaching the controllers' maximum capacity. In this case we observe the implementation of strategies favoring whatever is most efficient in terms of safe traffic management.

In the medium-density scenarios the controllers can adopt whatever strategies they wish, which might explain the increase in the crossing times and distances observed in the H1 scenario. In the high-density scenarios the controllers strive to perform their activities as efficiently as possible. They use the simplest solutions and direct routes are part of this. This is why we observe more significant reductions in time and distance for the highway scenarios.

Although these results are consistent with a cognitive logic of traffic management, they nevertheless differ from what we were expecting. Indeed, they show that the denser or more complex the traffic - i.e. where the constraints on controllers are tighter - the shorter the crossing times and distances. We might have expected traffic complexity to necessitate less fluid management strategies penalizing the traffic, but this is not the case judging from our observations in the experiment.

4) Dual Airspace and capacity

The results obtained from the medium and high-density scenarios were particularly interesting for the increase-in-capacity hypothesis. In fact, even if we note a difference in the controllers' activity and traffic safety between the medium and high-density scenarios, these differences do not demonstrate an effect linked to the sector structure. This means that the highway, regardless of the configuration under assessment, does not reduce sector capacity in comparison with a sector without a highway. Thus the traffic carried by the highway would represent a net increase in the total traffic managed within the geographical area of the sector.

It is thus possible to maintain the same traffic capacity in a sector with a highway as in a sector without a highway. Thus the traffic carried by a highway represents an extra gain within given geographical area.

C. Configuration of the Highway and turbulence

The highway must be resistant to disruptive factors including those linked to turbulence, which forces aircraft to change their flight level. For each structure type a way of vertically expanding the highway was assessed. For H1 (see Figure 2), expansion meant the use of intermediate flight levels for descending and climbing between the two highway flight levels. For H2, expansion meant adding two adjacent levels below the existing highway levels.

1) H1 configuration

In the light of the data collected on medium-density traffic we can see that there is no difference for the H1 highway between a situation with and one without activation of the airlock (Figure 6). Whether in terms of controller activity, the safety level achieved or the traffic management performance, similar values are noted for the two experimental conditions.

	Normal	**Turbulence**
Workload	2.25	2.50
ATCO instructions	1.76	2.06
Picture	1.75	1.50
STCA	0.25	0.50
AIW	0.00	0.00
Delta time(s)	0.75	-0.25
Delta dist. (1/10	0.00	0.00

Figure 6: Highway H1 turbulence-adjustment mechanism

2) H2 configuration

Unlike for H1, with H2 we can see a significant change in the values for the controller activity and the safety level achieved upon activation of the two flight levels (Figure 7). Where there is medium traffic density, equivalent or even slightly higher values are observed than those for the controller activity in high-density traffic. As regards safety, a significant increase can be seen in the number of STCA events triggered, exceeding the number without activation albeit in a high-density traffic situation, and above all we note that AIW events were triggered, which was not the case in any of the other scenario phases.

	Normal	**Turbulence**
Workload	1.75	***3.25***
ATCO instructions	1.28	***2.16***
Picture	1.50	***3.50***
STCA	0.00	***1.00***
AIW	0.00	***0.50***
Delta time(s)	-3.00	-2.50
Delta dist. (1/10	-4.50	-3.75

Figure 7: Highway H2 turbulence-adjustment mechanism

3) Summary of the two adjustment mechanisms

Clearly, adding two flight levels and thus blocking four adjacent flight levels creates a wall which disrupts controller activity. In the H1 hypothesis only three adjacent flight levels are blocked, which seems much more acceptable to controllers.

This may seem like stating the obvious, since blocking four levels is clearly a more penalizing measure than blocking three. Beyond this simple observation, however, it is especially interesting to see that a limit seems to exist in terms of blocked flight levels after which the control task is objectively penalized

Beyond the fact that the H2 solution seems problematic in the event of adjustments for turbulence, which is not the case for the H1 solution, the comparison between the two experimental situations makes it possible to set a maximum number of adjacent flight levels operationally acceptable to controllers. This figure of three is important to take into account when determining the height of the highways but also the highway adjustment mechanisms in order to deal with disruptive factors.

The average number of instructions given to each aircraft is slightly higher in the adjustment situation than under normal conditions. This is entirely logical and represents the additional instructions given by controllers in order to avoid the new flight levels which had not been declared in the flight plans.

Regarding traffic management efficiency, it is difficult irrespective of the highway configuration to identify any particular trend between the situations.

D. Highway and working methods

On the basis of the interviews conducted with all the controllers following the simulation and an assessment of the data, the consensus is that the presence of a highway in a sector does not change current control methods. The controllers acknowledged that they managed the aircraft in the same way as they currently manage them in the control room, and this posed no particular problem with regard to the constraints imposed by the highway. They made the same observation during the H1 and H2 scenario phases when the highway adjustment mechanisms for turbulence were activated.

However, the automatic management of aircraft vertical constraints linked to the penetration of the highway was strongly criticized by the controllers. Rather than help them, automatic management of vertical constraints increased their workload, since it was difficult to know what each aircraft would do and, above all, when and how it would do it. They unanimously preferred to control the vertical movements of aircraft themselves on the basis of their own perception of the traffic. The controllers are not, however, against traffic planning which takes highway-crossing into account. This means that vertical instructions may be introduced into the flight plan. These instructions would serve as a framework for the controller but it would always be the controller who gives the instructions. By giving these instructions, he/she is sure of aircraft movements and, in particular, of controlling future flight paths thus facilitating their memorization.

The controllers quickly integrated the flight-level constraints on leaving the sector into the management of the vertical constraints linked to the highway. They thus adopted strategies early on to avoid flight levels occupied by the highway, taking into account the exit flight levels (XFL). A corollary of this attitude is the desire they expressed for aircraft flight plans to be as consistent as possible in terms of the highway vertical constraints and constraints on leaving the sector so as to facilitate both highway-crossing and to simplify coordination with the adjacent sectors.

E. Control support tools and human-machine interface

1) Vertical representation of the traffic

This display, located to the bottom left of the radar screen, was considered useless by the controllers. The first criticism made against it was that it was too small to allow a clear picture of the traffic, even if the scroll bars to zoom in on and centre the image were very intuitive and useful. Moreover, a horizontal rather than a vertical display would better warrant a tool for analogue visual identification of flight levels in relation to one another. Horizontal lines marking the reference flight levels seem to be essential. Although the logic adopted for the projection in the vertical plane was considered the only viable one (projection axis perpendicular to the highway axis), the controllers did not use this tool because they were perfectly able to construct a vertical representation of the traffic without additional support. Vertical representation or, more to the point, spatial representation of the traffic requires know-how which is developed and acquired through training and experience. It is constructed using 2D analogue radar data and the alphanumerical information presented on the radar and/or paper strips. Since they already have this know-how, a 2D vertical representation does not help the controllers in any way. They have this aptitude: it is one of their skills and to be frank they do not understand how anyone could doubt their ability to construct a correct spatial representation of the traffic.

2) Aircraft-centered vertical representation

This display, located to the bottom right of the radar screen, was hardly used. Initially thought to be on a par with the vertical representation of the traffic, i.e. useless, the controllers revised their opinion because it allowed an aircraft's chosen vertical flight path to be displayed and therefore made it possible to anticipate the aircraft's future position in relation to the highway. This function was seen as useful for organizing the traffic entering the sector, in particular the aircraft’s vertical constraints were entirely managed by controllers, as they would like. Used in this way, the display could have two functions:

- Assessing the correct crossing of the highway on the basis of the flight plan.
- Helping the controller give instructions for movements in the vertical plane, on the basis of an aircraft's performance as shown in the display.

These functions were not implemented as part of the experiment and should be the subject of further study in this specific area.

3) Highway-level filter

The level-filter linked to the levels occupied by the highway was unanimously seen to be particularly useful for anticipating flight-level problems arising from crossing the highway. Used for entry to the sector, it helps the controller construct a spatial representation of the traffic. In fact, it acts more to back up the controller's own spatial representation than really to help construct one. The level filter allows the controller to check that he/she has properly understood the traffic, not in fact all the traffic but only the potential traffic problems caused by highway-crossing. The complexity of the level-filter tool depends on the extent to which the flight plan data is integrated. Either it functions according to the aircraft's flight level when the filter is activated, which is already satisfactory for the controllers, or it integrates the flight plan and functions according to the aircraft's anticipated flight level when it crosses the highway.

4) Area Infringement Warning (AIW)

This tool is very similar to STCA. Like STCA it acts as a "safety net" and as such was totally endorsed by the controllers.

5) Human-machine interface

The main comment concerned the visual display of the highway in the radar image. Even if what was presented was entirely acceptable, the controllers would have liked the highway to be represented in a more "alarming" way and requested, in particular, a reminder on the radar display of the flight levels occupied by the highway.

F. Airspace

The Human in the Loop experiment made it possible to confront controllers with the sector's design rules and those of the highway which crosses it.

1) Sector

The size of the sector as proposed in the experiment (80 NM x 60 NM – FL 290 to FL 430) was deemed satisfactory by the controllers, although two of them would have preferred a larger sector to allow better anticipation of traffic management and to reduce coordination with adjacent sectors.

The choice of the size and configuration of the sector was the result of a compromise which took into account:

- The width of the highway.
- The density and nature of the traffic crossing the sector.
- The position of the convergence points in the sector in relation to the highway, in order to manage highway-crossing.
- The position of route convergence points in relation to the sector entry point, to anticipate as far as possible the management of potential conflicts.

The main comment from controllers concerned the lack of flight level parity for routes in the sector. Parity would have made their work easier regarding the proposed convergence points.

The most contentious issue was the positioning of the highway in the sector. The ideal airspace according to the controllers would include:

- A central highway.
- Convergence points at a minimum distance of 20 to 25 NM from the highway in order to anticipate highway-crossing. This implies that there should be no route parallel to the highway at less than 20 NM distance.
- No route crossing points within the geographical area of the highway.
- No converging routes at less than 20 NM from the sector entry point.
- Routes with equal status.

The problem with this ideal airspace is its size, which, in high-density traffic areas, is no longer compatible with controller resources in terms of capacity. It emerged from the interviews that it was difficult to define a strict framework for designing sector airspace. Each sector will be a special case and will need to be optimized on the basis of local constraints. The solutions assessed in the experiment are relevant and must be adapted to the local context.

Alongside this sector-by-sector approach and specific sector design, it was unanimously acknowledged that the highway must in no way be shared by adjacent sectors owing to the risk of confusion on handover and controller responsibility for aircraft.

Another intermediate solution would be to situate the highway at the edge of the sector. This would make it possible to liberate as much airspace as possible in the sector to anticipate the crossing of the highway by traffic flows coming from the opposite side of the sector to that on which the highway is positioned. On the contrary, for flows arriving from the sector next to the highway, any anticipation and management of traffic crossing the highway would therefore have to be done by the entering sector. The controllers think that this solution may be feasible since very precise rules exist for the sharing of tasks and responsibilities between the various operators.

To summarize the issue of sector airspace, we can say that this experiment made it possible to better identify operational needs and thus define design rules. However, it is clear that beyond the local level of acceptability assessed here, airspace can only be envisaged and defined at a global level. Although airspace has to accommodate the application of the *Dual Airspace* concept, it must above all respond to air traffic demand with all the related constraints on punctuality and cost-effectiveness. On the continental scale, new constraints would certainly appear that are impossible to perceive at this stage.

Building on the experiment results, subsequent steps for developing design rules for *Dual Airspace* at sector level should now be the subject of specific studies on airspace at continental level and be based on design models.

2) Highway

The controllers involved in the experiment stated that the highway should occupy as little space as possible in the sector. However, the highway must be big enough to accommodate the traffic which it carries.

In the horizontal plane, the design rules used for the highway (width of 22 NM) have been accepted and leave almost no room for maneuver to reduce the width of the highway. Moreover, all the controllers found this width to be entirely acceptable, and not detrimental to efficient traffic management in the sector. Admittedly, finding an acceptable width for the highway also depends on the sector's size, as detailed in the previous section. This also shows that an operational and relevant *Dual Airspace* concept can only be designed at a more global level. It is at this level that assessments must be made, with all the peculiarities that will arise. From this perspective, the experiment that was carried out was an initial step, indispensable but not sufficient for the process of designing *Dual Airspace*.

In the vertical plane it is also necessary to determine the number of flight levels occupied by the highway as a function of traffic demand. The decisions made in the experiment in both the vertical and horizontal planes are deliberately focused on capacity. On the basis of more refined studies of potential traffic demand scenarios, it would be entirely possible to adapt the number of flight levels.

In the context of the experiment, the distribution of flight levels in the two highway configurations was considered acceptable by the controllers involved. Nevertheless, in the analysis of the two hypotheses H1 and H2, opinion was divided. For some, it would be preferable for traffic to have highway flight levels which alternate with the sector flight levels, as is the case in H1. This imposes fewer constraints on the management of traffic crossing through by not creating a wall and by limiting the scope of the instructions which may be given. Others felt that a single free flight level between the flight levels occupied by the highway (H1) was insufficient because it was too restrictive and not very practical at an operational level. This flight level is trapped between the other two and is difficult to work with.

In the light of these results, blocking three adjacent flight levels was acceptable to the controllers, but this was not the case with four flight levels. This means that any adjustment to the dual system to deal with turbulence must not exceed three adjacent flight levels. If this cannot be done as part of a limited adjustment to these three flight levels, additional adjustment solutions must be envisaged, probably on a more global scale, such as action taken on traffic organization and density.

3) Independent sector and highway

In the definition of the *Dual Airspace* concept, sector and highway are completely hermetic which means that the controllers of each system do not see the traffic in the other system. This was the hypothesis assessed in the experiment.

The controllers consented to this way of working but drew attention to degraded or critical situations (emergency descent for example). These cases were represented in the operational concept of *Dual Airspace* in the following way: as soon as an aircraft declares itself to be in a degraded situation (transponder switched to Emergency), it displays a transponder code which makes it visible in airspace other than that in which it finds itself but with which it may have to interact. From this moment on, the controllers of each airspace volume concerned can make use of communication means (datalinks, radio communication, etc.) suitable for optimizing the management of this aircraft depending on its intentions. Faced with this kind of situation we move away from a nominal operating framework to functioning in emergency mode. This operating mode was defined in order to take into account firstly the controller's workload in a nominal situation and secondly the controller's area of responsibility. It should be remembered that under the current legal framework a controller cannot be deemed responsible for aircraft which are not visible to her [5].

During the role-plays in the experiment scenarios, even where this was not one of the experiment's objectives, it emerged that some controllers would like to be able to visualize traffic in airspace which they are not controlling. To this end, they suggest display on request rather than permanent display of information. They would prefer this solution to the hypothesis proposed in the concept, because they feel that an overview of traffic in other airspace is important in order to best manage the emergency situation. This alternative will require on the one hand an examination from a legal point of view of whether a controller can be deemed not responsible when displaying information from another airspace, and on the other hand a clear definition of what we mean by an emergency situation and who establishes criteria defining such a situation. These hypotheses must be assessed in future developments of the *Dual Airspace* concept

VII. CONCLUSION & FURTHER STUDY

The general feeling about the concept is positive. Both the opinions of the controllers and the analysis of the data obtained indicate that the Dual Airspace concept presents a potential axis for increasing capacity in geographical areas with high traffic density. For controllers, therefore, the concept is perfectly "workable" and deserves to be developed in this type of experiment. To support this general feeling, we can summarize the results of the experiment via an analysis of the experimental hypotheses:

- The presence of a properly designed highway within a sector, in whatever configuration (H1 or H2), does not affect capacity, safety or efficiency in comparison with the same sector without a highway.
- The presence of the highway does not affect the controllers' workload or their understanding of the traffic.
- A sector's capacity is maintained when the sector is crossed by a highway. Therefore, if we add the sector capacity to that which can be managed in the highway, it is possible to increase the capacity of a single geographical area by using the *Dual Airspace* concept. This is possible because the principle of traffic sharing between controllers is no longer geographical but based on a functional separation of traffic.
- When the highway is modified to deal with disruptive factors such as turbulence, the vertical wall created should not exceed three adjacent flight levels. The adjustment solution assessed via the H2 highway configuration, occupying four adjacent flight levels,

proved too problematic in terms of safety and workload and was therefore unacceptable.

- The hypotheses for the design of the highway and sector are a coherent foundation for continuing work on the *Dual Airspace* concept.
- The working methods for traffic control and management are not fundamentally different from those currently in use. The controllers unanimously rejected the principle of automatic management by aircraft of the vertical constraints linked to the highway. They would prefer complete control of the vertical constraints in order to ensure optimum management of traffic anticipation and organization.
- Concerning the support tools proposed to controllers, only the AIW and the highway-level filter were considered of operational interest. The vertical representation of traffic is irrelevant and the aircraft-centered vertical representation could be useful in the framework of complete management by controllers of the vertical constraints.

To sum up, we can therefore conclude that the answers provided to the questions posed on *Dual Airspace* as assessed in the "proof of concept" experiment enable us to state that the concept is operationally acceptable to controllers and that it allows the capacity of a geographical area with high traffic density to be safely increased.

Even if these initial results confirm the relevance of the *Dual Airspace* concept and the steps taken to define it, they are only a first step in the validation process. They demonstrate that the concept deserves further examination. Equally, the issues tackled in this experiment do not cover all aspects of the concept. It would therefore be useful, following this experiment, to:

- Develop concepts for aircraft entry/exit between the sector and the highway.
- Define the functioning of a highway.
- Refine the definition of the rules governing airspace design, working methods and support tools.
- Define at a continental level, for both the highway and the sectors, an airspace structure which integrates local constraints.
- Lastly, assess the *Dual Airspace* concept using a real-time, full-scale simulation platform in order to integrate operational consistency between sectors, between sector and highway and with aircraft.

Throughout this definition and validation process, it is important to stress the leading role to be played by the operational controllers. They must be included as full partners in the various stages of the process.

REFERENCES

[1] L. Guichard, S. Guibert, H. Hering, D. Dohy, J.Y. Grau, J. Nobel, K. Belahcene; *« Paradigm SHIFT » Concept Opérationnel*, EEC Note No. 01/05

[2] L. Guichard, S. Guibert, H. Hering, D. Dohy, J.Y. Grau, J. Nobel, K. Belahcene *« Paradigm SHIFT » Research Agenda*, EEC Note No. 16/05.

[3] Dubey, G. (2000). *Social factors in air traffic control simulation*. EEC Report n° 348, July 2000.

[4] Sperandio, J.C. (1984*). L'ergonomie du travail mental.* Masson, Paris

[5] J, Laborde dit Bouriat *Impact de la Responsabilité juridique dans le domaine du contrôle du trafic aérien* EEC Note No 06/2005

[6] Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32-64

[7] Grau J.Y, Menu J.P., Amalberti R. (1995) La conscience de la situation en aéronautique de combat. AGARD, CP-575, Bruxelles, April 24-27, 17.1-17.9

[8] NASA Langley, *Highway in the sky*, FS-1996-07-01-LARC ,1996

[9] A.K.Barrows, J.D.Powell, *Flying a tunne-in-the-sky display within the current airspace system*, AIAA 2000

Laurent GUICHARD, Master in Electronics (ENAC, 1986), graduated in Human Factors in Aeronautics (University of Paris V, 2000) and in Stress (University of Paris V, 2001), was a software project leader at STNA/DGAC before joining EUROCONTROL in 1993. He has been successively Project Leader of the Multi-Cockpit Simulator, AudioLAN, LOOK, HADES, EXPLORER projects. He is currently Project Leader of Paradigm SHIFT in the Innovative Research Area in EUROCONTROL Experimental Centre in Brétigny.

Sandrine GUIBERT, B.Sc in Electronics and Information Technology (University of Nice, 1991), Mil. ATC Diploma (France, 1992), graduated in Human Factors (University of Paris V, 2002), was a Military ATCO before joining EUROCONTROL in 1998, as a simulation analyst (RVSM, Three States, FRAP, Look). She has been successively involved in various projects such as Stress, Implicit, Explorer, Supersector and she is currently Research Assistant and Deputy of the Paradigm Shift Project at the Innovative Research Area in EUROCONTROL Experimental Centre in Brétigny.

Didier DOHY, PhD in Physics (University of Paris XIII, 1987) is an ATM architecture & engineering specialist. NeoSYS' R&D managing Director, he has been involved in earlier EUROCONTROL Experimental Centre Innovative Research projects such as Supersector or Paradigm Shift. He has 15 years of experience in ATC center design and engineering as well as data-link software implementation and validation.

Jean-Yves GRAU, Doctor of Medicine, was an ergonomist at the French Aerospace Medicine Institute before becoming a consultant in Human Factors in ATM. His research interests include the design of decision supports, human error, flight safety and ergonomic assessment of complex systems. He was also involved in the Supersector project at the EUROCONTROL Experimental Centre as its experimentation lead.

Operational Feasibility of Traffic Synchronisation – Preliminary Results

Lenka Dravecka[1,2]

Abstract—**An anticipated increase in future traffic demand has stimulated an investigation of numerous concepts aimed at improving efficiency and gaining capacity by reducing controller workload. However, the geographic constraint limits the opportunities to increase the system capacity by building new routes. Therefore, some efforts to alleviate the problem have focused on optimising the traffic flow using already known practices. The hypothesis suggests a shift from current ATM concept, which is non-synchronised, to a synchronised system with constant distance separation between all aircraft evolving in a flow.**

The goal of the research is through experimental approach to achieve objective results of traffic synchronisation through flow capacity increase (i) and potential conflict situation decrease (ii) and hence improvements of safety to support the operational feasibility applied in case of Central European Upper Airspace.

This paper describes a simple synchronisation model together with rule-based algorithm, taking into account the quantitative indicators and presents first results showing the potential of traffic synchronisation in real time.

Index Terms—**en-route speed control, future concept, task sharing, traffic flow synchronisation**

I. Introduction

SOMETIMES definitions are not so crystal clear as they should be, and in the case of Synchronisation there are few interpretations. Synchronisation in the context of this research means *the tactical establishment and maintenance of a safe flow of traffic with higher throughput.* In other words, coordination between control sectors is a necessity since the distance required to synchronise a given traffic can span across several sectors, given the small sizes of sectors in core European airspace.

As the traffic density is approaching average European levels in many of those areas, managing such high traffic growth presents a challenge [1]. This estimation leads to a serious problem, especially in the core area. In the past, the solutions to this recurrent problem were relatively simple: split the airspace in smaller sectors, re-design the airspace routes network, and improve the technology, and thus increase the controllers' productivity, or regulate the traffic demand by allocating costly ground delays.

This system is now reaching its limits and as a consequence, either air route network has to be extended to accommodate increased air traffic flows, or traffic in flows shall be better regulated with fewer extensions to current route network.

Several studies and concepts appeared recently in Europe and USA aiming to support the future vision of synchronised ATM system [2] [3] [4] [5] [6].

Numerous organisational changes in the way to handle the air traffic have been proposed – *Multi Sector Planner* [7], *Free Flight* [8], *Free-route* [9] and more recently *Sector-less* [10] or *SuperSector* [11], *Super HighWay* [12]. Some other studies on traffic organization have investigated the potential of flow synchronisation; however they focus mainly on speed control and/or solely on the terminal areas [13] [14].

All afforecited investigations are based upon the hypotheses that when traffic is synchronised, better efficiency for the overall air traffic management system could be reached thanks to the focus on the management and monitoring of traffic flow instead of individual flights. Since the flights are already organised into the flows in the en-route environment, it directly affects the efficiency and as well improves the throughput of an airport.

Among previous studies, the FAM project at EEC has explicitly considered the same assumption [15]. However, none of the above investigations have seriously considered *controllers' acceptability*, and in particular for the task of monitoring aircraft's speed adjustment and assigning flight level to achieve synchronised flows. That may span over several sectors, and therefore requires collaborative work between controllers. The work presented in this paper is framed within a PhD research that investigates the operational feasibility of traffic synchronisation applied in a case of Central European Upper Airspace.

The two main questions arising by planning temporal traffic re-organisation changes are, first, *"when should the traffic synchronisation be adapted?"*, and secondly, *"how to arrange the traffic in the most efficient fashion?"* in order to increase the airspace efficiency in terms of airspace throughput while avoiding critical situations.

The paper focuses on answering the second question for

Manuscript received March 2, 2006.

This work is funded by the CEATS Research, Development and Simulation Centre (CRDS), Budapest, Hungary.

Lenka Dravecka is a Ph.D. candidate at EUROCONTROL CEATS Research, Development and Simulation Centre (CRDS), Hungary and University of Zilina, Slovakia. Address: EUROCONTROL CRDS, Ferihegy 1, "A" Porta, H-1185 Budapest, Hungary (phone: 0036-1-297-2240; fax: 0036-1-297-2243, e-mail: lenka.dravecka@eurocontrol.int

[1] University of Zilina, Zilina, Slovakia

[2] CEATS Research, Development and Simulation Centre, Budapest, Hungary

decision making process. The aim is to demonstrate a simple model and rule-base algorithm to analyse the efficiency (airspace throughput) of synchronising independent flows of traffic in the en-route environment. The next section will briefly describe the model and will emphasize the need for quantitative parameters to evaluate the benefits of synchronisation. The third section will discuss the rule-based algorithm and present first results. The paper ends with conclusions and future zooming in the possible ways to answer the question *When.*

II. MODELLING

It is widely recognised that the present air traffic management system will not accommodate the growth of traffic beyond a certain point that is predicted to be reached in the not too distant future. In order to meet the society's expectations for safe and economic air travel, further capacity enhancement will be required that might be possible through better use of already known practices of current system.

According to the comparison of US and European en-route environment, it was shown that the controllers' productivity is unlike. The difference arises in part from the fact that the US controllers can handle more traffic (flight-hours for each hour on duty) when working at their maximum throughput [16]. An essential difference is in techniques used to organise the traffic. In US En-route Miles-In-Trail Spacing (MIT) is the most common Air Traffic Flow Management (ATFM) measure defined by the distance between two consecutive aircraft on a given flow. They are used to distribute arrival delays upstream of destination airports and to mitigate local areas of en-route space congestion. They have a significant operational advantage; when flights are formed into in-trail streams, *controllers are able to visualise and control spacing at the sector without automatic assistance*. Nothing comparable to MIT restrictions is applied in European en-route environment, the aircraft start to be organised once they reach the Terminal Area (TMA).

In conclusion, the ATFM system in US tends to work towards better utilization of airspace capacity. This in turn has an impact on controllers' workload. Because peak-time flows are more predictable in the short term, and more regular, controllers in the US are able to handle a larger number of flights simultaneously, contributing to a greater productive efficiency [17].

MIT measures are taken as an example of traffic flow and safety enhancement. Improvement of safety in European environment through removing potential conflict situations and increasing the airspace throughput as a result of traffic synchronisation is not only desired but needed.

A. *Simplified synchronisation model*

To discover the benefits of synchronisation in an en-route environment a model considering different ATM system actors (airlines, providers, efficiency of airspace use itself) is required beginning with better utilization of existing route network with respect to all above mentioned actors of the ATM system. The simple form consists of three levels: the top is the goal of the research, the second level is represented by the criteria by which the indicators will be evaluated, and the third level consists of the measurable indicators themselves (Figure 1).

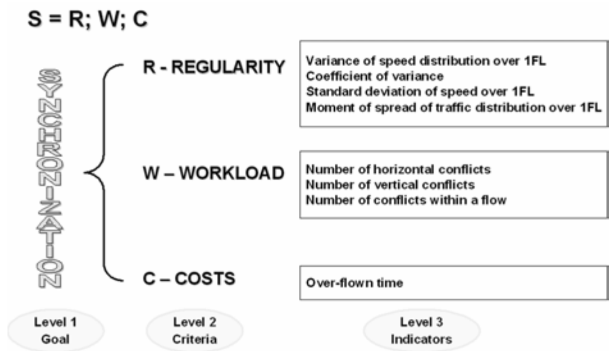


Fig. 1. Synchronisation mode

For deeper analysis measurable indicators characterising the traffic distribution after synchronisation process have been defined with respect to the limiting factors: aircraft performance envelope, number of conflict situations, controller's workload and costs of the over-flown time.

At present, the traffic is distributed on the flight levels mainly with respect to airliners preferences (fuel consumption, optimum speed). This tends to increase controllers' workload, because one flight level accommodates aircraft with several different speeds and each of these flights requires inevitable communication between pilot and controller; manoeuvres to be done.

On the contrary, reorganised traffic on flight levels allows controller to remove the conflicts occurring between flights following the same flow since they are maintaining the same speed.

With traffic synchronisation the irregularity of speed distribution on target flight level is removed with direct impact on speed diversity on neighbouring flight levels (as described later in the paper). The traffic on the target flight level is released from the aircraft not following the same target speed. This leads to more balanced traffic between flight levels, since the aircraft change the flight level from more occupied to those less occupied.

Four values are used for regularity assessment [18]:

Variance of speed distribution over 1FL is a measure of its statistical dispersion, indicating how far from the expected value its values typically are. Where the average (or mean) is a measure of the centre of a group of numbers, the variance is the measure of the spread.

Standard deviation of speed over 1FL is a measure of the degree of dispersion of the data from the mean value. In this case it is used as an indicator how balanced the traffic load is. Through the aircraft shifts into other FL, when an adjustment in speed is not applicable, the traffic distribution is changed.

Coefficient of variance is calculated for the speed distribution on selected flight levels. With this parameter the consequences of synchronisation on the surrounded traffic are examined.

Bunching index is used as a factor describing the grade of regularity in time on FL. A bunching index of 0 would describe an absolute regular flow. It is calculated by the sum of the square differences of the actual, to the theoretical capacity (of one aircraft) in each interval:

B. *Transition time*

To achieve synchronised traffic flow, a given traffic shall be transformed from a non-synchronised state to an ordered traffic. In this design, the flow isolates a part of the traffic on particular flight level selected according to predefined rules. The aircraft falling into the target speed range will be instructed to keep or adjust its speed to join the synchronised flow; otherwise it has to be instructed to change the flight level. In reality the flight level change is dependent on many factors as controller's workload and slot availability on neighbouring flight level or aircraft performance limitations. Figure 2[3] provides an example of an aircraft flight level change and an image of transition phase.

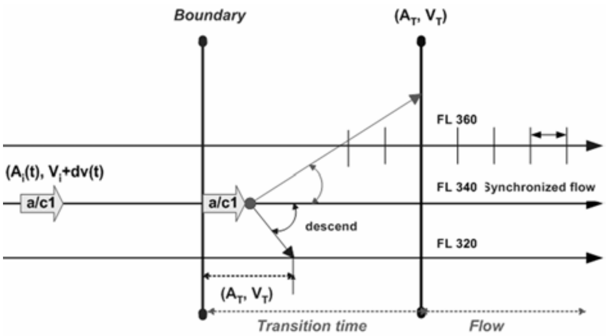


Fig. 2. Transition time

The main goal is not only to increase the route throughput without closer look on the time needed to re-organise the inbound traffic. But the impact of the route length to ensure that synchronisation is beneficial should be considered as well. The transition time (T_t) differs for each flight and depends on traffic density and the availability of the slot on target flight level. The question is:

How much time is needed to transform non-synchronised traffic into a synchronised one? Does the transit phase encompass one or more sectors?

The synchronisation process starts once the aircraft crosses the border of the airspace block. If this is not feasible within the boundaries of the first sector, the second sector follows the same procedures to achieve flow-wide improvement and so forth. The coordination and traffic organisation to achieve synchronised flow is the task of the controllers.

In order to simplify the investigation, other issues such as non-nominal weather conditions or military traffic are initially disregarded.

[3] $V_i(t)$ [NM/h] – actual speed, V_T [NM/h] – target speed, $A_i(t)$ [-] – actual FL, A_T [-] – target FL

C. *Airspace*

If en-route ATM is required to deliver traffic for the airports in a particular sequence at specific times, then it must have the ability to speed up or slow down aircraft. This will influence aircraft operating levels and, in consequence, flight efficiency. Conversely, organising and sequencing of traffic by speed and flight levels raise questions as to the optimum level of integration which is feasible and the most beneficial.

The Central European Upper Airspace[4] (Figures 3) is

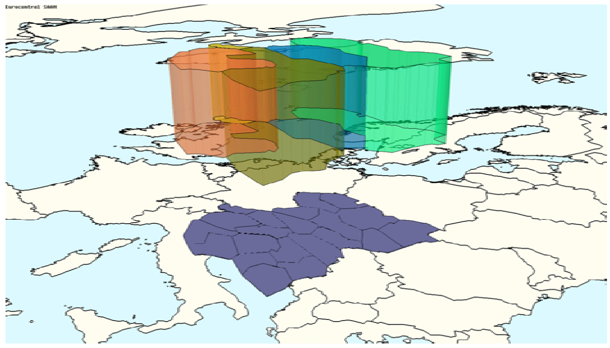


Fig. 3. CEATS Functional Airspace Block

chosen for this research because of the characteristics of its traffic: high growth and high percentage of over-flown traffic.

The chosen airspace offers high potential for performance improvement with synchronised ATM concept.

III. SELECTING ROUTE SEGMENTS

Flow consists of aircraft with a common part of the flight path or part of it spreading on more than one flight level. Adjusting the flow of traffic into given airspace, along a given route or bound for a given aerodrome, so as to ensure the most effective utilization of the airspace is very complex. To simplify this complexity an assumption has been made where *each synchronised flow has its own speed & flight level; the flow speed fixes the speed of the flight.*

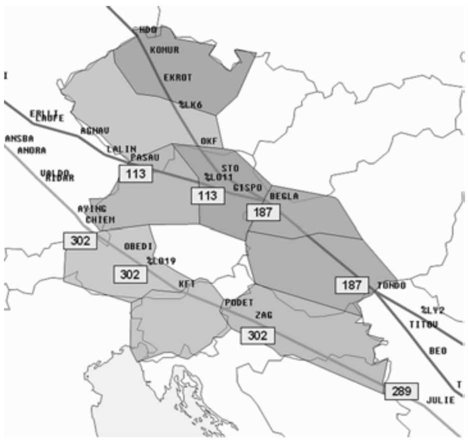


Figure 4. Daily loading of selected segments

[4] The activities of CEATS are centred towards improving the operation in this area, where Air Traffic Services are currently provided by different Area Control Centres in Vienna, Budapest, Bratislava, Ljubljana, Zagreb and Padua, and to be replaced by a unique common control centre in Vienna, called CEATS Upper Area Control Centre (CUAC), planned to start its operation in 2012.

The discovery of 'interesting flows' depends on both daily traffic distribution and frequency of the aircraft arrivals to CEATS Airspace Block. The initial investigations showed 2 main flows. These two routes carry the heaviest amount of traffic between Western and Eastern Europe; therefore they were selected for further investigation.

Aircraft performance envelopes play a major role in the ability of the aircraft to reach target flight level assigned to the aircraft regarding the speed it flies.

IV. PRELIMINARY RESULTS

A. *Traffic analysis*

Due to the fact that CEATS is not under the operation yet the modelling approach uses the data obtained with the last model based simulation (FTS4) as baselines for comparative study [19]. One 24-hours traffic sample (June 28, 2002) is used during the course of this study, increased of 36% (as estimated for the start of CUAC's initial operations). This simulation was performed with the Reorganized ATC Mathematical Simulator (RAMS Plus^{TM}) - a fast-time simulation tool commonly used at CEATS Research, Development and Simulation Centre (CRDS).

In this sample more than 7 000 flights are considered of which 5241 are in upper airspace (above FL285). To make a comparison, adequate aircraft performance data are necessary, especially true air speed during the cruising phase of flight. In reality, there is a great disparity of aircraft performances, so to simplify it those data were obtained from the aircraft performance summary tables for BADA 3.6 document [20]. For each aircraft type, the performance table specify the true air speed, rate of climb/descent and fuel consumption for each phase of flight. In the course of the study the cruising, plus cruising and descending aircraft are examined. Moreover those aircraft which have only a small offset of their entry or exit flight level to the cruising flight level were also added to the traffic sample[5]. This was done under the assumption that in this case still most of the cruising phase of flight is performed within the CEATS boundaries. To understand the traffic behaviour some additional measurements have been made including: over-flown time (mean value per sector); average speed versus % of traffic; aircraft types versus speed; speed versus FL.

The initial investigation shows that 54.42% of the traffic uses the speed 0.8M; flight level FL370 is highly preferred by 17.78% of the overall traffic. These speed and flight level have high potential to be assigned to the inbound traffic for synchronisation purposes. Nevertheless, presented analysis showed only the traffic distribution on the level basis. Another step was to look deeper and perform an analysis with the same approach on selected two main route segments.

[5] E.g.: Flight DLH2888: Entry = 250, Cruise = 370; Exit = 0

B. *Rule-based algorithm*

To perform the synchronisation a rule-based algorithm with four steps was developed according to which aircraft are organised into synchronised flows. Before discussing its details, it is useful to explain the motive. In synchronisation, if the speed change alone cannot solve the conflict, the controller can modify the flight path firstly by flight level change and if not possible by vectoring the aircraft. The algorithm implements spatial satisfaction by organising the traffic on flight levels depending on preferred flight speed. First, all aircraft on particular route involved in synchronisation process are organised on speed levels (each flight level has its own speed; flow fixes the speed of the flight). It this particular case two FL's were selected to be synchronised depending on the aircraft preferences (see later in the paper). Second step is to select speed range also based on current aircraft preferences on selected FL. Next step is to test whether the cruising speed of the aircraft can be adjusted. With respect to the result, either the aircraft has to adjust the speed or it has to change the FL. This happens under the consideration of the aircraft performance limitations.

Three scenarios of rule-based algorithm are considered:

1. Speed adjustment on FL of +/-0.02M;
2. Speed adjustment on FL of +/-0.01M;
3. Speed adjustment on FL of +/-0.02M, change FL of aircraft falling into target speed range in neighbouring FL into target FL selected for synchronisation.

The limitations are set to minus two flight levels to descent and plus one FL for climbing.

C. *CHIEM-JULIE –preliminary results*

Scenarios are compared resulting from different values for the control variable Speed in the algorithm. In order to asses the modifications, indicators are calculated before and after the synchronisation.

Analysed data of selected segments (CHIEM-OBEDI-ZAG; TONDO-BEGLA) contain 302 flights entering at CHIEM and flying via OBEDI to ZAG and 184 flights using the segment TONDO – BEGLA. In CEATS environment Reduced Vertical Separation Minima (RVSM) will be applied between FL290 and FL410 and the traffic must be split into odd and even flight levels. Both selected segments are uni-directional. The

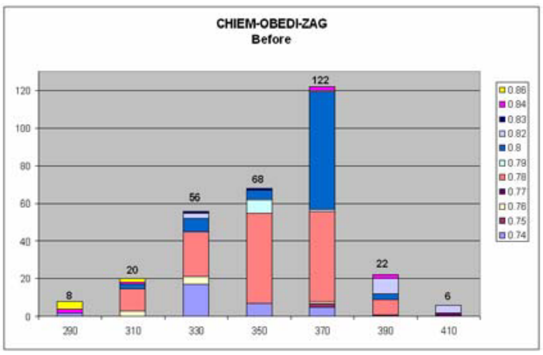


Figure 6. Traffic distribution on CHIEM-ZAG segment before synchronisation

current traffic distribution on the segment CHIEM-ZAG is illustrated in Figure 6 (the most preferred FL370 accommodates 122 flights; and 50.82% of the traffic on this FL prefers the speed 0.8M) and Figures 7 and 8 depict the situation after applying the algorithm.

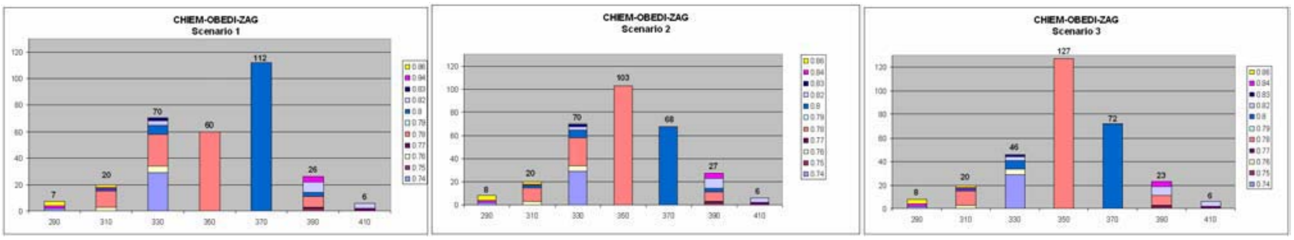


Figure 7. Scenario 1-2-3: Traffic distribution on CHIEM-ZAG segment after synchronisation

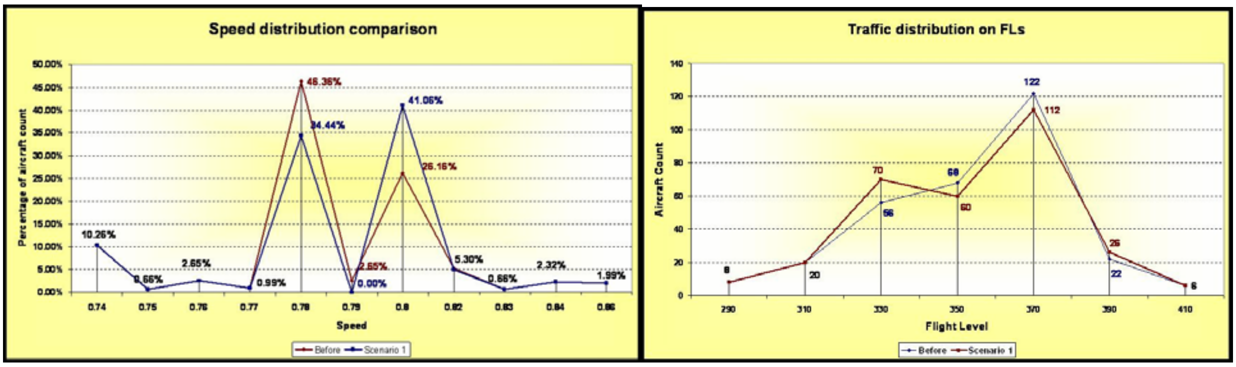


Figure 8. CHIEM-ZAG: Comparison of Speed and Traffic distribution between current situation and Scenario 1

For example for *the scenario 1* on the en-route segment CHIEM-ZAG (speed adjustment of +/-0.02M) two flight levels have been selected: FL 370 and FL350. FL 370 initially accommodates 122 and FL350 68 flights. The target speed 0.8M for synchronising the flow on FL370 is currently preferred by 50.82% and speed 0.78M on FL350 represents 70.59% of the traffic. All those flights that didn't fall into target speed range have been instructed to adjacent flight levels.

After the synchronisation, 60 flights remained on FL350 and 112 flights on FL370, to the rest of the flights was assigned another flight level.

Likewise, in *scenario 2* (speed adjustment of +/-0.01M) two flight levels have been selected: FL 370 with speed 0.8M and FL 350 with speed 0.78M. After synchronisation 68 flights remained on FL370 in comparison with double increase of the traffic on FL350 (103). This modification of the algorithm has caused increase of the vertical movements in order to build the synchronised flow. From the Figure 7 it is visible how the traffic distribution on the flight levels has changed compared to basic scenario and scenario 1.

Scenario 3 is modification of scenario 1 with additional rule. In this case the flights on neighbouring FL's falling into the target speed range have been instructed to target flight level to join the flow. From the Figure 7 it is evident that in scenario 3, FL350 accommodates much larger number of flights (127) what is in contrary with FL350 with less traffic (72) than in scenario 1 or scenario 2.

Next step was the calculation of indicators for each above mentioned scenario. The goal was to quantitatively find the most beneficial with respect to traffic balance between flight levels and minimising the number of level changes.

All scenarios show 100% regularity in speed distribution over target FL's and the variance of speed in these FL's is zero. All the other indicators have been compared and preliminary conclusions drawn (Figure 9).

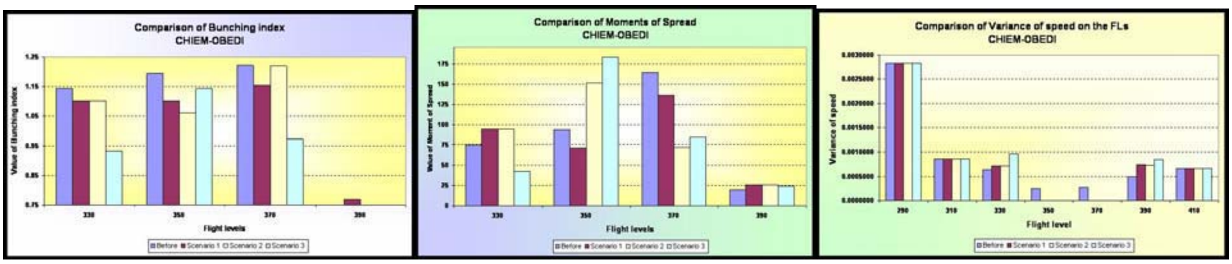


Figure 9. CHIEM-ZAG Regularity indicators

a) Regularity (R)

When prioritizing the aircraft speed adjustment on target flight level the best results can be seen in both: scenario 2 and 3 (of +/-0.001M) where only one flight is penalized in terms of speed reduction. Nevertheless it is not relevant to take into account this result independently since other indicators should be considered as well.

b) Workload (W)

In case of workload the initial investigation was concentrated on the number of vertical movements that could lead to increase of controller's workload and potential conflict situations. Prioritizing this criterion scenario 1 could be used for synchronising the traffic in real time.

c) Costs (C)

Because of speed changes, there is also an impact on flight time costs considering the change of the over-flown time through the airspace. Due to the fact, that after the synchronisation only one speed per one flight level will be used, the over-flown time can be defined precisely. It can be accurately calculated with simple formula and results of both over-flown times (current situation and after the synchronisation) compared:

$$t := \frac{s}{v_{TAS}} \qquad (1)$$ [6]

This can be appreciated by the airlines since they will be eventually able to optimise (modify) integrated times in the flight planes when crossing given airspace.

Table 3 depicts the difference in flight time a reference flight[7], out of the used traffic sample, from CEATS entry at CHIEM, to its destination at Milas Bodrum, Turkey (LTFE) (954 NM). After application of the synchronisation the reference flight gained 1minute and 19 seconds to cross the CEATS Airspace from navaid CHIEM to navaid JULIE and more then 3 minutes to reach its destination.

However, through the moderate speed adjustment allowed, over maximum +/- 0.02 M, only little changes in flight times are caused. Thereat only a little influence on the flight time costs can be expected.

As the highest priority is to synchronise the inbound traffic with the lowest number of aircraft shifts it is suggested to apply scenario 1 in both route segments.

In this study, the variances of speeds on the FLs of the two discussed routes are assesd in order to describe the degree of synchronisation (ΔG_s)In the case of a complete synchronisation of target speed, the value of the variance for this FL is zero. But adjacent FLs must be considered as well, as they are also effected by synchronisation. During the synchronisation process, aircraft which can not adjust to the target speed are sent to adjacent FLs. This might increase the number of speeds found in these FLs, thus decreasing their regularity. This is a contrary effect to the absolute regularity created on the target FLs. As the variance has no maximum, no absolute value can be calculated. Thus, changes in the degree of synchronisation of the routes are used and are expressed as a relative value (ΔG_s). ΔG_s can be calculated with formulas intended for variance analysis [18]. The variance is the root-mean-square deviation. In order to compare the conditions between flight levels, the square deviation of each FL is added up (Formula 2) and then divided by the number of the degrees of freedom (Formula 3) (as the variance of one aircraft on one flight level cannot be calculated, this value is never 0). ΔG_s is calculated as a relation of MQW after synchronisation to the value of MQW before. Comparing two degrees of synchronisation can result also in negative values, in case the situation after synchronising is worse then the current situation.

TABLE 3. DIFFERENCES IN FLIGHT TIME THROUGH ACCELERATION OVER +0.02 M

CHIEM - JULIE							
FL	initial TAS [M]	initial TAS [kt]	sync. TAS [M]	synch TAS [kt]	initial flight time	sync. flight time	time difference
390	0,78	446,350	0,8	457,795	0:52:39	0:51:20	0:01:19
Reference flight: CHIEM - LTFE							
FL	initial TAS [M]	initial TAS [kt]	sync. TAS [M]	synch TAS [kt]	initial flight time	sync. flight time	time difference
390	0,78	446,350	0,8	457,795	2:08:14	2:05:02	0:03:12

TABLE 4. DEGREE OF SYNCHRONISATION

	CHIEM - JULIE			TONDO – BEGLA	
Change of the degree of synchronisation [%]	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2
	20.27	20.01	27.74	20.8	-7.49

The degree of synchronisation improved in comparison with current situation approximately about 20% in all

[6] v_{tas} [kt] - true air speed, t [min] – flight time, s [NM] – flight distance

[7] Flight AMM0549 from London Gatwick, Great Britain to Milas Bodrum, Turkey

scenarios except scenario two of segment TONDO-BEGLA. This downgrade is a result of shifting a large number of flights flying at speed 0.8 Mach from FL360 to FL340 which represents on FL340 an additional major speed group. That increases the square deviation from the average speed on this FL affecting the ΔG_s in the scenario. In scenario 3 of segment CHIEM-OBEDI with additional shifts of flight to target flight reached an improvement of 27.74% compared to current situation.

$$SQW = \sum_{i=1}^{p}\sum_{j=1}^{n}\left(x_{ij} - x_{mi}\right)^2 \qquad (2)^{8}$$

$$MQW = \frac{SQW}{(n-p)} \qquad (3)^{9}$$

$$\Delta G_s = \frac{MQW_S}{MQW_0} \qquad (4)^{10}$$

V. Conclusion

The traffic complexity in the 'core area' requires defining a specific mode of operation by separating level flying aircraft from climbs, descent. It is reasonable to suppose that the increase in demand will result in an increase in en-route traffic over the core area, which is already congested to the point where it gives rise to numerous regulations. The aim is to relieve pressure on the main axes forming synchronised flows with flights flying the same speed. This could span the continent and they could be reserved for steady aircraft in level flight. Traffic management on these routes will be flow-based, with closure conflicts but no convergence conflicts. The benefit of this organisation is simplified traffic, making it possible to assimilate a larger number of aircraft.

This study is framed within a PhD research with the main objective to increase the system efficiency to achieve throughput benefits from present-day air traffic control procedures.

Traffic flow synchronisation can be used to smooth the traffic operations in areas with mainly long-hauls flights.

When assessing operational feasibility of traffic synchronisation three areas have been discussed: 1. criteria, 2. operational conditions (rule-based algorithm) and 3. indicators. Improvements in utilization of current route network have largely been focused on proposal of the model with respect to these criteria and limiting factors. By drawing up the preliminary results from a synchronisation feasibility investigation, it has been shown how the rule-based algorithm being assess exercises an influence on the results and the relative importance of each factor considered.

This paper presented an example of the sensitivity of the results due to algorithm modification under specific condition. The evaluation showed the importance of number of vertical movements and traffic balance between flight levels for the solution choice.

An important consideration is the ability to validate the model by comparing calculated values with available values under actual conditions (in this case, last updated data from FTS4).

VI. Future Work

To build synchronised flow, inter-sector coordination is needed and may induce an increase of the controllers' workload in comparison with the maintenance of already organised flow. The hypothesis lies upon the operational feasibility of synchronisation in an en-route environment.

Experimentation is to set up and empirically validate the hypothesis. Validation method is aimed on: making sure that synchronisation is effective and estimates the potential benefits of the concept.

The first point can be checked by considering the conflicts detected between flights when synchronising. The potential benefits of the concept can be assessed by several means, including real-time simulations with air traffic controllers. However, it is possible to start with simpler methods, for example by analysing the number and nature of the conflicts.

The second validation method consists in running fast-time simulation, first with reference traffic, and second with a modified traffic scenarios, where flights belonging to the main flows are to be reorganised. The number of detected conflicts allows comparing these traffic situations, considering that flights following same flow have the same speed and should be able to maintain their separation with other along-track traffic. In the reference traffic the flight plans are considered as they are. In the modified traffic, the flights belonging to the target flow follow the same speed. The other flights follow their standard route, as declared in their flight plan.

The research follows the second method based on running 'synchronised' fast-time simulation (SFTS). The validation is made only on few sectors of CEATS airspace (those effected by the synchronisation). In order to test this hypothesis, the traffic scenarios are set up with different traffic levels up to increase of 200%. The objective is to verify through measurements the ability of the en-route airspace (particular routes) to **accommodate substantial increases in traffic volume through the increase of route throughput (1).** The second profit of the synchronisation is

[8] x_{ij} [-] – speed of aircraft j on the FL i; $x_{m,i}$ [-] – average of speeds of the FL i; n [-] – aircraft count on the route; p [-] – number of flight levels of the route

[9] MQW – mean square deviation within FLs (variance); SQW – sum of square deviations on the FLs; n [-] – aircraft count on the route; p [-] – number of flight levels of the route

[10] 0 – index for value before synchronization; S – index for value after synchronisation

computed by **comparing the number of conflict detected in the reference traffic, to the number of conflicts in the reorganised traffic (2)**, *removing the conflicts occurring between flights following the same flow*. This factor directly affects **controller's workload (3)** and is very important for its reduction. Currently in fast-time model simulations the measurement of workload is derived from the mathematical calculation of the total working times recorded for each ATC task category (Flight data management, Co-ordinations, Conflict search, Routine R/T, Radar). This main categorization in CEATS studies and each category consists of ATC task set. These tasks are assigned to defined actors, i.e. planning and/or executive controller. The last measure is the potential impact of synchronisation on **minimising over-flown time (4)**.

Measures proposed above study impact of new ATM element (e.g. new traffic re-organisation in en-route environment); propose changes to increase airspace throughput and controller productivity by decreasing the number of conflict situations between flights following the same flow. The required degree of accuracy of the results strongly depends on the usage of the method. Since it does not directly take into account the dynamic nature of the traffic; the next step is to consider this issue and look on the question *When to apply it*.

Additional measurement could be done considering the needs of airlines, the extra costs they have to pay (e.g. fuel consumptions) in the synchronisation process. And as well for the future to consider the time (minimum and maximum) the flight should be on the cruising level to become part of the synchronised flow.

ACKNOWLEDGMENT

The author is grateful to the discussions, and knowledge transfer, with Mr. Vu Duong (EUROCONTROL EEC, Bretigny) and with all the personnel of CRDS Budapest, Hungary. Especially to Philippe Debels, Sophie Dusire, Borce Dvojakovski, Gabor Fugedi and Iosif Ackermann for their advices, time and continuous help.

REFERENCES

[1] *EUROCONTROL PRU Performance Review report 8*, Performance Review Commission, EUROCONTROL., 2004.

[2] N.Suarez, M. Poppe and N.Makins, "From Strategic Planning to Real-Time Optimization," European Commission, 2002.

[3] R. Graham, A. Marsden, I. Pichancourt, F. Dowling, "Controller Roles – Time to change," *in Proc. 3rd USA/Europe R&D Seminar*, Napoli, Italy, 2000.

[4] S. Stoltz, P. Ky, "Reducing Traffic bunching through a more Flexible Air Traffic Flow Management," *in Proc. 4th USA/Europe R&D Seminar,* Santa Fe, USA, 2001.

[5] R. Graham, E. Hoffman, Ch. Pusch,K. Zeghal,"Absolute versus Relative navigation: Theoretical Considerations from a ATM Perspective*," in Proc. 5th USA/Europe R&D Seminar*, Budapest, 2003.

[6] L. Guichard, S. Guibert, D. dohy, J.-Y. Grau, K. Belahcene, "Paradigm SHIFT – Dual Airspace Concept Assessment," *in Proc. 4th EUROCONTROL Innovative research Workshop & Exhibition*, Bretigny-sur-Orge, 2005.

[7] C. Meckiff, R. Chone, J.P. Nicolaon, "The Tactical Load smoother for Multi-Sector Planner," *in Proc. 2nd USA/Europe R&D Seminar*, Orlando, Florida, USA, 1999.

[8] J. Post, D. Knorr, "Free Flight Program Update," *in Proc. 5th USA/Europe R&D Seminar*, Budapest, 2003.

[9] D. Schafer, "Human factors perspective on free routing and airborne separation assurance in the Mediterranean Airspace," in *Proc. 5th USA/Europe R&D Seminar*, Budapest, 2003.

[10] D. Nace, J. Carlier, N.L. Doan, V. Duong, "A linear programming approach for route and level flight assignment," in *Proc. 5th USA/Europe R&D Seminar*, Budapest, 2003.

[11] G.Gawinowski, V Duong, "Bridging the Predictive and Adaptive Issues in Air Traffic Management: The Synchronous ATM," *in Proc. 22nd DASC*, Indianapolis, Indiana, USA, 2003.

[12] N. Suarez, "Super HighWay," *in Proc. 4th EUROCONTROL Innovative research Workshop & Exhibition*, Bretigny-sur-Orge, 2005.

[13] R. Ehrmanntraut, "Towards an Operational Concept for Integrated Adaptive and Predictive Air Traffic Management," *in Proc. 22nd DASC*, Indianapolis, Indiana, USA, 2003.

[14] T.S. Abbott, "Speed control law for precision terminal area in-trail self spacing," NASA Langley Research Centre, Virginia, USA, 2002.

[15] *Future ATFM Measures,* Operational Concept Document, Eurocontrol, 2003.

[16] K.-A. White, "A study of Chicago O'Hare's March 2005 restrictions using an MIT Analysis tool," *in Proc. 24th DASC*, Washington D.C, USA, 2005.

[17] *A Comparison of performance in selected US and European En-route Centres*, EUROCONTROL Performance Review Commission, 2003.

[18] K. Nachtigall, "Verkehrssystemtheorie, Manuskript zur Vorlessung im WS und SS 2001/02, Technische Universitat Dresden, Dresden, Germany, 2001.

[19] D. Brain, "CEATS Fast-Time Simulation No.4," CRDS-CEATS Research, Development and Simulation Centre, Budapest, Hungary, 2004.

[20] Aircraft Performance Summary Tables for the Base of Aircraft Data (BADA) Revision 3.6, EUROCONTROL, 2004.

Lenka Dravecka is a Ph.D. candidate at EUROCONTROL CEATS Research, Development and Simulation Centre (CRDS), Hungary, under the supervision of Dr. Vu Duong from EUROCONTROL Experimental Centre (EEC). She holds a Master Degree in Civil Aviation Engineering from the University of Zilina. Address: EUROCONTROL CRDS, Ferihegy 1, "A" Porta, H-1185 Budapest, Hungary (phone: 0036-1-297-2240; fax: 0036-1-297-2243, e-mail: lenka.dravecka@eurocontrol.int

Effect of Realistic Speed Change Maneuvers on Aircraft Economics

Cem Cetek, MSME

Abstract—**Sharp increase in air traffic densities in the world over the last decade made many researchers prompted to investigate innovative air traffic management concepts to improve the current air traffic capacity and economic performance of aircraft. Providing aircraft operators more freedom to choose their trajectories is one of them. In such an air traffic environment, resolution of aircraft conflicts is a key element for both flight safety and aircraft economics.**

This study analyzes the economic performance of a specific conflict resolution strategy based on speed change between two aircraft in terms of extra time and fuel consumption. In this strategy, minimum time solutions are investigated without violating safe separation and realistic operational constraints for various flight and conflicts conditions in horizontal plane. Maneuver models include detailed aerodynamic and engine characteristics to make a more accurate analysis, which is very critical for both safety and economic efficiency concerns.

Using speed change alone seems unattractive since it requires longer resolution time than heading change. On the other hand, speed change maneuvers can support heading change maneuvers and extent their range of application especially for small route crossing angles. Study of these isolated conflict resolution cases can be an effective tool for aircraft operators and air traffic system designers to make a complete economic performance analysis of flight in possible future scenarios.

Index Terms—**Aircraft Economics, Air Traffic Control, Conflict Resolution, Trajectory Optimization**

I. INTRODUCTION

IN the United States and Europe, number of the civil flights increased over 40% in the last decade [1],[2]. Various projections for the next ten years show that the demand for air transportation will grow at about 30-47% [3]-[5]. These numbers indicate that the problems of the current ATM system regarding capacity and efficiency will become much more serious in the future.

In order to overcome these problems, researchers have been working on innovative ATM concepts such as free flight, providing more freedom to aircraft operators to choose their trajectories [6]. In such a flexible air traffic environment, resolution of aircraft conflicts is a key element for both flight safety and aircraft economics.

Cem Cetek is with School of Civil Aviation, Anadolu University, Eskisehir 26470, Turkey (phone: 90-222-321-3550 (ext. 6811); fax: 90-222-322-1619; e-mail: ccetek@ anadolu.edu.tr).

There are numerous studies regarding conflict resolution from both algorithmic and operational perspectives. Most of these approaches are summarized in [7]-[10]. Despite this extensive literature, studies investigating conflict resolution maneuvers in terms of aircraft economics are still limited in number and context with few exceptions [11]-[12].

The aim of this study is to analyze the economic performance of a specific conflict resolution strategy based on speed change maneuvers. This strategy includes non-cooperative avoidance and recovery maneuvers between two aircraft in horizontal plane. In the analysis, minimum time solutions for resolution and recovery are investigated for various flight and conflict scenarios. Delays and extra fuel consumption are calculated for the each case with respect to the reference flight conditions. Maneuvers subjected realistic operational constraints include more detailed aerodynamic and engine models for an accurate analysis.

II. MATHEMATICAL MODEL

A. Problem Statement

Air traffic conflict arising between two commercial aircraft at the same flight altitude with intersecting routes is shown in Fig. 1. Their encounter geometry can be characterized by their relative angular

$$\theta = \psi_1(t) - \psi_2 . \tag{1}$$

$$s(t) = \left((x_1(t) - x_2(t))^2 + y_1^2(t)\right)^{1/2} . \tag{2}$$

In (1), θ is the path crossing angle; and ψ_1, ψ_2 are the headings of the aircraft 1 and 2, respectively. In (2), s is the distance between the aircraft; and x_1, $y1$, y_2 are linear positions of the aircraft.

The ratio of initial airspeeds of the aircraft is described as

$$VR = \frac{v_{20}}{v_{10}} . \tag{3}$$

where v_{10}, v_{20} are the airspeeds of the aircraft at time, $t = t_0$.

The closest distance between the aircraft, s_{min} is another parameter indicating the seriousness of the conflict. For en route airspace any encounter results in conflict $0 \le s_{min} < s_y$ for any $t \in [t_0, t_f]$ where s_y is the minimum horizontal safe separation and equals to 9.26 km (5nm).

It is supposed that if no avoidance action is taken, at t_c, aircraft 1 will reach point C whereas the aircraft 2 will be at

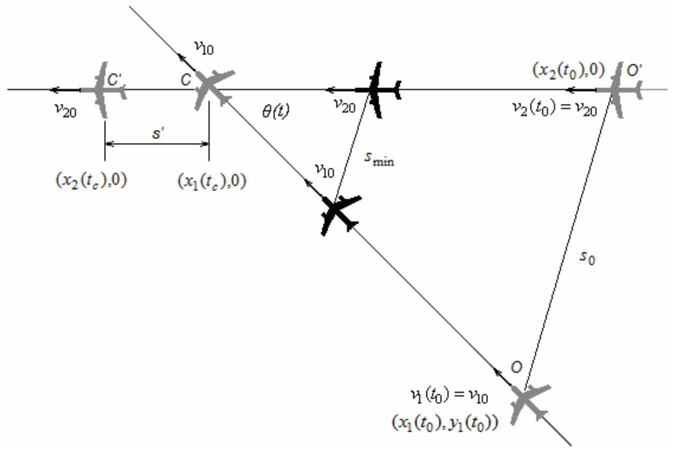


Fig. 1. Conflict geometry of the aircraft with intersecting linear routes

point C'. Let the distance $s' = | CC' |$, can be calculated in terms of VR, θ and s_{min} as

$$s' = \frac{s_{\min}}{\left(\left(1 - \frac{(VR - \cos\theta)^2}{VR^2 - 2VR\cos\theta + 1}\right)^2 + \left(\frac{(VR - \cos\theta)\sin\theta}{VR^2 - 2VR\cos\theta + 1}\right)^2\right)^{1/2}} \quad (4)$$

.

The time t_c can be found using the following relation

$$t_c = \frac{|OC|}{v_{10}} = \frac{|OC'| + s'}{VRv_{10}}. \quad (5)$$

where the distances $| OC |$ and $| OC' |$ are unknowns. They will be determined using an optimization process in part III.

B. Rules of Road

Since non-cooperative conflict resolution is investigated, a set of rules of road are necessary. The following assumptions are made as the rules of road in the encounter described in Fig.1.

1. $VR \geq 1$
2. No overtaking or head-on encounter is considered.
3. Slower aircraft is always supposed to maneuver.
4. If the airspeeds are equal, the aircraft on the starboard has the priority of the road.

According to these assumptions, aircraft always maneuvers maneuver, while aircraft 2 takes no action.

C. Equations of Motion

The following point mass equations of motion are used to describe the system shown in Fig. 1.

$$\dot{x}_1 = a(\sigma)M_1(t)\cos\theta(t). \quad (6)$$

$$\dot{y}_1 = a(\sigma)M_1(t)\sin\theta(t). \quad (7)$$

$$\dot{x}_2 = a(\sigma)\cdot VR \cdot M_1(t_0). \quad (8)$$

$$\dot{M}_1 = \frac{1}{m_1 \cdot a(\sigma)}(\eta_1 F_1(\sigma, M_1) - D_1(\sigma, M_1, \alpha_1)). \quad (9)$$

$$L(\sigma, M_1, \alpha_1) - m_1 g = 0. \quad (10)$$

where a is the speed of sound; σ is the relative density at the given flight altitude; g is the gravitational acceleration; $M_1=v_1/a(\sigma)$ is the Mach number; m_1 is the reference mass; F_1 is the nominal engine thrust; D_1 is the aerodynamic drag force; η_1 is the throttle setting and α_1 is the angle of attack of the aircraft 1.

Equations (6)-(10) are derived assuming standard atmospheric conditions, zero wind, constant mass and small angle of attack controlling aircraft to maintain the flight altitude constant.

For horizontal flight, aerodynamic drag force, D on the aircraft including effects of compressibility and cambered wing can be formulated as

$$D_1(M) = d_1 M_1^{\,2} + \frac{d_2 M_1^{\,2}}{\sqrt{1 - M_1^{\,2}}}\left(\frac{d_3}{M_1^{\,2}} - d_4\right)^2 \quad (11)$$

where d_1, d_2, d_3 and d_4 drag force parameters depending on flight level and aircraft's physical and aerodynamic properties.

Total nominal engine thrust, F_1 is modeled based on a high by-pass ratio gas turbofan jet engine data from [13] as

$$F(\sigma, M) = c_1 M_1^{\,2} + c_2 M_1 + c_3 \quad (12)$$

where c_1, c_2 and c_3 are thrust parameters as a function of flight level.

The system described in (6)-(10) is subjected to the following constraints in [t_0 ,t_f]

$$s(t) \geq s_y \quad (13)$$

$$M_{1md} \leq M_1 \leq M_{10} = M_{cr} \quad (14)$$

In (14), M_{1md} is the minimum drag Mach number, and M_{cr} is the optimal cruise Mach number for fuel consumption at the chosen flight altitude.

The control variable, η_1 is also limited by the following operational constraints.

$$\eta_{\min} \leq \eta_1 \leq \eta_{\max} \quad (15)$$

These minimum and maximum throttle setting values depend on the following acceleration limits imposed by passenger comfort limits [14] and engine constraints.

$$-0.02g \leq a_{\min} \leq \dot{v}_1 \leq a_{\max} \leq 0.02g \quad (16)$$

The accuracy of the point mass model described in (6)-(10) can be increased if throttle setting transients are taken into account. Transient changes in $\Delta\eta_1$ can be described using a first-order response model described in [15] as

$$\Delta\eta_1 = \Delta\eta_1^c(1 - e^{(-t/\tau_c)}) \quad (17)$$

where $\Delta\eta_1^{\,c}$ is the throttle command and τ_c is the time constant which is taken as 3.5 seconds.

From Fig. 1, initial conditions of the system are described as

$$(x_1(t_0), y_1(t_0)) = (|OC|\cos\theta, |OC|\sin\theta) \quad (18)$$

$$(x_2(t_0), y_2(t_0)) = (|O'C|, 0) \quad (19)$$

$$v_1(t_0) = v_{10} = v_{cr} \quad (20)$$

$$v_2(t_0) = v_{20} = VR \cdot v_{cr} \quad (21)$$

From Fig. 2, final conditions are defined as follows

$$(x_1(t_f), y_1(t_f)) = (|CD|\cos\theta, |CD|\sin\theta) \quad (22)$$

$$(x_2(t_0), y_2(t_0)) = (|CD'|, 0) \quad (23)$$

$$v_1(t_0) = v_{10} = v_{cr} \quad (24)$$

$$v_2(t_0) = v_{20} = VR \cdot v_{cr} \quad (25)$$

In (18)-(25), t_0=0 and distances $| OC |$, $| OC' |$, $| CD |$ and $|$

CD'| depend on free final time, t_f.

III. Resolution Strategy

A. Speed Change Maneuvers

Speed change strategy is graphically described in Fig. 2. In order to perform speed maneuvers in minimum time with satisfying the constraints (13)-(16) and boundary conditions (18)-(25), a three step procedure is proposed. These steps are slowing down with maximum deceleration (from point *O* to *A*), steady flight at the minimum allowable speed (from point *A* to *C*) and speeding up back to the original speed with maximum acceleration (from point C to D).

Total resolution (avoidance and recovery) time, therefore, can be described as

$$t_f = T_{dec} + T_{ss} + T_{acc} \ . \qquad (26)$$

where T_{dec} is the period of deceleration in $[t_0,t_1]$; T_{ss} is the period of cruise at the minimum speed in $[t_1,t_2]$ and T_{acc} is the period of acceleration back to the original speed in $[t_2,t_3]$.

T_{dec} and T_{acc} can be formulated with the following relations

$$T_{dec} = \frac{a(\sigma)\left|M_{cr} - M_{md}\right|}{\left|\dot{v}_1\right|_{t_0}^{t_1}} = \frac{a(\sigma)\left|M_{cr} - M_{md}\right|}{\left|a_{\min}\right|} + t_{thrt} \ . \qquad (27)$$

$$T_{dec} = \frac{a(\sigma)\left|M_{cr} - M_{md}\right|}{\left|\dot{v}_1\right|_{t_2}^{t_3}} = \frac{a(\sigma)\left|M_{cr} - M_{md}\right|}{\left|a_{\max}\right|} + t_{thrt} \ . \qquad (28)$$

In (27)-(28), the term, t_{thrt} is the delay due to throttle setting transients described in (17).

Using the conflict geometries before and after the resolution in Fig. 1 and Fig. 2, T_{ss} can be described as

$$T_{ss} = \frac{\left|OA\right| / v_{10} + (s'' - s')/(VR \cdot v_{10}) - T_{dec}}{(1 - v_{1md} / v_{10})} \qquad (29)$$

where

$$\left|OA\right| = \int_0^{T_{dec}} v_1(t)dt. \qquad (30)$$

The only unknown parameter in (29) is the distance *s"* and in order to find it, the following constrained optimization problem is supposed to be solved.

$$\min\left(\left((v_2^0 - v_{1md}\cos\theta)\tau - s''\right)^2 + \left(v_{1md}\sin\theta\cdot\tau\right)^2 - s_y^2\right) \qquad (31)$$

$$s(\tau) \geq s_y$$

$$s'' \geq s_y$$

where $\tau = t - t_1$.

B. Reference Time and Fuel Consumption

During the resolution process, aircraft 1 flies the distance |*CD*|. Reference time is calculated based on this distance as

$$t_{ref} = \frac{\left|OD\right|}{v_{10}} \qquad (32)$$

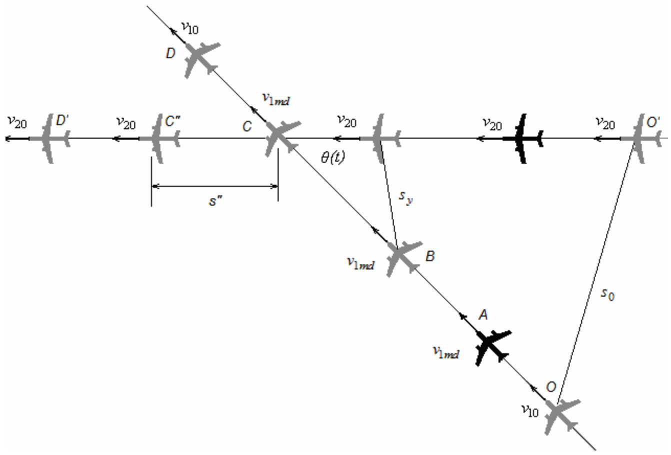


Fig. 2. Non-cooperative speed change avoidance and recovery maneuvers

Therefore, the delay due to the maneuvers can be written as

$$\Delta t = t_f - \frac{\left|OD\right|}{v_{10}} \qquad (33)$$

In order to calculate fuel consumption, fuel flow rate of the engines should be known. Based on the engine data [13], fuel flow rate is modeled as a quadratic function of η_1 as

$$\dot{\omega}_f = (f_1(\sigma, M_1)\eta_1{}^2 + f_2(\sigma, M_1)) \cdot F_1(\sigma, M) \qquad (34)$$

In (34) f_1 and f_2 are fuel flow parameters and can be written as a function of Mach number as

$$f_1(M_1) = a_1 M_1^2 + a_2 M_1 + a_3 \qquad (35)$$

$$f_2(M_1) = b_1 M_1^2 + b_2 \qquad (36)$$

In (35)-(36), a_1, a_2, a_3, b_1 and b_2 are specific fuel consumption (SPC) constants for the given flight altitude.

Therefore, reference and extra fuel consumptions are as follows

$$(\omega_f)_{ref} = \int_0^{t_{ref}} \dot{\omega}_f \, dt \qquad (35)$$

$$\Delta\omega_f = \omega_f - (\omega_f)_{ref} \qquad (36)$$

IV. Results

A. Simulation Parameters

Resolution maneuvers are tested for a narrow body commercial jet transport based on Boeing 737-400 physical and aerodynamic date in [14] and [16]. Thrust and specific fuel consumption data in [13] is adapted for this aircraft.

In calculations *VR* are limited in between 1.0 and 1.4. For commercial jet transports, it seems unlikely to have *VR* values greater than 1.4, while they flight at their optimal cruise speed in en route airspace.

As it is mentioned earlier, no head-on or overtaking conflicts are considered in the study; therefore, values of the path crossing angle, θ are taken in between 30˚and 150˚in the analysis. Minimum horizontal separation before the resolution, s_{min} is set to 0, 3.7 and 7.4 km (0, 2 and 4 nm) during the tests. Simulation parameters regarding aircraft model and flight conditions are listed in Table I.

TABLE I

SIMULATION PARAMETERS

	Parameter	Value
h	Altitude, m (ft)	7600 (25000)
M_{md}	Minimum Drag Mach Number	0.572
M_{cr}	Optimum Cruise Mach Number	0.645
m	Mass of Aircraft (kg)	58,000
d_1	Drag Constant (N)	62,338
d_2	Drag Constant (N)	122,390
d_3	Drag Constant	0.2372
d_4	Drag Constant	0.2
c_1	Thrust Constant (N)	32,504
c_2	Thrust Constant (N)	-46,086
c_3	Thrust Constant (N)	78,720
a_1	SPC constant (kg/N/s)	$1.698{\cdot}10^{-5}$
a_2	SPC constant (kg/N/s)	$-2.038{\cdot}10^{-5}$
a_3	SPC constant (kg/N/s)	$0.699{\cdot}10^{-5}$
b_1	SPC constant (kg/N/s)	$0.976{\cdot}10^{-5}$
b_2	SPC constant (kg/N/s)	$1.023{\cdot}10^{-5}$

B. Test Results

During the simulations, total resolution time; total time delay and percentage extra fuel consumption are estimated for the given conflict configurations. In each test, variations of these measures with respect to θ, VR and s_{min} are analyzed. Simulation code is written using Optimization Toolbox of MATLAB software.

In Fig. 3, variation of total resolution (avoidance and recovery) time, t_f with θ is shown for VR=1.0, 1.2 and 1.4, respectively. For the smaller values of θ, t_f is about 11 minutes and VR has no significant effect on it. As θ increases, t_f reaches to unacceptable values for the deterministic conflict resolution. Higher values of VR result in shorter t_f, especially when $\theta \geq 60^\circ$.

Pre-solution minimum horizontal separation, effect of s_{min} on t_f is stronger than VR, as it can be seen from Fig. 4. As s_{min} decreases, naturally the conflict becomes more dangerous and longer t_f is required for resolution. For larger values of s_{min}, t_f becomes less sensitive to the variation of θ.

Similar to t_f, time delay increases drastically with increasing θ (Fig. 5 and 6). For $\theta < 90^\circ$, it is under 60 seconds, which proves that speed change is effective for the encounters having narrow crossing angles.

Fig. 7 and 8 show that percentage extra fuel consumption is not as sensitive as time delay to the variation of θ, VR and s_{min}. It remains at about 3% in most of the cases. Therefore, in terms of total cost, speed change maneuvers are less costly for smaller crossing angles.

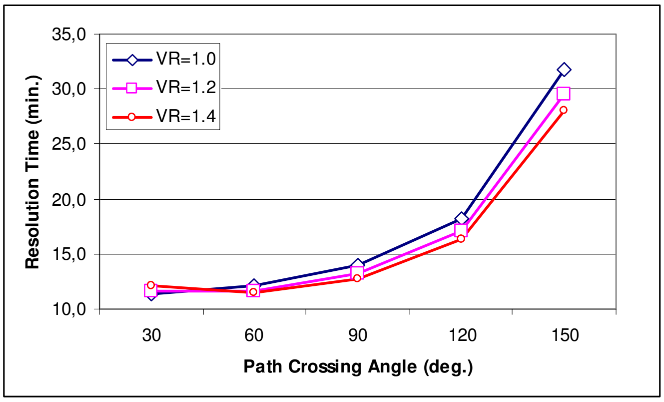


Fig. 3. Variation of the total resolution (avoidance and recovery) time with path crossing angle and initial speed ratio of the aircraft.

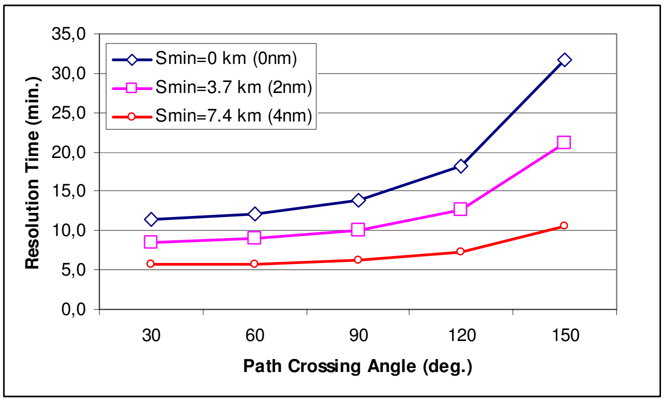


Fig. 4. Variation of the total resolution (avoidance and recovery) time with path crossing angle and minimum horizontal separation before the resolution.

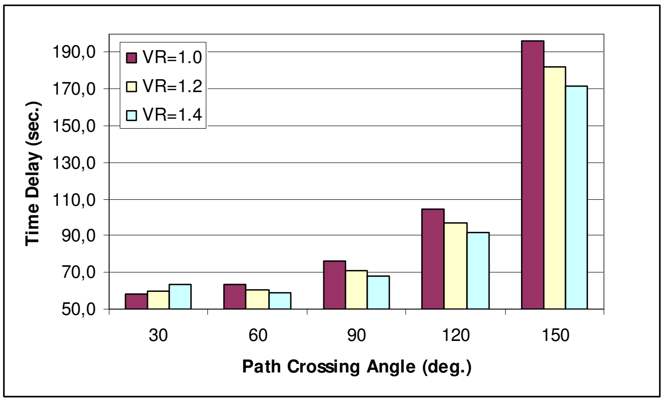


Fig. 5. Variation of total time delay with path crossing angle and initial speed ratio of the aircraft.

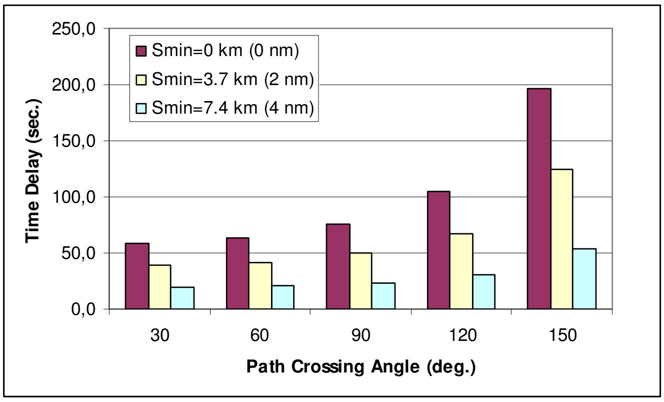


Fig. 6. Variation of total time delay with path crossing angle and minimum horizontal separation before the resolution.

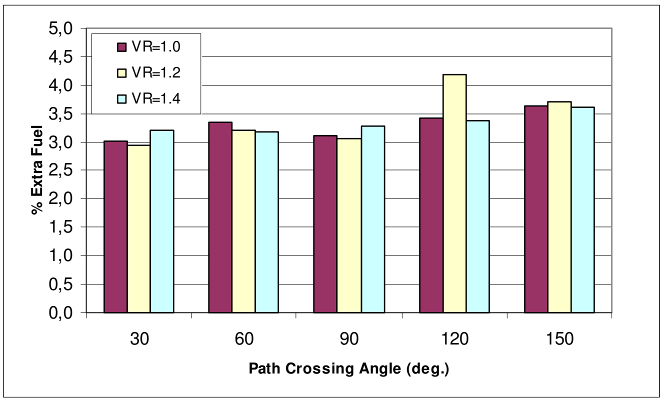


Fig. 7. Variation of % extra fuel consumption with path crossing angle and initial speed ratio of the aircraft.

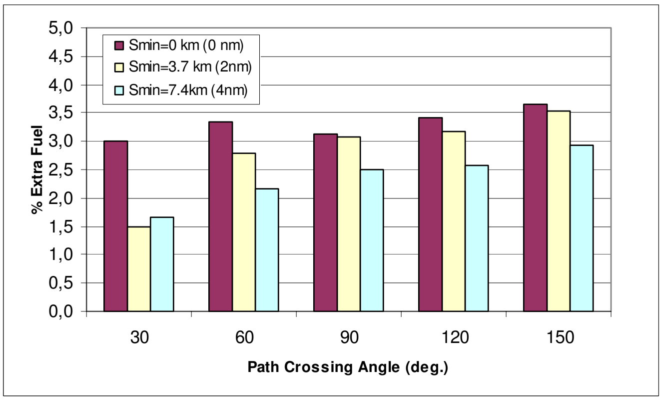


Fig. 8. . Variation of % extra fuel consumption with path crossing angle and minimum horizontal separation before the resolution.

V. CONCLUSION

Minimum time speed change maneuvers have a certain drawback regarding the total resolution time. It takes longer to resolve a conflict especially when the routes of the aircraft intersect with large angles. Therefore using the speed changes for the conflict resolution alone is less preferable than heading changes. On the other hand, it can be helpful to use them combined with heading maneuvers. For the crossing angles less than 90˚, significant cuts in time can be obtained using them. Combined use of speed and heading maneuvers may also help to use less lateral space and therefore increase the capacity of the airspace. These advantages make use of speed change maneuvers for the conflict resolution interesting.

Though this study on speed change strategies focuses on a very restricted case of conflict resolution, further investigation on this topic can provide essential tools for aircraft operators and air traffic system designers to make a more complete economic performance analysis of flight in possible future scenarios.

REFERENCES

[1] http://www.bts.gov/programs/airline_information/indicators/top.html

[2] Eurocontrol, *Medium-Term Forecast Flight Movements (2005-2011) Vol.1*, STATFOR DOC.110.v01, Brussels, Belgium (2005).

[3] Eurocontrol, *Long-Term Forecast Flights (2004-2025) Vol.1,* STATFOR DOC.103.v01, Brussels, Belgium (2005).

[4] Eurocontrol, *Medium-Term Forecast Flight Movements (2005-2011) Vol.2,* STATFOR DOC.110.v02, Brussels, Belgium (2005).

[5] http://www.faa.gov/data_statistics/aviation/aerospace_forecasts/2004-2015/.

[6] RTCA, *Advancing Free Flight Through Human Factors,* Technical Workshop Report, FAA AAR-100 (1995).

[7] Kuchar, J.K. and Yang, L.C., A Review of Conflict Detection and Resolution Modeling Methods, IEEE Transactions on Intelligent Transportation Systems, **1**, 179-189 (2000).

[8] Dimarogonas, D. and Kyriakopoulous, K., *Inventory of Decentralized Conflict Detection and Resolution Systems in Air Traffic*, HYBRIDGE Task 6.1 v.2.0, IST-2001-32460 of European Commission (2003).

[9] Mendoza, M., *Current State of ATC Conflict Resolution*, Eurocontrol Experimental Center, ECC Note No. 12/99, Bretigny, France (1999).

[10] Eurocontrol, *Towards a Controller-based Conflict Resolution Tool-A Literature Review*, CORA 2 Project, ASA.01.CORA2.DEL04-A.LIT, v.01, Brussels, Belgium (2002).

[11] Valenti Clari, M.S.V, Ruigrok, R.C.J., Heesbeen, B.W.M. and Groeneweg, J., *Research Flight Simulation of Future Autonomous Aircraft Operations*, Proceedings of Winter Simulation Conference, San Diego, USA (2002).

[12] Krozel, J. and Peters, M., *Strategic Conflict Detection and Resolution for Free Flight*, Proceedings of the Conference on Decision and Control, San Diego, USA (1997).

[13] McCormick, B.W., *Aerodynamics, Aernautics, and Flight Mechanics*, J. Wiley, New York, USA (1979).

[14] Nuic, A., *User Manual for the Base of Aircraft Data (BADA) Revision 3.5,* Eurocontrol Experimental Center, ECC Note No. 11/03, Brétigny-sur-Orge CEDEX, France (2003).

[15] Etkin, B. and Reid, L.D., *Dynamics of Flight-Stability and Control 3rd Edition*, Wiley and Sons Inc., Toronto, Canada (1996).

[16] GONG, C. and CHAN, W.N., *Using Flight Manual Data to Derive Aeor-Propulsive Models for Predicting Aircraft Trajectories*, AIAA Aircraft Technology., Integration and Operations Conference, Los Angeles, USA (2002).

Cem Cetek . BSME, Middle East Technical University, Turkey in 1994, MSME, Ohio University, 1999. He is a research assistant and PhD candidate at Anadolu University-School of Civil Aviation in Eskisehir, Turkey.

Potential of Speed Control on Flight Efficiency

Thomas Günther, and Hartmut Fricke

Abstract — **Speed control applied in ATC is considered to be a valuable means to reduce controller workload. This paper reveals the potential of speed control onto flight efficiency if applied pre-tactically for delay absorption. It summarizes the dependencies of the operational speed chosen onto the total operational costs for airlines. It is shown, that speed control may improve flight efficiency when demand is exceeding capacity at destination airports. Based on the experiences of ASCAPE (Arrival Sequence Control and Planning Engine), a system using speed control to optimize arrival traffic at congested airports and operationally tested by Lufthansa, the ongoing research at TUD in this field is presented promising improvements for both controller workload and flight efficiency.**

Index Terms — **Air Traffic Flow Management, Delay Sharing, Flight Efficiency, Speed Control**

I. Introduction

THE potential of speed control in ATC has intensively been investigated in recent research studies: Two objectives were identified in this field: Conflict resolution [1], [2] and Traffic Synchronization [3], [4]. Both applications of speed control can create so called “lucky traffic” for controllers as it is prospectively used to resolve aircraft conflicts. The use of pre-tactical speed adjustments by an automation system will lead to less workload and increased airspace capacity [1]. As such, the obvious positive effects of speed control onto controller workload are focused in various researches.

In this context, it must be clarified that effects of speed control onto flight efficiency shall rather be considered as effects onto Total Cost efficiency including time and fuel costs, this being one of the high level objectives defined in the ATM 2000+ Strategy. It clarifies: “Indirect costs for airspace users, which include the costs of ATM-related delays, flight inefficiency and on-board equipment, need to be considered.” [5]. Previous investigations revert to this objective by a subsequent evaluation of differences in fuel consumption. This paper goes one step further and includes flight efficiency as one key objective of speed control based on the experiences of previous studies.

Manuscript received February 26, 2006.

Thomas Günther works at the Institute of Aviation, Chair Air Transport Technology and Logistics at Technische Universität Dresden, Germany, (phone: +49 351 463-36740; e-mail: guenther@ifl.tu-dresden.de).

Hartmut Fricke leads the Institute of Aviation, Chair Air Transport Technology and Logistics at Technische Universität Dresden, Germany, (phone: +49 351 463-36739; e-mail: fricke@ifl.tu-dresden.de).

II. Background

Speed control is almost unused in European upper en-route sector control. This is mainly due to the small potential for ad hoc conflict resolutions because of the limited margin in accelerating or decelerating aircraft. However, speed control used in a (large space) multi-sector-planning environment so acting pre-tactically may achieve conflict resolution rates of about 80 percent [1].

As such, speed control can be identified as one gap filler between current Air Traffic Flow Management on ground and ad hoc Air Traffic Control [14]. A special working position, described as the multi sector planner (MSP), could act as the planning authority for speed control in current ATC environment. MSP will be able to detect conflicts with a look ahead time from 10 up to 60 minutes and recommend conflict resolution advisories [6]. Because of the high success rates for resolutions achieved by speed advisories this seems to be an option for a multi sector planning function. Thereby it is assumed that such pre-tactical traffic organization shall be highly automated, as putting complete responsibility on appropriate MSP controller could lead to an excessive workload especially in high density sectors.

The multi sector planning function has to consider aircraft performance envelopes as they are the major constraint in speed variation of aircraft. Hence, aircraft performance data are introduced in previous studies. In addition some studies set up fixed and constant constraint in speed variances, exemplary +/-15% [1] or +/- 0.02 Mach [4]. Fig. 1 shows the speed envelope of an Airbus A320 (at a gross weight of 45t) based on aircraft performance data as given in [7]. It must be stated, that all data are based on mathematical models and more precise results could prospectively be achieved by using performance data provided by the aircraft manufacturer.

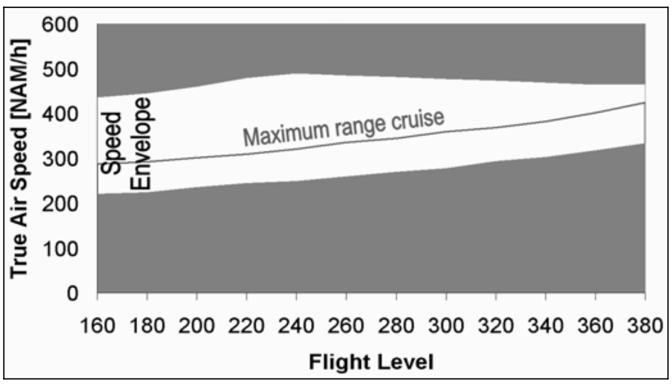


Fig. 1. A320 speed envelope

The acceleration potential regarding maximum range cruise (MRC) speed is decreasing with the altitude but is in FL 380 still about 9%, which corresponds to the results of previous studies [1]. On the other hand, chosen speed by airlines is according to the consideration of time-related costs higher than MRC speed. Hence, for a detailed analysis of the potential for an aircraft to speed up or to slow down, it is necessary to determine the economic cruise (ECON) speed. For this reason chapter 3 gives an overview on aircraft operating costs.

III. Aircraft Operating Costs

Minimum aircraft operating cost is always a trade-off between minimum time costs and minimum fuel consumption. The so called cost index (CI) has been defined for use in the Flight Management System (FMS) to allow presetting that relationship for each mission. In essence, the CI takes into account the relationship between fuel costs (C_{Fuel}) and time-related costs (C_{Time}) with the mathematical expression CI = C_{Time} / C_{Fuel} and unit mostly given in kg/min. ECON speed is always calculated in dependence on the chosen CI. Hence, the CI provides flexibility to control fuel burn and trip time between two extreme cases: A CI = 0 corresponds to MRC speed and a CI = MAX (or practically 999) corresponds to maximum operating speed [8]. Within this study, the CI is always assumed to be set to "30" and fuel costs are estimated as 1.00 EUR/kg. This is equivalent to time-related costs of 30 EUR per minute of flight, which corresponds to the results of previous studies [9].

In the following, total operating costs are calculated exemplarily for an Airbus A320 (at a gross weight of 45t) based on aircraft performance data given in [7] (no wind, no deviation from ISA). Firstly, Fig. 2 shows the dependencies between Mach number and fuel costs respectively total operating costs in Flight Level 240.

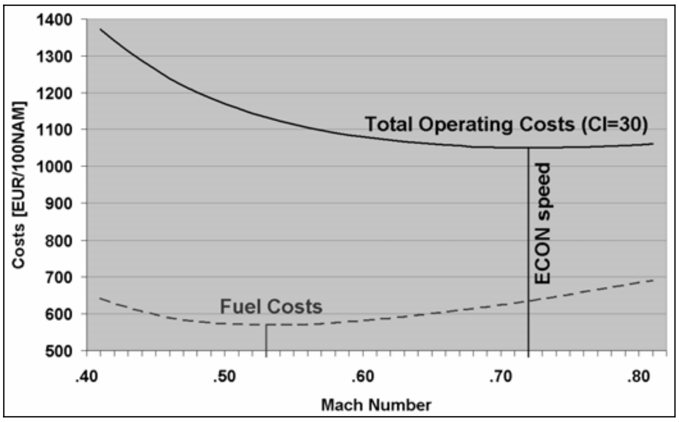


Fig. 2. A320 Operating Costs (FL 240)

Speed envelope covers the range from Mach 0.41 to Mach 0.81. ECON speed is at Mach 0.72, which is due to the consideration of time-related costs significantly higher than MRC speed (Mach 0.53). Nevertheless, it is still less than maximum operating speed and as such the acceleration potential in FL 240 is still about 12 %.

But as the following figures show, ECON speed in higher altitudes is getting closer and perhaps equivalent to the maximum operating speed. Fig. 3 shows this fact exemplary for Flight Level 340 (Airbus A320, gross weight 45t, no wind, no deviation from ISA) [7].

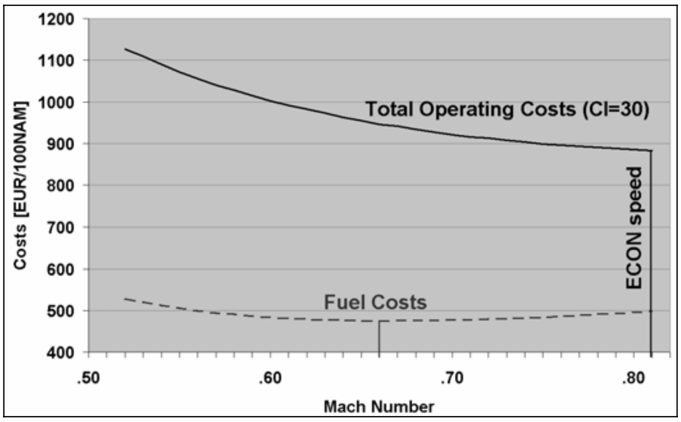


Fig. 3. A320 Operating Costs (FL 340)

Because of the slightly changing fuel consumption, time-related costs have strong impact onto ECON speed, which therefore is equivalent to maximum operating speed (Mach 0.81). This corresponds to operational experiences at Lufthansa, where an increase from ECON speed to maximum operating speed does not lead to a significant acceleration of the aircraft [10]. Both results indicate that in upper Flight Levels the speed chosen by an airline is close to maximum operating speed.

Regarding speed control measures, two conclusions can be drawn at this point: First, speed control advisories in high altitudes can only invoke the deceleration of aircraft, as there is no significant potential in accelerating aircraft. Second, a certain deceleration of aircraft leads to a reduction in fuel burn but to increasing operational costs for airlines due to the increasing time-related costs. This has to be considered, when speed control is applied. Nonetheless, chapter 4 reveals high potential of speed control on flight efficiency in case of pre-tactical delay absorption.

IV. Delay Absorption by Means of Speed Control

Speed advisories are already commonly used for delay absorption at high density airports. In times where demand exceeds capacity, arrival flights are decelerated to avoid congestion in the TMA. This is mostly done with automation support (Arrival Management Systems, AMAN) for ATC controllers. AMAN calculates an optimized arrival sequence and provides controllers with time to lose advisories for each flight in order to manage the optimized sequence over the arrival stack. In dependence on the amount of delay and sectors involved, the AMAN determines how this delay is shared between them. Finally, the specific delay per flight respectively recommended times over stack points are presented to the controllers [11]. Fig. 4 illustrates the principle of delay sharing (source: EUROCONTROL Institute of Air Navigation Services, Training Division, Luxembourg).

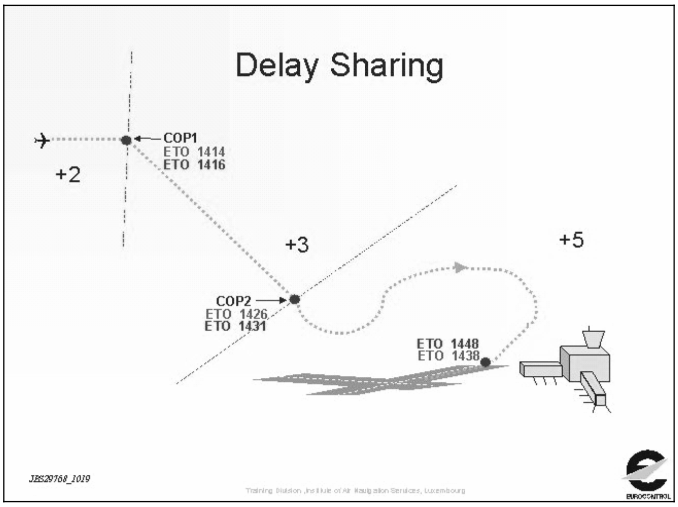


Fig. 4. Delay sharing

Current AMAN horizons are limited to the approach sector and at best to one adjacent ACC sector. This is mainly due to the increasing complexity of controller tasks in case of a further extension. For example at Frankfurt Main Airport this leads to look ahead advisory horizon of about 85 NM. Hence, AMAN enables the controller to minimize deviations from an economic flight profile, but currently this is done within a very limited range.

With regard to flight efficiency it will be useful to apply speed advisories outside current AMAN advisory horizon. In case of expected holding at the destination airport, useful application of speed control improves total operating costs as holding time can be reduced. Fig. 5 shows fuel burn in case of delay absorption by means of speed control in cruise flight. Exemplarily 20 minutes delay absorption is assumed and speed control is applied within a range of 250 NM to the airport. Calculation is based on aircraft performance data for an Airbus A 320 (gross weight 45t, no wind, no deviation from ISA) given in [7] with cruise in FL 340 and holding in FL 100. This is a simplified calculation as no descent profile is considered.

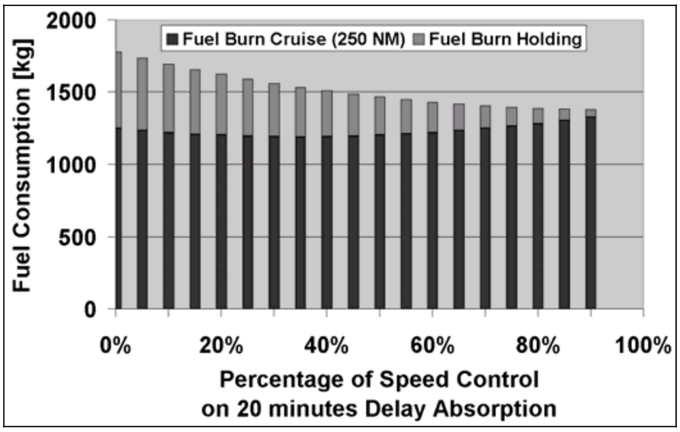


Fig. 5. Fuel consumption used for delay absorption by means of speed control

The application of speed control for delay absorption has two positive effects on fuel consumption: First, fuel burn in cruise (black columns) can be reduced if aircraft is decelerated to minimum fuel speed (see also chapter III). This leads to delay absorption of 7 minutes (35% of total delay absorption) in the presented calculation. A further deceleration consequently leads to increasing fuel consumption in cruise. Second, holding fuel consumption can be reduced as holding time is reduced (grey columns). Concerning time-related costs no negative effect occurs as delay is not increased but displaced. Fig. 6 summarizes the potential on fuel savings in case of delay absorption by means of speed control (no descent profile considered).

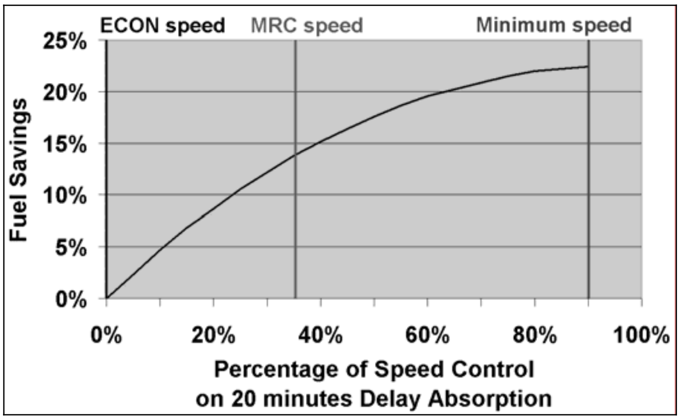


Fig. 6. Speed control potential on fuel savings

If cruising speed is reduced to MRC speed, which correspondents to a flight with a CI of 0, fuel savings are about 14%. Holding time will be reduced by about 35%. Even 22% of fuel can be saved by flying minimum speed, and holding time is reduced by about 90 %. It can be stated that delay absorptions by means of speed control seem to have a remarkable potential for flight efficiency improvements. More detailed investigations would be required to confirm these results, especially the consideration of the descent profile and its effects on fuel consumption.

V. LUFTHANSA SYSTEM ASCAPE

In spring 2004 Lufthansa performed an operational trial to assess the benefits of an initially airline operated system applying speed control within the legal framework of +/- 5% TAS. Objective was the optimization of inbound traffic at congested airports. The system called ASCAPE (Arrival Sequence Control and Planning Engine) communicated via datalink with the cockpit up to 2 hours in advance of the estimated time of arrival (ETA) [12]. ASCAPE was tested in addition to the tactical arrival management at Frankfurt/Main airport. Fig. 7 classifies the operational horizon of ASCAPE:

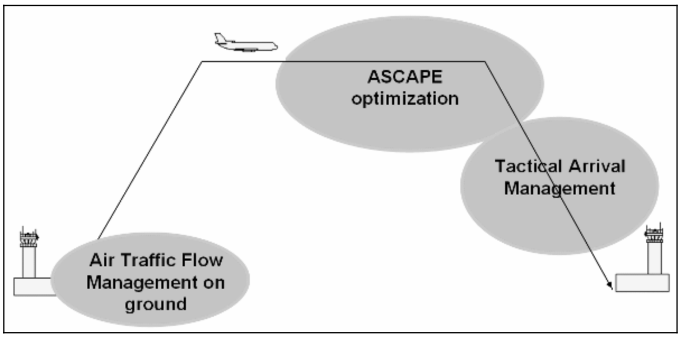


Fig. 7. ASCAPE operational horizon [10]

Key objectives are to improve predictability of estimated times of arrival and to avoid bunching situations at the arrival metering fixes. For this purpose, the ASCAPE process consists of three steps. At a first step, relevant data of all arrival flights are collected and a natural arrival sequence is determined. Thereby data with the highest reliability are used. For example, the estimated time over metering fix, calculated by the FMS, is submitted to ground via the Aircraft Communication Addressing and Reporting System (ACARS).

The second step covers the calculation of an optimized arrival sequence. Outcome is a smoothen sequence with less bunching situations and recommended times of Arrival (RTA) at metering fixes for each participating flight. The times are sent via ACARS to the aircraft and displayed at the Multipurpose Control and Display Unit (MCDU). Finally they will have to be accepted or rejected by the cockpit crew. In case of an acceptance the speed is automatically adopted by the FMS in order to manage the RTA. Thereby a fly-over-precision of +/- 6 seconds can be achieved [13]. As such, current FMS technology enables the achievement of target times and hence the application of speed control. Fig. 8 shows how the RTA recommendation is being displayed on the MCDU.

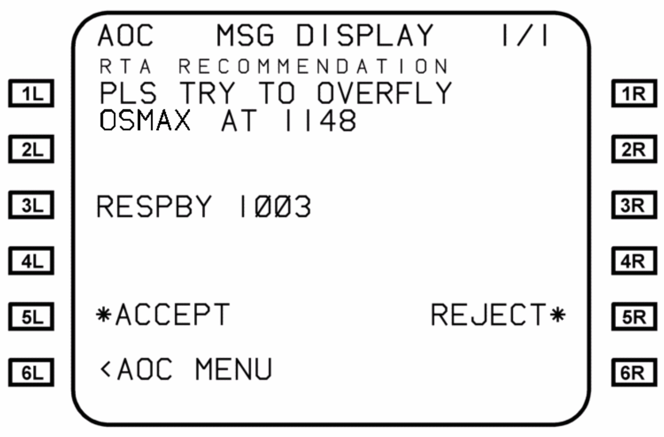


Fig. 8. RTA Recommendation on MCDU

The willingness of the cockpit crews to participate to the ASCAPE trial was more than 90 percent, which demonstrates a deep motivation on pilot's side. Nevertheless, key objectives of ASCAPE could not be achieved completely. This was mainly due to the following reasons:

The system was only tested within Lufthansa. No other airline and especially the German Air Navigation Service Provider (ANSP), DFS, did participate. Consequently, the specific RTA was not known by the controllers and it was not considered when conventional ATC advisories were given. As such, flights were often prevented in managing the RTA due to the actual traffic situation and the corresponding ATC advisories and clearances.

Furthermore the system was activated only to flights with at least 75 minutes of flight time. This, together with the fact that at this stage only Lufthansa flights did participate, lead to a quota of only 25 % of all arrival flights that received an RTA. According to that, the potential in optimizing the arrival traffic was very limited. As a conclusion, it must be stated, that a system like ASCAPE shall rather be operated by the ANSP and not by an airline. If done so, this could lead to much better results due to a coordinated planning with tactical ATC.

Despite the slight success of the operational test phase, ASCAPE demonstrated the technical possibility to apply speed control, underlined the significant interest of airlines in such a system and detected some major tasks to be worked on. Further research in that field can and should revert to the experiences of ASCAPE in order to explore the highest benefit for ATC and airlines when speed control is applied.

VI. CONCLUSIONS AND FURTHER RESEARCH

Previous investigation in speed control was mainly done with the objective to reduce controller workload. This will lead to increased capacity and afford a remarkable contribution to the further progress in air transport. Nevertheless, due to the deviation from the most economic speed restrictions imposed will in most cases lead to less economic flight profiles.

On the other hand, the example of the Lufthansa system

ASCAPE has shown the clear interest of airlines to apply speed control in situations where demand exceeds capacity at the destination airport. As explained in chapter IV, there is a remarkable potential on flight efficiency improvements in case of pre-tactical delay absorptions by means of speed control. It is understood, that this might certainly lead to increasing controller workload due to the rising number of constraints in traffic coordination within the ATC sectors. Consequently, future developments regarding speed control should consider both objectives. Only a common development of a support tool and an integrated speed control algorithm seems to enable improvements both in workload and in flight efficiency.

Hence, our further research on that area will be based on previous studies concerning speed control (e.g. [1] – [4]), revert to our extensive experiences on Multi-Sector Planning functions (e.g. [14]) and consider the experiences of the operational trial of ASCAPE [10]. A first step will be a more detailed analysis of flight efficiency potential. It is supposed, that the consideration of descent profile will decrease the potential of speed control due to less fuel burn. This could imply an investigation on the adaptation of the descent profile and its potential on flight efficiency. Also level changes in cruise flight in order to manage speed adjustments will have to be investigated.

Following, a speed control algorithm reverting to existing results and the integration of optimization criteria considering the situation at the destination airport will have to be developed. To serve this purpose, a simulation platform, consisting of an A320 FAROS Part Task Trainer, pseudo pilot workstations as well as an ATC Simulator (radar and multi-sector-planning controllers), will be used. Fig. 9 gives an overview on the simulation platform.

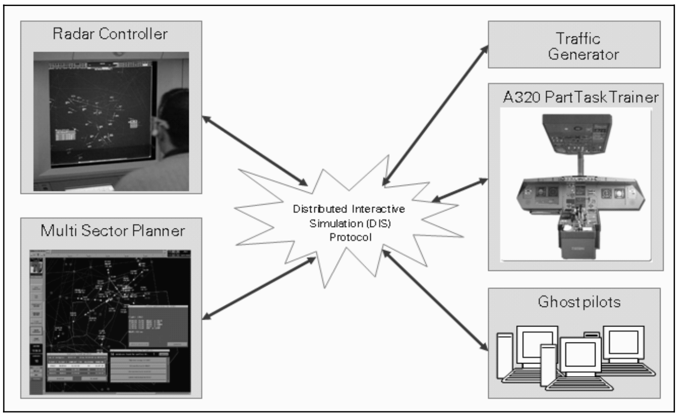


Fig. 9. Simulation platform at Technische Universität Dresden consisting of ground and air components

Supplementary to the development of speed control optimization algorithm, this simulation environment will enable research on data communication/visualization strategies as well as human-in-the-loop investigations. Together with an aspired cooperation with German ANSP (DFS) and Lufthansa this will allow judging the reliability and operational feasibility of the speed control algorithm that will be produced.

REFERENCES

[1] R. Ehrmanntraut, F. Jelinek, *Performance Parameters of Speed Control & the Potential of Lateral Offset*, EUROCONTROL Experimental Centre, EEC Note No. 22/2005, Bretigny-sur-Orge, 2005.

[2] S. Constant, *Minimizing potential conflict quantity with speed control, INRETS*, in Proceedings of the 4th EUROCONTROL Innovative Reasearch Workshop, Dec. 6-8 2005, Bretigny-sur-Orge, 2005.

[3] M. Kupfer, *An Approach for Air Traffic Synchronisation in the Central European Upper Airspace*, Project Thesis, TU Dresden, 2005.

[4] L. Dravecka, *CEATS for traffic synchronisation*, EUROCONTROL CRDS, in Proceedings of the 4th EUROCONTROL Innovative Reasearch Workshop, Dec. 6-8 2005, Bretigny-sur-Orge, 2005.

[5] EUROCONTROL, *Air Traffic Management Strategy fort he years 2000+ (Edition 2003, Volume 2)*, Brussels, 2003.

[6] N.-H. Stark, *Multi Sector Planning in the co-operative Air Traffic Management*, Technische Universität Berlin, 1999.

[7] EUROCONTROL Experimantal Centre, *Base of Aircraft Data (BADA) Version 3.6*, Bretigny-sur-Orge, 2005.

[8] Airbus Costumer Services, *Flight Operations Support & Line Assistance, Getting to grips with the cost index*, Blagnac, 1998.

[9] Institut du Transport Aérien, *Costs of Air Transport Delay in Europe*, Paris, 2000.

[10] T. Günther, *Valididierung der Interdependenzen eines Systems zur Ankunftszeitoptimierung von Luftfahrzeugen in Ergänzung zu einem Arrival Management an einem Verkehrsflughafen*, TU Dresden, 2004.

[11] EUROCONTROL, *Operational Requirements Document for EATCHIP Phase III ATM Added Functions Volume 3 – Arrival Management*, Brussels, 1999.

[12] ATRiCS, *ASCAPE Arrival Sequence Control and Planning Engine*, Freiburg, http://www.atrics.de, 2004.

[13] H. Mikuda, *Analyse eines RTA (Required Time of Arrival) Landeanflugverfahrens für den Flughafen Frankfurt am Main*, RWTH Aachen, 1999.

[14] H. Fricke, *Entwicklung und experimentelle Realisierung einer kooperativen Schnittstelle zwischen Verkehrsflusssteuerung (ATFM) und Flugverkehrskontrolle (ATC)*, Technische Universität Berlin, 1995.

Thomas Günther works at Technische Universität Dresden, Institute of Aviation, Chair Air Transport Technology and Logistics as a scientific assistant since 2004. During several internships at Frankfurt Main Airport, German Aerospace Center DLR and Lufthansa he accumulated experiences in the field of ATM. The evaluation of the system ASCAPE was part of his diploma thesis supervised by Technische Universität Dresden and Deutsche Lufthansa AG.

Hartmut Fricke performs MSP research since 1991. He investigated the interface between Air Traffic Control and ATFM Processes in his PhD Thesis [14], worked on distributed Air Ground Communication Projects with Eurocontrol (Station Keeping, ASAS, Free Flight 2000) and also elaborated the effects of speed control operationally onto Vienna TMA (ATC Benchmark Study for Austrian Airlines in 2001). He leads the Institute of Aviation, Chair Air Transport Technology and Logistics at Technische Universität Dresden since 2001.

Minimizing Potential Conflict Quantity with Speed Control

Sophie Constans, Bastian Fontaine, and Rémy Fondacci

Abstract—**Conflict resolution is a crucial problem in Air Traffic Management, since conflicts are critical situations where passenger safety is at stake. Improved management of these situations is all the most relevant that traffic is continuously increasing, leading to crowded or saturated airspace where conflicts are more and more likely to occur, and controller workload is raising. We present a method for minimizing the potential conflict quantity during a complete traffic day, so that the remaining occurrences are maintained at a level in compliance with low controller workload. Our method is based on an optimal control framework with a sliding horizon loop process. Optimization sub-problems are solved at each iteration, giving optimal arrival times at reference points for all the flights so that a forecast potential conflict quantity is minimized; optimal speed modulations are then determined. The final intent of this work is to implement a real-time closed loop process minimizing conflict risks by dynamically imposing feasible modifications on the velocities of the aircraft.**

Index Terms—**ATM, Conflicts, Linear Programming, Optimal Control, Speed Control.**

I. Introduction

Conflict resolution is one of the key-problems arising in Air Traffic Management. A conflict happens when two aircraft or more get too close to each other, no longer respecting the separation norms, which endangers the passenger safety. These situations are managed by the air traffic controllers as emergency situations and summon up a large part of their attention. Solving a conflict is a heavy task, since it implies to be able to detect future occurrences, and to avoid them by choosing efficient avoidance maneuvers. Air traffic is currently characterized by a continuous increase in the number of operated flights. Teams of controllers are thus in charge of increasing numbers of aircraft, generating even higher conflict risk, and hence have increasing workload. However, in order to do their job properly and ensure a perfect safety to the aircraft and the passengers under their responsibility, they need to be protected from overload.

For this reason, short- or mid-term detection and resolution of conflict is a challenging research topic, as well for researchers, as for public institutions in charge of a safe and efficient flow of air traffic [1]-[5]. The present work is in keeping with this pattern; it is aimed at providing air traffic controllers with traffic situations that are easier to manage, by anticipating and avoiding conflicts thanks to a global multi-aircraft method, considering the entire traffic situation as well as the deposed flight plans, in a sector-less environment. Heading or level change are the procedures most commonly used by air traffic controllers when facing a potential conflict configuration, whereas speed control seems to be seldom used [6]. However, recent studies tend to prove that speed maneuvers have great potential in solving conflict situations [7] and that they could be somehow automatically applied in such a way that controller workload could be lightened [8], with a time horizon slightly greater than that used by the controllers, thus anticipating their action. We choose to focus on the application of speed control to the aircraft to lower a global conflict quantity in spite of perturbations, so that this action remains complementary to further actions made by the controllers.

We aim at minimizing the overall conflict quantity of a traffic day, so that the remaining conflict occurrences are maintained at a level in compliance with low controller workload. Our method consists in a sliding horizon loop process where optimization sub-problems are solved at each iteration. These sub-problems aim at separating the aircraft at best, over the considered horizon, considering the results of the control procedures applied at the previous steps. They are based on the fact that conflicts happen around fixed reference geographical points, and force the aircraft flying a same flight level to cross these points with a time gap $|\Delta T|$ large enough to ensure horizontal separation, thanks to controls applied to cruising flight speeds. Each sub-problem is solved thanks to linear programming techniques, providing us with instructions on the instants the aircraft should cross the reference points situated on their routes. With these set points, speed is determined by a predictive functional controller, which gives more accurate results than open loop computation. Our final goal is to update speed set points every 10 minutes and consider the whole airborne traffic at a time. This approach rather corresponds to a subliminal control approach [8] than to an automatic control one; interactions between this system and the air traffic controllers have to be further investigated.

The next section describes our optimal control approach and the algorithmic framework for the sliding horizon process.

Manuscript received April 26, 2006.

S. Constans, B. Fontaine and R. Fondacci are with the Traffic Engineering Laboratory (LICIT) at the French National Institute for Transport and Safety Research (INRETS), 25, av. François Mitterrand, 69675 Bron cedex, France. (phone: + 33 4 72 14 25 42 or + 33 4 72 14 25 47; fax: + 33 4 72 14 25 50; e-mail: {sophie.constans, bastian.fontaine, remy.fondacci}@inrets.fr).

Section III details the optimization sub-problem to be solved at each supervisory layer iteration, to get optimal crossing times above the reference points, whereas section IV provides a description of the control algorithm used to compute speed variations. Sections III and IV give first results.

II. Optimal Control Approach

Conflict detection and resolution is based on a distributed hierarchical control process made of two layers. The upper layer deals with conflict detection and computation of reference point crossing times so as to avoid conflicts. The lower layer computes speeds for every aircraft involved in a conflict so that they reach their reference points at the times provided by the upper layer.

In the upper layer, the set point decision is an optimization problem which provides optimal controllers with reference point crossing times so that minimum separation is achieved. In the lower layer, there is one controller associated to every aircraft involved in a conflict. As a consequence, the aircraft travel times are controlled independently.

Those two layers have different sampling periods. The sampling period for the supervisory layer is called p_1 and that for the localised control layer is called p_2. The starting values for p_1 and p_2 are set to 10 min and 10 s respectively; they are tuning parameters and still have to be refined. Because maximum speed variations are quite low, the set points should be determined in advance enough so that those variations have an effect on travel times.

Assume t_0 is the current time period beginning in the upper layer. At any iteration, the optimisation problem is solved to lower the conflict quantity between $t_0+x.p_1$ and $t_0+y.p_1$, $x \in \mathbf{N}$, $y \in \mathbf{N}$. We call $H = [x.p_1, y.p_1]$ the optimisation horizon (see Fig. 1). The optimisation problem solution on the whole optimisation horizon or part of it is applied at the start of the next iteration. The x parameter should be chosen so that significant travel time variations are achievable between t_0+p_1 and $t_0+x.p_1$; thus, it should be greater or equal to two. The y parameter depends on prediction uncertainty. At time t_0+p_1, a new optimisation problem is solved to lower conflict quantity between $t_0+p_1+x.p_1$ and $t_0+p_1+y.p_1$. This is a sliding horizon loop process.

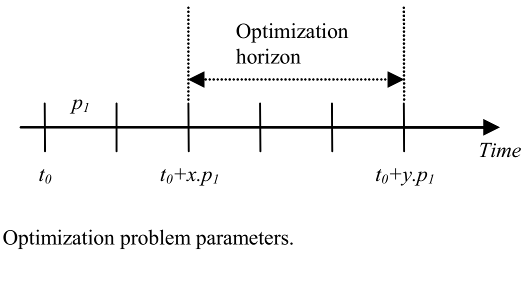


Fig. 1. Optimization problem parameters.

In the lower layer, the optimal controller associated to an aircraft provides speed variations till it reaches the reference point where there is a potential conflict as soon as the aircraft gets on cruise and the optimal travel time is known. The actual arrival time at the reference point should be close enough to the set point to solve a conflict efficiently in spite of disturbances such as unknown wind. The speed variations should be computed so as to take into account the aircraft performances and the computational cost should allow many speed computations simultaneously. Among control techniques, an advanced control technique called Predictive Functional Control (PFC) fulfils the requirements.

III. Optimization Sub-Problem

This section describes the sub-problem posed and solved during an upper layer iteration to minimize a global conflict indicator taking into account the aircraft that are either already airborne or with take off times that can be precisely forecast within a specified time window T_T. This indicator represents a global conflict risk, over the whole airspace considered and for the next T_{OPT} minutes. The boundary condition of a sub-problem is the instantaneous state of the traffic at instant T_{REF}, the beginning of the iteration, making it dependent of the controls previously imposed.

A. Formulation

For two aircraft having intersecting routes, conflict risk is localized at the intersection of the routes; it may be physically materialized by a beacon, or simply be the intersection point of two inter-beacon segments. In the following, beacons and inter-beacon segments intersections are called reference points. For two aircraft sharing a common route or segment, there is no intersection and a conflict may happen anywhere in the common path. However, the relative distance between them along a segment may be expressed as a function of that at the previous or at the next reference point. So, in any case, conflict occurrences can be characterized by the situation of the aircraft around the reference points. The method takes advantage of this by concentrating on reference points, and trying to force separation at these locations. In that way, better satisfaction of the separation distances over the whole considered airspace is ensured.

Since we focus on the situation at specific points and act on the speeds of the aircraft, we choose to use a conflict characterization based on the arrival times of aircraft at these locations. Let us first consider a point where two segments s_1 and s_2 intersect with an angle $\theta \neq 0$, and two flights f_1 and f_2, using s_1 and s_2 respectively. Using the Euclidian distance around the intersection, we derive a simple criterion on the instants when these flights cross the intersection characterizing whether or not there is a conflict between these flights [9]. This characterization indicates that there is a conflict if

$$|\Delta T| < \sqrt{\frac{D_s^2\left(u_1^2 + u_2^2 + 2\alpha u_1 u_2\right)}{u_1^2 u_2^2\left(1-\alpha^2\right)}}\,, \tag{1}$$

where $|\Delta T| = |T_1 - T_2|$ is the time gap between the instants T_1 and T_2 when the flights f_1 and f_2 cross the intersection point, u_1 and u_2 are the velocities of f_1 and f_2, respectively, and $\alpha = \cos\theta$. Then, let us consider two aircraft sharing a common segment.

Equation (1) is not valid any more since $\alpha = 1$. Using the Euclidian distance around one of the extremities of the segment gives the following condition for a conflict to occur:

$$|\Delta T| < \frac{\delta_{12} D_s - (u_1 - u_2)t}{u_2}, \tag{2}$$

for at least one t in the period during which the aircraft share the segment; δ_{12} is equal to 1 if f_1 passes the reference point before f_2, and to -1 otherwise. In both cases, space separation is turned into time separation, according to the characteristics of the aircraft types and of the intersection configuration.

A local potential conflict cost can then be defined for each pair of aircraft having intersecting flight paths, and having close altitudes around the intersection. This cost should model the fact that the conflict is all the most severe that $|\Delta T|$ is small. Noting Γ the right hand side of either (1) or (2), we propose the following local potential conflict cost for two flights above a given reference point:

$$Max(\Gamma - |\Delta T|, 0). \tag{3}$$

This cost is zero as soon as $|\Delta T|$ is large enough to ensure separation, and increases as $|\Delta T|$ decreases. Then, we propose to define the global conflict indicator as the sum of all the local ones, for all the flights and all the reference points.

The variables are the arrival times t of the flights at the reference points. Since we want to adjust speed differently along the flight path of an aircraft, we derive expressions for the travel times between those reference points from the arrival times t. Noting t_{fi} the arrival time of flight f at the i^{th} reference point remaining in its trip and τ_{fi} its travel time from $(i\text{-}1)^{th}$ to i^{th} reference point, we get the relations (4) and (5):

$$t_{fi} = t_{fi-1} + \tau_{fi}, \quad \forall i \neq 1 \tag{4}$$

$$t_{f1} = T_{fREF} + \tau_{f1} \tag{5}$$

T_{fREF} is equal to the reference instant T_{REF} if the aircraft is already airborne, and to the flight take off time otherwise. Equations (4) and (5) give a linear relationship between variables t and quantities τ of the form:

$$Bt_f = \tau_f + T_{fR} \tag{6}$$

where T_{fR} is a vector of zeros, except for its first component, equal to T_{fREF}, and where B is a square matrix such that $b_{ij} = 1$ if $i = j$, $b_{ij} = -1$ if $i\text{-}1 = j$ $(i > 1)$, and $b_{ij} = 0$ otherwise.

Travel time control is made by modifying the τ quantities around their nominal values. These modifications have to be limited, due to aircraft technical characteristics. We define two vectors to limit the range of these modifications for flight f, and name them $\tau^m{}_f$ and $t^M{}_f$, as lower and upper bounds on τ_f.

Finally, from (3) and (6), the definition of B and of the bounds, we are led to the following formulation:

$$\begin{aligned} & Min \sum_{i \in I} \left(\sum_{\substack{f, f' \in F_i \\ f < f'}} Max\left(\Gamma_{ff'i} - \left|t_{fi} - t_{f'i}\right|, 0\right) \right) \\ & \text{s.t.} \quad \tau_f^m + T_{fR} \leq Bt_f \leq \tau_f^M + T_{fR}, \qquad \forall f \in F \end{aligned} \tag{7}$$

where the subscripts in the objective function indicate the index of the reference point where the potential conflict is located, and the indices of the flights involved. I is the set of the reference points and F_i the set of the flights crossing reference point i.

B. Resolution and first tests

We choose to take advantage of the linear part of formulation (7). Imposing the aircraft to cross any reference point with the same order as initially planned leads to a quite simple linear expression for (7), that can then be solved with a linear programming tool. With this simplification, the flight that arrives first at an intersection is still first after the optimization process. It helps solving the problem on the one hand, but on the other, it reduces the solution space and might deprive us of good solutions. A perspective of this work is to implement a lower level resolution tool, for example inspired from the method presented in [10], where we could handle the absolute values directly.

Our first tests concern sample traffic situations appearing at 3, 5 and 7 a.m. when simulating the traffic of September 4, 2003 over Europe. The aircraft taken into account are those already airborne and those having a take off time scheduled within the next 10 minutes ($T_T = 10$ min). The number of flights taken into account in each case is indicated in Table I.

The aircraft trajectories, and particularly the nominal arrival times and nominal travel times are computed using the theoretical performances of the aircraft types considered. The parameters used have been set in the following way:

--the separation distance value is set to 10 NM, which is not statutory, but leaves a margin for controller intervention,

--the optimal travel times between two consecutive reference points have to remain in the range of -10% / +5% of the nominal ones,

--the aircraft crossing a given reference point cross it with the same order as initially planned,

--optimization is made on the travel times corresponding to cruising phases,

--to ease the problem settlement procedure, T_{OPT} is temporarily set to infinity, which means that the optimization horizon corresponds to the furthest landing time.

The procedure is implemented in a C++ code; numerical tests are carried out on a Pentium IV Computer, 3 GHz, 2 Go RAM. The computational time for the largest instance is of 1 minute for the problem settlement, where all the reference points are determined, as well as the nominal travel times and the coefficients of the objective function. The optimization phase, carried out with CPLEX, requires a 2 minute computational time. Notice that most of the operations made during the problem settlement phase can be done once for all at the beginning of the whole process, after reading the flight plans. The numerical results obtained are given in Table I. We remark a relatively small enhancement of the solution, particularly for large instances of the problem. We believe it is due to our choice of T_{OPT}, which is only temporary, and to the constraint added, forcing the aircraft to cross the reference points with the same order as initially planned. Notice also that the coefficients of the objective function are determined for the nominal velocities, and are kept unchanged during the optimization process in spite of the alterations induced on the travel times and hence on the velocities; this might cause difficulties in interpreting the results. Finally, efficiency of the method cannot be fully assessed until integration of the optimization sub-problem in the control loop and simulation tests. Improvement of these numerical results is a perspective of this work, motivated by the good computational times obtained so far.

TABLE I
POTENTIAL CONFLICT COSTS

Traffic sample	Base cost	Optimized cost
Sept. 4, 2003 – 3h (720 flights)	$244{,}6.10^6$	$196{,}2.10^6$
Sept. 4, 2003 – 5h (2244 flights)	$3776{,}3.10^6$	3475.10^6
Sept 4, 2003 – 7h (3168 flights)	$6633{,}7.10^6$	$6223{,}7.10^6$

IV. TRAVEL TIME CONTROL WITH SPEED CHANGES

Controlling a flight travel time between two points of its trajectory by acting on its speed is an attractive way for reducing conflict quantity. Still, aircraft performances heavily restrict travel time variations. Small travel time variations combined with uncertainty in trajectory prediction make conflict solving with speed variations a difficult problem. Trajectory uncertainty is mainly due to unknown wind. Using a closed loop control algorithm is recommended to tackle disturbances and this section is dedicated to show how an optimal control algorithm, Predictive Functional Control, might be used to control a flight travel time between two points of its route by acting on its speed. PFC background is provided first, then the way it is adapted to control travel time is developed, and a trial is finally carried out.

A. Introduction to Predictive Functional Control

The first predictive control strategies were developed in the late 1970's by industrial groups [11], [12]. They had applications in the petrochemical industry. Among advanced control methods, Predictive Functional Control (PFC [13]) enables robust control with low computational cost.

In predictive control, the controller output is computed so that a prediction of the system output gets as close to a reference trajectory as possible over a time window called prediction horizon. A model is required to predict the future system outputs and the control efficiency depends very much on the model quality. Fig. 2 shows the control loop while Fig. 3 provides details on the predictive controller. On Fig. 2 and next, the system output is called y_p. The controller acts on the system input u to change the system output and bring it as close as possible to a reference trajectory that depends on the set point s. The model has the same input as the system. On Fig. 3 and next, the model output is called y_m.

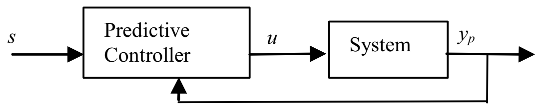


Fig. 2. Control loop.

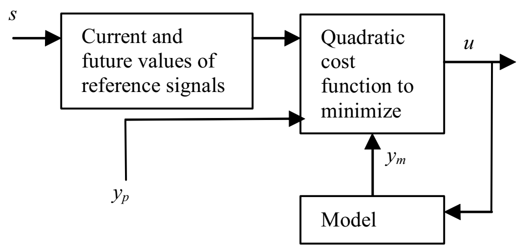


Fig. 3. Predictive controller.

B. Control of Travel Times with PFC

Details are provided next on PFC implementation to control a flight travel time between two points A and B in its trajectory. C gives the current flight position. Only flights on cruise are considered.

The model output is the flight travel time between the two points, assuming that the flight travels at nominal cruise speed from C to B. Its input is the difference between the actual speed and the nominal speed. The nominal speed v_0 is defined as the nominal cruise true airspeed plus the known component of wind speed. It is thus an approximation of the ground speed. The flight speed $v(n)$ used by the model to assess travel time is then defined as:

$$v(n) = v_0 + u(n), \tag{8}$$

where $u(n)$ is the controller output, which goes to model and system inputs.

The model output is defined as the sum of a linear model output y_l and a term to compensate for the difference between the system output and the model output:

$$y_m(n) = y_l(n) + e(n) \tag{9}$$

The linear model output is obtained in the following way: if the flight travels the distance $d(n\text{-}1)$ during time interval $[(n\text{-}1).p,\ n.p[$ the model output at time $n.p$ is equal to the system output at time $(n\text{-}1).p$ plus the advance or minus the delay with respect to the case where there is no control (i.e.

the flight travels at nominal speed):

$$y_l(n) = y_l(n-1) + \frac{d(n-1)}{v_0 + u(n-1)} - \frac{d(n-1)}{v_0} \quad (10)$$

Calling p the sampling period, $d(n\text{-}1)$ might be expressed as:

$$d(n-1) = (v_0 + u(n-1)) \cdot p \quad (11)$$

Replacing (11) into (10) and rearranging yields

$$y_l(n) - y_l(n-1) = -p.\frac{u(n-1)}{v_0} \quad (12)$$

Define the delay operator q^{-1} as

$$q^{-1}.y_l(n) = y_l(n-1) \quad (13)$$

Using the delay operator in (12) leads to the transfer function

$$\frac{y_l(n)}{u(n)} = -\frac{p}{v_0}.\frac{q^{-1}}{1-q^{-1}}. \quad (14)$$

The model described by (14) is an integrator with delay and negative gain and is stable. It has been turned into a state space representation and used to predict the future system outputs. The controller tuning parameters have been chosen so as to get a robust control.

The term to compensate for the difference between the linear model output and the system output is obtained in the following way:

$$e(n) = e(n-1) + p - \frac{d(n-1)}{v_0 + u(n-1)} \quad (15)$$

There are constraints on maximum and minimum speeds as well as maximum acceleration and deceleration. The minimum speed depends on stall speed. The maximum acceleration and deceleration depend on thrust and drag and might be obtained with the total energy model. Those constraints depend on aircraft performance as well as passenger comfort and are currently computed with BADA [14] (Base of Aircraft Data).

C. *Example*

A trial has been carried out with an A320. The aircraft has a cruise flight level of 390 and a cruise true airspeed of 447 KTS. Its travel time is controlled between A and B over a distance of 140 NM. The lower bound for speed is equal to 360 KTS whereas the upper bound is equal to 470 KTS. Maximum acceleration and deceleration are computed on-line as a function of speed. The sampling period is equal to 10 s. Fig. 4 gives the wind speed as a function of the distance to travel. It is a sinusoid with a continuous component of 54 KTS and amplitude of 27 KTS. The weather forecast provides an average wind speed of 59 KTS. As a consequence, the nominal speed used in the model is $v_0 = 447 + 59 = 506$ KTS.

The travel time at nominal speed is equal to 1005 s. The aim is to get a delay of 120 s with respect to the case where there is no control. Then, the set point is set to 1125 s. The travel time obtained with PFC control is equal to 1120 s. The difference between the set point and the travel time obtained is due to the fact that only part of unknown wind is tackled by the method.

The ground speed and the true airspeed are represented on Fig. 5. The computational time of speed is equal to about 2 ms per time sample. The true airspeed is the sum of the nominal cruise speed and controller output. The ground speed is the sum of the true airspeed and wind speed. The true airspeed goes down to increase travel time with respect to the case where the aircraft flies at nominal speed, reaches the lower bound and goes up to nominal cruise speed. As the model is an integrator, constant model output means zero model input. For this reason, the controller output goes back to zero at the end of regulation and the flight goes back to nominal speed. This property is interesting in the sense that the flight cruise speed is usually the most economic one.

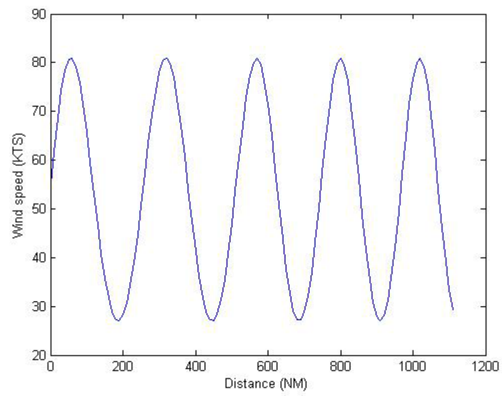


Fig. 4. Wind speed as a function of the distance to travel.

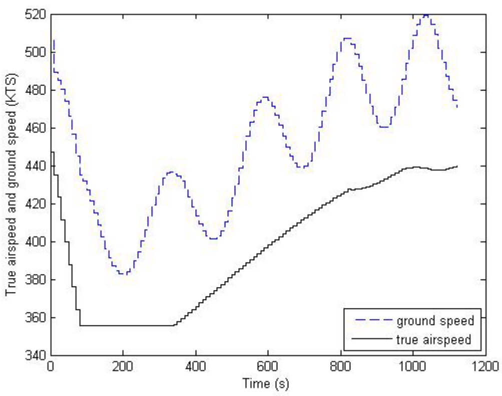


Fig. 5. Flight ground speed and true airspeed.

Further trials with more realistic wind speeds should be carried out to better evaluate the method efficiency. With this method, speed computation should be airborne as it requires too much computation to be ground based. Indeed, higher sampling period would reduce accuracy.

In this section, travel time control is only carried out on cruise. Still, one can imagine extending it to climb and descent by acting either on speed once again or rate of climb and descent or both.

V. CONCLUSION

This paper describes a first work on the problem of minimizing potential conflict quantity with speed control. The aim is to determine feasible velocities alterations so that aircraft having close altitudes cross at intersection points with time gaps sufficiently large to ensure horizontal space separation. The algorithms have to be quick enough to be inserted in a sliding horizon loop procedure, making it possible to compute the speed alterations in real-time, for the whole traffic over Europe, airborne or supposed to take off within a specified time window T_T.

The control framework involves a closed loop control and an optimization sub-problem, aimed at minimizing conflict risk at each supervisory layer iteration, considering the latest information available on the traffic state. The results obtained so far are quite encouraging. On the one hand, the optimization sub-problem is solved with a computational time compatible with a real-time application. On the other, optimization gives an enhancement of the objective function that should be improved thanks to a better definition of our optimization horizon and to an alternative solution algorithm.

Future work is however necessary. First, the optimization phase of the problem has to be enhanced to get better improvement of our objective function, and thus of the traffic conditions the air traffic controllers have to face. Then, this phase has to be integrated in the control loop, so the whole process can be run in a real-time basis. Its boundary conditions are then to be updated at each upper layer iteration, and its results have to be converted into travel time control by acting on speeds. Further trials should be carried out with more realistic wind models and perhaps different travel time control strategies to determine the best way to compute speed depending on accuracy and computational load requirements. These speeds could also be computed on the basis of optimal crossing times obtained for a predefined set of reference points, so that smoother actions are taken, avoiding for example systematical return to nominal cruise speed at these points. Moreover, further enhancements can also be included in the control process. For example, one could impose additional constraints in the optimization sub-problem to link the results obtained at consecutive supervisory layer time steps. In that way, speed controls would be smoothened from one iteration to the next. We could also custom the constraints bounds to make the possible speed alteration depend on the aircraft type, and have more realistic set points, easier to achieve. Besides, adjustment of the parameters is another topic of further investigations. Finally, the results of this procedure should be tested within an air traffic simulation framework.

ACKNOWLEDGMENT

The authors thank Eurocontrol, the European Organization for the Safety of Air Navigation, for providing flight plan and aircraft types performance data.

REFERENCES

[1] N. Durand, and J.-M. Alliot, "Optimal resolution of en route conflicts," in *Proc. 1st Eurocontrol / FAA Air Traffic Management R&D Seminar*, Saclay, France, 1997.
[2] J.K. Kuchar, and L.C. Yang, "Survey of conflict detection and resolution modeling methods," in *Proc. AIAA Guidance, Navigation and Control Conference*, New Orleans, USA, Collection of technical papers (AIAA-97-3732), 1997.
[3] C. Tomlin, G.J. Pappas, and S. Sastry, "Conflict resolution for air traffic management: a study in multiagent hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 509-521, 1998.
[4] G. Dowek, A. Geser, and C. Munoz, "Tactical conflict detection and resolution in a 3-D airspace," in *Proc. 4th USA / Europe Air Traffic Management R&D Seminar*, Santa Fe, USA, 2001.
[5] L. Pallottino, E.M. Feron, and A. Bicchi, "Conflict resolution problems for air traffic management systems solved with mixed integer programming," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 3-11, 2002.
[6] H. Hering, "Technical analysis of ATC controller to pilot voice communication with regards to automatic speech recognition systems," Eurocontrol Experimental Centre, EEC Note No. 01/2001, 2001, Available: www.eurocontrol.int/eec/public/standard_page/2001_note_01.html.
[7] R. Ehrmanntraut, "The potential of speed control," in *Proc. Digital Avionic Systems Conference*, Salt Lake City, USA, 2004.
[8] J. Villiers, "Automatisation du contrôle de la circulation aérienne, "ERASMUS", une voie conviviale pour franchir le mur de la capacité," Collection Etudes & Documents, Institut du Transport Aérien, 2004, 60p.
[9] S. Constans, N.E. El Faouzi, O. Goldschmidt, and R. Fondacci, "Optimal flight level assignment: introducing uncertainty," in *INNOVATIVE RESEARCH Activity Report 2004*, Vu Duong Ed., 2004, pp. 261-276.
[10] H. Arsham, G. Cimperman, N. Damij, T. Damij, and J. Grad, "A computer implementation of the Pull-and-Push algorithm and its computational comparison with simple LP simplex method," *Applied Mathematics and Computation*, vol. 170, pp. 36-63, 2005.
[11] J. Richalet, A. Rault, J. L. Testud, and J. Papon, "Model based heuristic control: application to industrial processes," *Automatica*, vol. 14, pp. 413-428, 1978.
[12] C. R. Cutler, B. L. Ramker, "Dynamic matrix control- a computer control algorithm," in *Proc. AIChE Annual Meeting*, Houston, Texas, USA, 1979.
[13] J. Richalet, *Pratique de la Commande Prédictive*, Hermès, 1993, pp. 121-128.
[14] A. Nuic, "BADA Version 3.6," Eurocontrol Experimental Centre, 2004, Available: www.eurocontrol.fr/projets/bada/

Session 6

Innovative Ideas

Developing Visualizations to support Spatial-Temporal Reasoning in ATC

Simone Rozzi, William Wong, Peter Woodward, Paola Amaldi, Bob Fields, Emanuele Panizzi, Alessio Malizia, Alessandro Boccalatte, Antonio Monteleone, Luigi Mazzuchelli

Abstract: **In this paper we present research into the development of display visualizations to support spatial-temporal reasoning in real-time air traffic control tasks. We refer to such displays as 4D displays, incorporating both visualizations of 3D space and time. This work is based on ethnographic observation studies, contextual inquiry-type studies, as well as in-depth cognitive task analyses, of air traffic controllers in two different ATC domains: a tower control room and a terminal area control centre. Guided by human factors principles, we identified a number of visualization display requirements and a framework for considering the ATC controller's spatial-temporal reasoning needs. We then applied an iterative approach to the design of candidate visualization prototypes, to demonstrate how temporal decision requirements can be included in spatial representations, and how such 4D spatial-temporal representations can be embedded within the familiar 2D Plan Picture Indicator type of radar display. The designs were guided by human factors principles such as Ecological Interface Design.**

Index Terms — **Display Design, 4D HMI, Spatial Temporal Reasoning, Approach and Tower Control.**

I. INTRODUCTION

IN this paper, we report on a case which describes the insights we have gained from the development of 4D user interfaces for Air Traffic Control (ATC). A 4D display is a one that integrates and renders 3D-space and time information into a visual representation. In ATC, controllers are regularly called upon to coordinate and control the movements of aircraft not just in 3D space, but also over a period of time. They frequently need to perform mental calculations to project ahead in time aircraft locations (3D space) and to assess the likelihood of conflicts. Performing such spatial-temporal reasoning in an expert manner as demonstrated by the ATC controllers we studied, requires a high level of training and experience. Being able to 'perceive' the 3D space, and to then develop a good 'mental picture' of the activities of the aircraft, projecting ahead future traffic situations, and including other demands within that 3D space (such as weather changes) within that 3D space, and over different time bands [4], is essential to maintaining good situational awareness or SA [8]. Situation Awareness is crucial for ATC controllers to make good informed decisions [7]. ATC controllers, therefore assess the current situation, anticipate likely future situations in an appropriate time band, and then developing plans to deal with the new situations that could develop.

Other research in air traffic transportation has been concentrated in the definition of 3D types of information. One of the main advantages of 3D visualization is its capacity to accommodate more information than practical in a traditional 2D display, such as slope distances between aircraft, and pictorially realistic renderings of terrain such as mountain ranges. 3D displays have also been found to make the understanding of climbing and descending traffic easier than with only a 2D radar display which requires the reading and interpretation of numerical data in terms of 3D space [15].

On the other hand several arguments exist against the use of 3D in air traffic control. According to Smallman and his colleagues [13], positions and altitudes of aircraft are difficult to understand without some artificial enhancement, such as shadows or drop line; detection of heading and pitch can be confusing unless you look very closely; finally recognition or even detection of objects is difficult because of infelicitous viewing angles. Smallman and St. John in 2005 [12] also goes further to explain that the pictorial realism often seen in 3D visualisations created by readily available digital terrain maps, or actual aircraft models with airline livery markings, can be counter-productive. The excessive detail, which they refer to as naïve realism, tends to distract and make it harder for the observer to extract necessary information, for example, size of the pictorially realistic airliner is often portrayed larger in proportion in relation to the ground features, or the indication of the direction of travel and the attitude (nose up or nose down, and roll) of the airliner. This level of realism makes it difficult to accurately extract relevant information (e.g. size and position of aircraft over ground), or raises the expectation

Manuscript received 26 February 2006. This work was supported in part by the European Commission's Sixth Framework Programme, Priority 1.4 Aeronautics and Space, "AD4: 4D Virtual Airspace Management System" Project Contract No. 012328.

S. Rozzi, W. Wong, P. Woodward, P. Amaldi and B. Fields are with the Interaction Design Center, Middlesex University, London, UK (tel: 0044 +44 (0)20 8411 610; fax: +44 (0)20 8411 5216; e-mail: s.rozzi, w.wong, p.woodward, p.amaldi-trillo, b.fields@mdx.ac.uk).

E. Panizzi and A. Malizia are with Department of Computing Science, University of Rome "La Sapienza", Rome, Italy (e-mail: panizzi, malizia@di.uniroma1.it).

A. Boccalatte is with Space Application Services S.A/N.V, Brussels, Belgium (e-mail: ab@spaceapplications.com).

A.Monteleone and L. Mazzucchelli are with Next Ingegneria dei Sistemi, Rome, Italy (e-mail: antonio.monteleone, luigi.mazzucchelli@next.it).

of providing information at a level of precision it is not capable of (nose up or down information or roll information).

Based on the display technologies of the time, Wickens and May [15] have also found that a planar view allows better judgments about time to collision with respect to a perspective 3D view. This is highly plausible due to the distortion of horizontal distances when viewed in a 3D perspective view.

Another important issue is the transfer of experience. It appears that a negative transfer of experience could result for an ATC controller when he or she moves to a 3D display as their experience is based on the 2D radar display [15]. This suggests, perhaps, the need to maintain a 2D frame of reference of the visual field of view for the ATC controller within which to view a 3D or 4D display. The process of developing such a design is discussed later in this paper.

Thus, it appears that there are many issues still open with the design of a new 3D and 4D system for air traffic control. We present in this paper a user-driven, iterative design approach to a development of 4D interface for air traffic. In this approach we also investigated the way time could be accommodated in a 3D system. For example, we translated the dimension of time into a spatial parameter of distance, and the development of techniques for visualising different time values and time reference points.

The paper presents a description of the 4D visualisation system developed for the AD4 project. We will outline the methodology we had undertaken, and then to report on our findings, in particular, a brief description of the spatial temporal nature of the tasks of the air traffic controllers and its representation by the spatial-temporal framework. We will then describe the development of the 3D and then 4D designs, and finally discuss some of the lessons we have learnt about the creation of such visualisation interfaces.

II. Methodology

This section briefly reports on the methodology used to understand the nature of the ATC controllers' cognitive work in order to define a novel 4D interface for air traffic control. For this project, we conducted a field study at two locations: the approach control at the Ciampino Area Control Centre in Rome, and in the tower control room at the Fiumicino Airport Control Tower, also in Rome.

Field Study and Data Collection. The work analysis of the approach control was conducted over two periods of three days each. During that time, we conducted about 15 contextual inquiry-type interviews [3] and this resulted in about 15 hours of recordings on audio tapes.

We also conducted four sets of ethnographic observations of a group of controllers working together in their natural settings at their work stations; and another set of four observations of controllers working individually, also in their natural settings, at their work stations. This resulted in six hours of video tape recordings showing controllers' performance of various routine collaborative and individual activities in the approach control area.

We had also conducted three individual in-depth cognitive task analysis interviews using a technique called the Critical Decision Method, CDM [10], to identify and to probe more deeply into the individual cognitive strategies invoked by the ATC controllers during critical and non-routine situations. These CDM interviews resulted in another three hours of audio tape. We also conducted five unstructured interviews with the controllers.

At the control tower in Fiumicino, we interviewed all a total of 15 controllers, five of which were based on the Critical Incidents relating to a safety critical scenario; and a further 10 Contextual Inquiry and observations of controllers who were engaged at their work stations. In these occasions the controller explained their work strategies. This was done over a total of eight days in two periods. During this field study we collected more than 500 pictures and 80 hours of extensive ethnographic observation of controllers showing the different aspects of real-time collaborative actions and information handling in ATC episodes; and we also studied 10 Critical Incidents experienced by controllers in the Tower to understand the demands that these incidents placed on the controller.

Data analysis. We used a qualitative data analysis approach to make sense of the data we collected from the field study. The video and audio recordings were transcribed (from the Italian) and analysed using the Emergent Themes Approach, ETA [16]. This method broadly involves the identification of common themes of user behaviours, needs and strategies, which arise from different controllers dealing with different incidents and situations. The ETA provides a process for systematically collating evidence from different incidents studied and people interviewed, in order to understand the extent to which various types of behaviours and strategies exist to deal with the variety of circumstances that are likely to confront the ATC controller, under both routine and non-routine situations.

Subsequently these descriptions were clustered into operational scenarios, i.e. a scenario representative of challenging situations for the user, such as when de-conflicting multiple aircraft trajectories when attempting to change the route an aircraft was originally intended to take. These exemplar operational scenarios are likely to contain episodes that reflect error prone situations, or cognitively demanding mental computations. These operational scenarios serve as a basis from which to derive the information and visualisation requirements. Through our investigations, we identified 13 different but representative operational scenarios for approach control activities, and another 11 operational scenarios for tower control activities [1]. For the purpose of this paper, we will only cite selected extracts to illustrate the data analysis process.

Based on these operational scenarios, we used an iterative design process to sketch the earlier 4D concepts of the user interfaces. The process of sketching provided a means of working with the other designers and developers on the project, and with a number of users from the ATC centres, to articulate the design concepts, clarify the spatial-temporal representations, and to understand how the designs could be interpreted in a real operational context. These sketches and the process of use have been reported elsewhere [11]. These considerations have been encapsulated in the candidate designs presented in the later part of this paper.

These scenarios have been further analysed to identify the spatial-temporal characteristics of the ATC task. The nature of these characteristics and their relationships to each other have been developed into a framework for considering the visual design from a 3D-space and time perspective. In the following sections, for reasons of brevity, we will not provide a detailed description of the thematic collation and concept distillation, but will instead only describe the key outcomes that are relevant to this paper.

III. Controller's challenge

Real time control of air traffic is a very complex activity. This in fact entails the control of a highly variable process whose variables can combine in several ways, thereby determining several possible levels of difficulty of traffic. Although the complexity and the variability of the control tasks involved in such activity, it is possible to identify at least a demanding task whose computation is mandatory in order to construct safe traffic: the detection and prevention of conflict- the controller needs to ensure that aircraft will never be closer than 100 feet vertically and 3 miles horizontally.

The conflict detection activity is probably the most important task required of an air traffic controller. Kirwan [9] proposed a model which describes controllers' conflict resolution process. Overall this can be summarised in three phases:

- *Analysis and characterization of the conflict*: In this phase the controller assesses the current situation and projects what will happen in the near future, where, with which aircraft, when, how close and in which sector;
- *Determination of the physically possible solutions*: The controller works out a solution for a resolving a conflict situation considering the ATC constraints such as airspace constraints, aircraft performance, and sequencing aspects of arrival;
- *Selection of an optimal solution for the current situation*: The controller then selects an optimal solution considering safety, controller's own workload, fairness to aircraft. Controllers according to Kirwan have a set of guiding principles to balance these aspects.

A. Estimation of separation

1) Scanning for likely conflicts in time

Controllers continuously scan the radar display in order to verify the existence of possible conflicts or intersection points, i.e. points where the trajectories of two or more aircraft are going to intersect. In one of the investigated episodes a controller was monitoring a low traffic situation in Rome Approach control when he noticed two aircraft, one at the top of the radar display and another at the bottom. Although they were very far apart, the controller noticed that their trajectories were converging and therefore could intersect at some point in the immediate future:

> *C: you see the A1489, this is at FL 240 (is climbing) and has no traffic problems…uhmm… no wait…the Ryan is problematic which is at FL 278 and is descending… ok the distance is large and you will have no problems there, but he (pointing to the colleague in charge of the traffic at the moment) has always to monitor overall traffic.*

The controller therefore has to be aware of this potential conflict and bear it in mind when planning and sequencing the traffics.

2) Grouping and scanning

Controllers also consider aircraft in terms of *groups* of aircraft. In another episode an aircraft had just taken off from Fiumicino airport, and the controller could vector the departing aircraft along a 'short cut' to reach its target en route flight level more quickly. During his routine scanning for conflicts, he noticed that there were two other groups of aircraft, each with three aircraft flying loosely together, whose trajectories would have crossed the trajectory of the departing aircraft if it had token the intended short-cut:

> *C: …this aircraft can go along the standard route, or I can take it along this route. But here I must ask the coordinator; these three aircraft (referring to the first group of crossing aircraft) make it more difficult. You must always consider also this descending traffic (referring to the second group of aircraft).*

This is an example of controller grouping, or clustering, and has also been observed and explained in a previous work [2] which pointed out that controllers usually distinguish between two groups of aircraft in general: those who are likely to be in conflict with the aircraft that need to be monitored, and those that do not. Although the aircraft may be monitored in groups, the conflict assessments must still be performed for each corresponding pair of aircraft.

3) Assessment of separation in time and space

With reference to the example given above, although the paths of the two groups of aircraft will intersect with that of the departing aircraft, the controller also has to assess whether this interception of trajectories will occur at the times when the different aircraft are at the respective intersecting points in space. Once evaluated the existence of a possible crossing point in time and space, the controller evaluate whether there the separation exists. This assessment is achieved by projecting mentally the position of the aircraft in space, and to evaluating how close the aircraft are in that point, i.e. evaluating the separation between the aircraft at the intersection point.

C:First I ask whether an aircraft will be there in three minutes, then I ask myself: who is going to be here in three minutes? At which altitude?

Once the controller has imagined the position of the aircraft in the immediate future, the controller asks himself which other aircraft is likely to be nearby. Note that the controller here also mentions the need to know the altitude of the aircraft, as aircraft move both horizontally and vertically, climbing or descending, as well as flying straight and level. Therefore a projection of the future position of a group of aircraft must be achieved along both these two dimensions. However, while a projection of the horizontal position of an aircraft can be carried out by imagining the position from a plan view of the 2D radar, the vertical position requires a mental calculation as it cannot be inferred visually by the same display. In particular the controller must make the following calculation, mentally:

(Rate of Climb x Time to Crossing Point)+Current Altitude=Expected Altitude in the Crossing Point

In this way the controller estimates the vertical position of the aircraft in a given point and assesses whether the separation between the aircraft are within the safety minimas. Furthermore it is noteworthy that the controllers are not aware of aircraft rate of climb, and they must acquire this data from the pilots at the same time that they need to assess the existence of vertical separations. The following episode refers to an interview where a controller was describing the computation he usually makes when controlling two independent parallel approaches. In this operation the controller must verify that two aircraft will reach the final approach fix with acceptable separations:

C: For us it is important the separation…but we have not the climb and descent rate, if I can avoid this computation…I would save time!

Another assessment of aircraft positions projected in the near future has been observed to occur in tower control. In this control centre often the ground controller has to take an aircraft that is about to miss its slot (scheduled time of departure) from its parking bay to the runway. Along this path there are some taxiways and some junctions. When this aircraft approaches these junctions, the controller needs to know whether the aircraft that is about to miss its slot could safely pass without interfering with the other planned aircraft movements, i.e. no other aircraft crossing the junction at the same time. This assessment is necessary so to avoid to stopping the aircraft along the taxiways. In order to do so, he identifies any other aircraft that are likely to converge in the same junction, and whether it was possible to stop them before entering the junction in question.

B.3D picture in the head

As the examples and other research (for example see [8]) demonstrate, controllers were observed to develop and maintain a mental picture of the traffic situation. In the approach sector, controllers must de-conflict traffic that are moving both vertically and horizontally in the air space while converging towards the same point in space, i.e. the airport. As a result "aircraft are in this phase of flight at their closest proximity to one another than at any other point during their flight" [6].

Fig.1. Controller's gesticulation emphasizes the nature the aircraft maneuver in 3D space.

During the description of a vectoring action of an outbound aircraft, a controller explained how his colleague was ensuring the separation. He was allowing the departing aircraft to climb towards the en route level, while climbing over the inbound aircraft that was passing below.

C: "The Delta is in bound, hence the departure has to climb over the in bound, and then he can take it (the aircraft) in (towards the airport)…He is separating it from the descending traffic"

It is noteworthy that such descriptions of how separations are maintained were frequently accompanied by gestural explanation, such as in Fig. 1. This highlights the 3D configuration of the space and their mental model of the space that they operate in, using terms like to fly through, above or below a volume of (restricted) airspace:

C:Cumulus nimbi, military restricted airspaces can combine in different ways in the airspace, and for me is important to know where the aircraft can fly, as it can fly below or above but not through"

As the airspace in which the aircraft is allowed to fly in is of a limited volume of 3D-space, the controller will need to find transition spaces. These transition spaces are often narrow and offering limited room to maneuver. Therefore to de-conflict a situation, controllers need to have a good understanding of the 3D configuration of the space and terrain they are operating in, i.e. spaces they can or cannot go into, or by how much they can venture into restricted airspaces.

C: In that zone and the zone where the aircraft was flying towards, the minimum altitude was higher. One of the things which restrained me from taking the aircraft on the left was the much higher safety *minimum altitude, where really I had no guarantee whether the aircraft would crashed or not. In that zone I though that I had 3-400 meters (below the safety minimas)*

This controller describes the time constraints and the 3D-spatial considerations he had to make when he had two aircraft converging at the same point in space and time, due to bad weather.

C: ...it is easy to be in a complex situation when weather is bad. It is inevitable that aircraft fly in very narrow portion of air space.... They hand you an aircraft and you realize that this is exactly at the same flight level as another one in front of it. This is an extremely dangerous situation, because ... the aircraft reduce the space quickly (between each other) . There were no solutions, moving the aircraft aside would had implied other conflicts with other aircraft. ... "and now what can I do?!" ... and there my instinct told me put the aircraft below safety minimum altitude. ... I made the aircraft descending, they cross each other, then one had to go in a direction and this one (the one with whom the controller was in contact) had to entered a sequence for Ciampino.

With only about two minutes before a collision would have occurred, and because he had a good mental model of the 3D configuration of the area, and therefore how much altitude he had available, he descended the aircraft that he was just handed, below the minimum safety altitude for the terrain in the area, in order to avoid a collision and to avoid conflict with other aircraft operating nearby.

IV. Controller's Visualisation Requirements

From the example scenarios cited above and others, we identified a number of visualisation and representation design requirements which were used to drive the design of the 4D displays. Four directly relevant ones have been extracted from our report [1] and are summarized below for illustrative purposes:

a. Show the point of intersection between converging aircraft in 3D space on demand and within the context of the 2D radar plan picture indicator type of display.
b. Show and make visually salient, the estimated vertical, horizontal and angular separations between converging aircraft expected at the predicted intersection point to improve separation assessment.
c. Visualize the rate of climb (or descent) of the aircraft, which in this way make it available to the controller an important piece of data that today is only available from the pilot.
d. Show the heading of the aircraft and whether the aircraft is navigating on an assigned flight level, or is climbing/descending to a new FL, and therefore indicate the future position of the aircraft.

Having understood the visualisation needs of the ATC controller in terms of spatial-temporal reasoning, we then identified a total of 11 design recommendations for translating the visualisation needs for spatial-temporal reasoning into practical and implementable 4D user interface designs, for both the approach and tower control environments. Some of these recommendations were concerned with not restricting collaboration and shared awareness, the capacity to handle high workloads with multiple parallel activities, readability of the display.

The recommendation most relevant to this paper is that of integrating the 4D-view within the context of the existing 2D-radar PPI type of view. In follow-up discussions with various ATC controllers on the subject of a possible way of presenting the 4D display, it became apparent that despite the limitations of the current 2D radar displays, it is capable of providing good information about all the activities in the sector they are responsible for controlling. Any new display, such as the 4D display being proposed, must not detract from this. Any new display must instead augment their ability to control, coordinate and de-conflict the air traffic. Presentation options considered include having a separate display for the 4D display located immediately beside the 2D radar display, or to have the 4D view pop-up as a 'window' over the 2D display.

In the first option, the beside view, the controller would need to shift his or her locus of attention from the 2D radar display to the 4D view, orientate him/herself to the information presented in a different perspective, and then to revert back to the 2D display. This shift of attention was considered a problem by controllers, as had been described in an earlier example operational scenario, it is possible to lose track or miss important changes to the traffic situation, while one is attending to something else even for a very short moment.

In the second option where the 4D view would pop up as a window within the same field of view as the 2D radar display, the controllers considered this to be also a risky option as the 4D window would obscure possibly information located behind it on the 2D display. This would interfere with the way they work and would restrict the controller's access to vital air traffic situation information being presented in the 2D radar display.

A final option considered was the concept of a "picture within a picture", i.e. to visually and interactively integrate the 4D view with the 2D radar view. In this way, the 2D radar display will still present information about the overall traffic situation and that no part of this situation will be obscured by the integrated 4D perspective picture within a 2D planar picture view. A number of problems can immediately be expected with this design, such as boundary transitions, i.e. visual continuity of visual elements between the 4D perspective space and that of the 2D planar space. These and other related issues are being investigated further, some of which will be reported later in this paper.

Thus, to make a visualisation work, designers and developers have to consider more than just the visualisation of the information, but how that visualisation would interact with work practices, and related information needs.

V. VISUALIZING 4D INFORMATION

The type of cognitive activities mentioned above implies a controller's awareness of both spatial and temporal information (ST) simultaneously. Such a framework, the ST Framework, has been proposed and described in greater detail elsewhere [11], and reproduced in Fig. 2. The ST framework is based on principles from Ecological Interface Design [14, 5], particularly the visualisation of constraints and of key functional relationships between operational entities, functions, processes and purposes of the system. These principles of design were used to guide the visualization and representation of important informational relationships within the new 4D display.

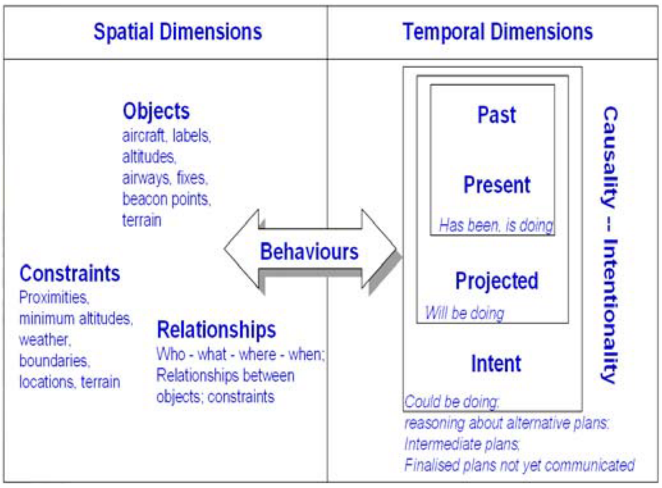


Fig. 2. The Spatial-Temporal Framework [11].

The ST Framework proposes that the work domain can be described by a spatial dimension and by a temporal dimension. The spatial dimension is static. Objects exist within the 3D space, but only exist or move in relation to time. These objects have customarily been represented as aircraft, labels, altitude information, beacon points, terrain, airways and airports. These objects operate with given constraints, some of which are hard constraints such as an aircraft cannot fly into a mountainside, and others are softer, operational ones, such as the declaration of a restricted airspace for duration for military use, or safety minimas between aircraft. However, what is important to the ATC controller are the positions and activities of the objects and the relationships that exist among them and the constraints.

This then introduces the temporal dimension. The objects, constraints, and their relationships can be described in terms of their past, e.g. track history, indicating what these objects have been doing, or an aircraft current or present heading and speed, and if it continues in this manner, its current behaviour, one can project into the future, where this aircraft is likely to be, or likely to be doing. Usually in most modern ATC systems, such information is readily available in the system. Such information has been used for a variety of things such as short term conflict detection and warning. However, a set of temporal information that is less readily available is that of intentions, of both the pilots and the controllers. The operational scenario of the short cut briefly described earlier is an example of intentional information. However, what is important to the controllers is not the separate dimensions, but the interaction of the objects, constraints, and relationships, over time. This is the resulting or emergent behaviors which are what the controller is required to manage and coordinate.

Thus, this framework identifies the information that need to be computed collectively by the controllers, and as such it provides a framework within which to consider the spatial and temporal information needed to be represented in a 4D visualisation that would help controllers develop and maintain an integrated mental model of activities or traffic behaviours in 3D space over time.

VI. 4D VISUALISATION FOR TOWER CONTROL

The human-machine interface (HMI) 4D visualisation prototyped for the control tower consists of a display providing the controller with a 3D view of the airport area, in particular, looking at the airport from a standpoint located at a convenient altitude above the ground, so that the camera field of view can cover the whole airfield. The controller can however leave the default viewpoint by switching to other significant predefined views, focusing on a few critical portions of the airfield ground, such as taxiway junctions and runway thresholds.

The objects of an airport domain include all the relevant ground elements, such as buildings, taxiways, junctions, parking areas and runways, with constraints that include the indication of runway or taxiway centrelines, stop-bars, borders and other restricted areas.

Airplanes which are moving on the ground (i.e. rolling) can be represented by flat 2D geometry, while flying airplanes can be represented by a simplified 3D shape (see Fig. 3 and 4).

The use of 2D shapes for rolling airplanes is justified by the fact that ground traffic is an implicitly bi-dimensional problem, while the 3rd spatial dimension (the height of aircraft) does not convey any useful information to controllers at the present. It is important to stress that the critical information about the objects of the airport domain are related to the position and direction of each airplane, its size and weight category and the actual area it occupied by the wing span and the position of nose and tail. This is so that the controller can readily estimate the spatial or horizontal separation between any two airplanes.

Thus it becomes possible to allocate additional visual cues to represent critical time-relevant (4th dimension) information in the 3rd dimension of the representation.

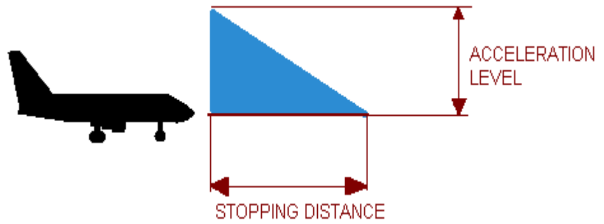


Fig. 3. Illustration of the triangle concept for visualising stopping distance and acceleration.

The vertical dimension of the visual representation in this example can be used to include time-related information. Aeroplanes are provided with an indication of the estimated stopping distance, i.e. the distance within which an aircraft is able to completely stop, according to its current speed, and the intensity of the acceleration or deceleration along the direction of motion. Stopping distance is visualized by a horizontal line, starting at the nose of the plane and extending to the estimated stop point, while acceleration is represented by a vertical line whose length is proportional to the acceleration magnitude, pointing up for positive values and down for negative values. Stopping distance and acceleration are merged together as a translucent triangle in a synthetic representation. This is illustrated in Fig. 3 and 4.

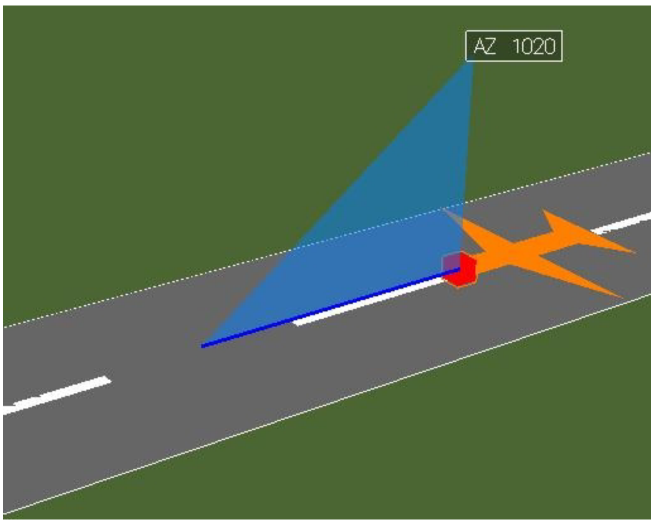


Fig. 4. A 3D view of the stopping-distance and acceleration triangle.

Displaying stopping distance is helpful whenever controllers need to predict early, and possibly avoid, a potential collision between aircraft. The visualisation of time in terms of the behaviours relevant to the spatial-temporal reasoning problem – will the aircraft be able to stop in time? – will off-load the cognitively demanding computations, to a more efficient perceptual task of visual comparison. Similarly, visualization of acceleration is helpful in understanding the pilots' intentions and to therefore foresee near future situations, such as, is the aircraft accelerating? is it going to take off? is it braking and going to stop?

Another aspect of operations that is major concern to

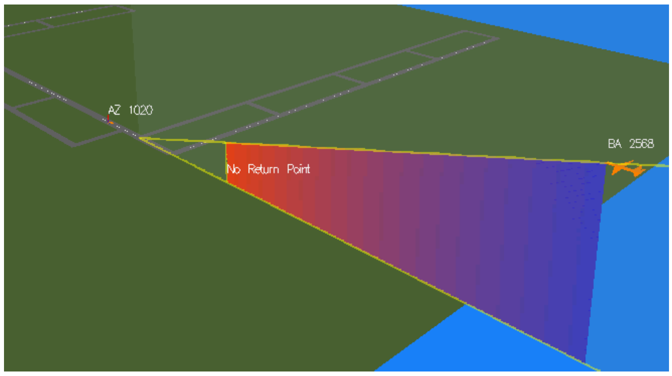


Fig. 5. A 3D view of the landing path, with time-based coloring and indication of point of no return for landing.

controllers is that of aeroplanes in flight that are going to landon a runway.

Typically, controllers have to make estimations about the time remaining before touch down and, more importantly, the time remaining before the approaching aircraft passes the so-called "point of no return" for landing. In fact, when the tower controller has to coordinate interleaved landing and departures from a runway, he has to ensure that the runway is free before authorizing an aircraft to land. If such a condition is not met, i.e. aircraft on the ground are still occupying the runway, a landing-abort is issued to the pilots.

The 4D visualization support for controllers in this critical task is provided by visualizing the landing path with runway aligned, for the incoming aircraft, and displaying the point of missed approach along the path. To increase awareness of timing (4th dimension), vertical coloured surfaces are drawn between the path line and the ground, where color represents the time remaining to the critical point. This is illustrated in Fig. 5.

VII. Picture-in-a-Picture Visualisations for Approach Control

In order to address the design recommendation for the "picture within a picture" concept described earlier, we developed a prototype concept we have called the "3D bubble". The 3D bubble concept presents both the planar 2D radar display, and a 3D view on demand of a selected part of the airspace. Such a visualisation is illustrated in Fig. 6. With the integration of both views into a single display space, we expect that the design will better support the complex spatial-temporal reasoning requirements we have identified earlier.

Fig. 7 shows a typical 2D radar planar view of three aircraft. The controller selects, say, the area in the centre of the display to activate the 3D bubble view. When the area is selected, note how the radar blips change into a triangle in preparation for the transition from 2D to 3D to represent the spatial dimension characteristics such as direction of travel and altitudes. Once the 3D bubble is activated, the view point of the display shifts, and only showing that the aircraft within the 3D bubble has been rendered in 3D, while the other two aircraft outside the

3D bubble remain in 2D perspective. In early evaluations of the 3D bubble concept, a number of "bubble" forms were also considered. These included a circle, a semi-sphere, a sphere, and an infinity cone. These different shapes are illustrated in Fig. 8.

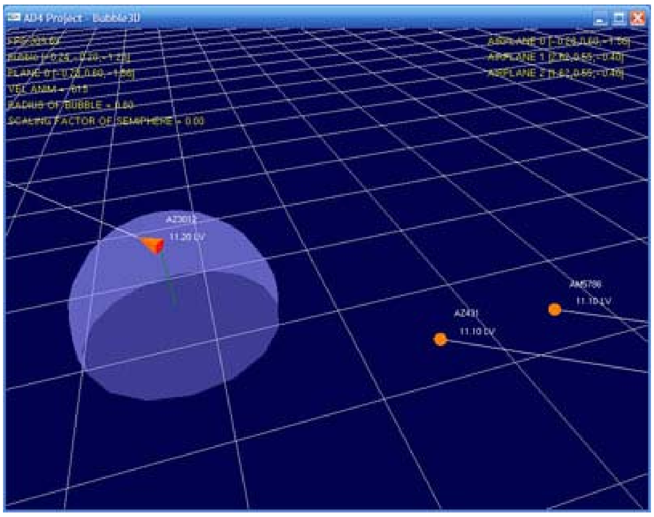


Fig. 6. The "3D Bubble" concept.

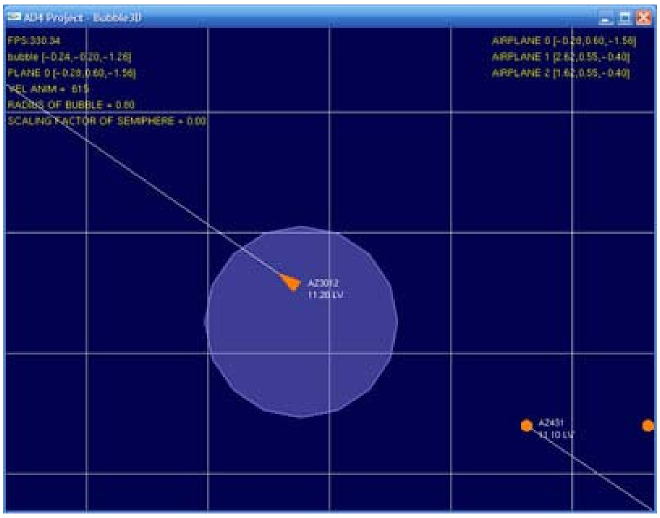


Fig. 7. Selection of 2D area in which to present the 3D bubble.

Fig. 7 shows a typical 2D radar planar view of three aircraft. The controller selects, say, the area in the centre of the display to activate the 3D bubble view. When the area is selected, note how the radar blips change into a triangle in preparation for the transition from 2D to 3D to represent the spatial dimension characteristics such as direction of travel and altitudes. Once the 3D bubble is activated, the view point of the display shifts, and only showing that the aircraft within the 3D bubble has been rendered in 3D, while the other two aircraft outside the 3D bubble remain in 2D perspective. In early evaluations of the 3D bubble concept, a number of "bubble" forms were also considered. These included a circle, a semi-sphere, a sphere, and an infinity cone. These different shapes are illustrated in Fig. 8.

In the prototype we built, the 3D bubble (and its variants) may be manipulated by changing the bubble model (disk, cylinder, hemisphere, sphere); increase or decrease radius of bubble; increase or decrease height of hemisphere model, and move bubble. The 3D camera viewpoint could also be freely moved according to users preferred point of view; for example

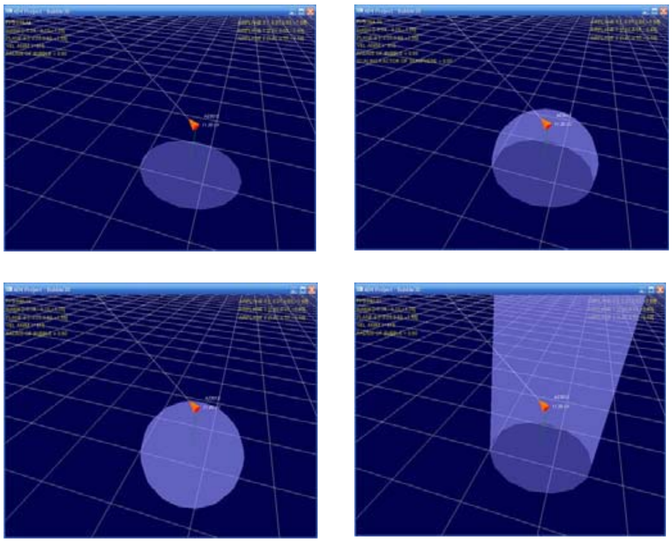


Fig. 8. Alternative shapes for the 3D bubble: circle, semi-sphere, sphere and infinity cylinder.

having different perspective when trying to understand the air traffic geometry under different flight situations. It is also possible to switch instantaneously between 3D and 2D views for managing situations that are more detectable in classic views. Camera (viewpoint) operations, such as rotating the perspective and changes between an orthographic view and a perspective, are managed with direct manipulation approach by a mouse and all other operations could be performed by key short-cuts.

In this way, we achieved the "picture within a picture" recommendation. However, at this stage of the design development, we have yet to incorporate the additional temporal dimensions to make this design truly 4D. This will be discussed next.

VIII. Visualization to support Spatial-Temporal Reasoning

In this section we illustrate how the concept of time can be integrated into the "3D picture within a 2D picture" display for approach control described in the previous section. The rationale for showing time-relevant aspects of the situation is to enable controller to minimise effortful mental computations needed to project a situation ahead in time, such as evaluating whether acceptable separations between aircraft exist, e.g. *if I take this aircraft along this direction, will the separation be guaranteed?*

In the proposed picture-in-a-picture concept, temporal information can be provided initially in the 2D planar view by means of future trajectories based on the aircraft's short track

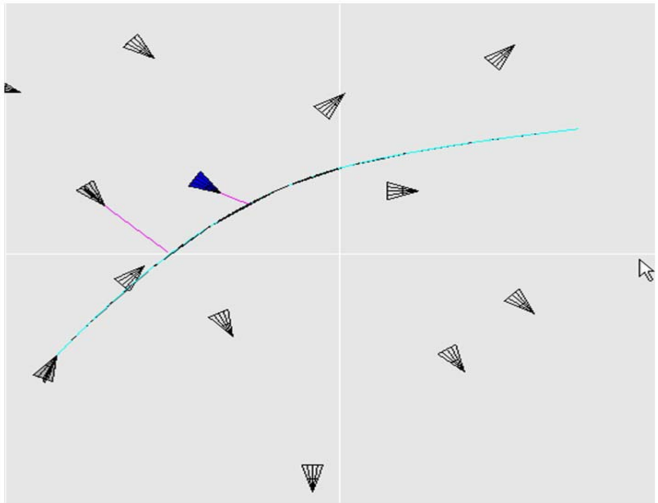


Fig. 9. Incorporating temporal information by projecting ahead intersecting aircraft trajectories.

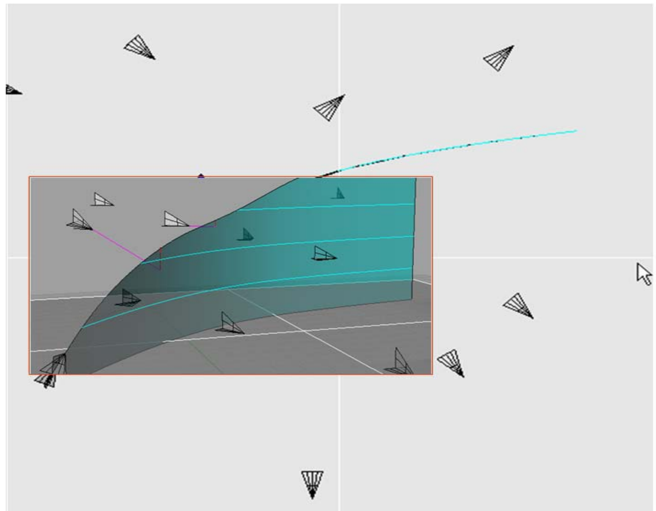


Fig. 10. A 4D-picture in a 2D-picture view showing in perspective aircraft positions in relation to each other in the selected area, with trajectories and constraints projected into the immediate future.

history, as provided by the current radar system. Such projections should facilitate the controller's detection of horizontal separations. This is illustrated in Fig. 9. For example, in the event that a controller wants to allow an aircraft which has just taken off to climb according to its planned departure flight path towards the en route flight level, based on the projections, she or he can verify whether trajectories of other aircraft are likely to cross the projected trajectory for the aircraft in question. The other aspect of the temporal dimension, is the conversion of the time it takes an aircraft to reach a given point, e.g. the intersection with the trajectory of the aircraft in question, can be converted into the distance travelled within the time in question, with these distances indicated on the display as lines projecting from the apex of the triangles representing an aircraft (Fig. 9).

If multiple crossing points exist, the controller could select and access a more detailed 4D view of the desired portion of airspace, so that other traffic in the surrounding area can continue to be monitored. This is illustrated in Fig. 10.

The 4D view portrays a more realistic view of traffic as it occurs in the selected portion of the 3D airspace. In terms of spatial considerations, the controller does not need to imagine, or to conjure up a mental model of the disposition of the traffic in the vertical dimension, as the controllers currently do. They will simply be able to see the traffic situation in the display. In terms of temporal considerations, the future projected trajectory of the aircraft in question is shown. This is depicted as a semi-transparent 'curtain', with horizontal lines to indicate the flight levels with intervals of 1000 feet, which are the usual intervals considered by controllers.

In the scene illustrated in Fig. 10, the trajectories of the other aircraft will cross this curtain, as the visualization includes only the potential conflicting traffics, i.e. only aircraft whose trajectories are projected to cross with the trajectory of the aircraft in question on the planar view are presented.

With such a representation the controllers can evaluate whether the aircraft will be adequately separated at the crossing points, i.e. when a possible conflicting aircraft will intersect the curtain (the intended trajectory), will acceptable separations exists at this point in time?

It is perhaps worth mentioning that the number of aircraft depicted in the 4D view is implicit in the change from the planar view to the spatial temporal view. The 4D picture does not show all the possible aircraft included in the 2D point of view, but only the possible conflicting aircraft instead. This approach mitigates the clustering problem intrinsic in complete 3D visualization of air traffic.

The visualization of separation in time is not concerned only with situation where the controller has to evaluate whether is possible to take a an aircraft which has just taken off along a possible short cut in order to reach the en route flight level. 4D is useful for example to evaluate whether two aircraft which are climbing towards the same point in space will be vertically separated in correspondence of the point in question; or again whether instructing an aircraft to descend along a given trajectory would cross trajectories of other traffics and whether the separations-horizontal and vertical- exist in correspondence of these point in time.

IX. Conclusion

In this paper we have shown how findings from the investigations about how controllers do their work has led to the development of a novel interface that can support the spatial-temporal reasoning needs in ATC. The ethnographic study and accompanying cognitive task analyses were fundamental in providing structure to the visual design of the 4D display. It seems difficult to envisage the development of a useful information portrayal design without a preliminary understanding of the controllers' spatial and temporal reasoning requirements, through the analysis of their strategies and information handling behaviour actually exploited during their everyday working practices. In the process of this research, our efforts have also led to the development of a Spatial-Temporal framework that, although still to be fully developed, can be used to guide design decisions about what aspects of the situation need to be portrayed, and how they should be related, including behaviours and temporal projections based on intentionality. However, there are still a number of issues yet to be resolved – what better forms are there for visually representing time? over what time bands should this information be presented? how do we overcome the visual continuity problems that arise from integrating 4D perspective view within a 2D planar view? how would a controller interact within the 4D space? what visual indicators and mappings between objects in the 4D world and the 2D space are needed to minimise cognitive mapping inefficiencies and errors? These questions set the stage of further research into visualisation support for spatial-temporal reasoning in ATC.

References

[1] P. Amaldi, B. Fields, S. Rozzi, and B. L. W. Wong, "Operational Concepts Report: Approach and Tower Control. Deliverable 12. Report prepared for AD4: 4D Virtual Airspace Management System, Sixth Framework Programme, Priority 1.4 Aeronautics and Space. Project Contract No. 012328. 31 January 2006, OCR1-AD4-WP2-MU," Interaction Design Centre, Middlesex University, London, UK 2005.

[2] P. Amaldi and M. Leroux, "Selecting Relevant Information in a Complex Environment: the Case of Air Traffic Control", *Proc. of the 5th European conference on cognitive approaches to process control*, Helsinki, 1995, pp. 95.

[3] H. Beyer and K. Holtzblatt, "Contextual design," Interactions, vol. 6, pp. 32-42, 1999.

[4] A. Burns and G. Baxter, "Time bands for systems structure," in Structure for Dependability: Computer-based Systems from an Interdisciplinary Perspective, D. Besnard, C. Gacek, and C. B. jones, Eds. London: Springer, 2006, pp. 74-90.

[5] C. M. Burns and J. R. Hajdukiewicz, Ecological Interface Design: CRC Press, 2004.

[6] J. H. Davison, J. M. Histon, M. D. Ragbarsdoittir, L. M. Major and R. J. Hansman, *Impact of Operating Context on the Use of Structure in Air Traffic Controller Cognitive Processes,* 2003. Available: http://icat-server.mit.edu/Library/fullRecord.cgi?idDoc=180

[7] M. R. Endsley, B. Bolte, and D. G. Jones, Designing for situation awareness: An approach to human-centered design. London: Taylor & Francis, 2003.

[8] M. R. Endsley and M. W. Smolensky, "Situation Awareness in Air Traffic Control: The Picture," in Human Factors in Air Traffic Control, M. W. Smolensky and E. S. Stein, Eds. San Diego: Academic Press, 1998, pp. 115-150.

[9] B. Kirwan, "Towards a controller-based conflict resolution tool - a literature review," EUROCONTROL, European Organisation for the Safety of Air Navigation ASA.01.CORA.2.DEL04-A.LIT, EATMP Infocentre Reference 020607-01, 14 March 2002.

[10] G. A. Klein, R. Calderwood, and D. Macgregor, "Critical decision method for eliciting knowledge," IEEE Transactions on Systems, Man and Cybernetics, vol. 19, pp. 462-472, 1989.

[11] S. Rozzi, W. Wong, P. Amaldi, P. Woodward, and B. Fields, "Design sketching for space and time," in Proceedings of the First Conference on Human Work and Interaction Design HWID 2006, A. Mark-Pjetersen, T. Clemensen, R. Orngreen, P. Campos, and W. Wong, Eds. Funchal, Portugal, 2006, pp. 181-205.

[12] H. S. Smallman and M. St. John, "Naive realism: Limits of realism as a display principle," in Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society, 2005.

[13] H. S. Smallman, M. St. John, H. M., Oonk and M. B. Cowen. (2001, September/October). Information Availability in 2D and 3D displays. *IEEE Trans. in Computer Graphics and Applications*, 2001 p.56.

[14] K. J. Vicente and J. Rasmussen, "Ecological Interface Design: Theoretical Foundations," IEEE Transactions on Systems, Man, and Cybernetics, vol. 22, pp. 589-605, 1992.

[15] C. D. Wickens and P., May, "Terrain Representation for Air Traffic Control: A Comparison of Perspective with Plan View Displays," Federal Aviation Administration, System Operations & Engineering Branch, Washington, D., C., Tech. Rep. ARL-94-10/FAA-94-2, September 1994. Available: http://www.aviation.uic.edu/institute/research/arl/technical-reports/terrain/terrain.html

[16] B. L. W. Wong and A. Blandford, "Analysing ambulance dispatcher decision making: Trialing Emergent Themes Analysis," in Human Factors 2002, the Joint Conference of the Computer Human Interaction Special Interest Group and The Ergonomics Society of Australia, HF2002, F. Vetere, L. Johnston, and R. Kushinsky, Eds. Melbourne, 2002, pp. CD-ROM publication.

Remote Airport Tower Operation with Augmented Vision Video Panorama HMI

M. Schmidt, M. Rudolph, B. Werther, N. Fürstenau

Abstract— **In this paper the initial results of the DLR project Remote Airport Tower Operation Research (RapTOr) are described. Within this project an augmented vision video panorama system and a corresponding HMI is developed for remote surface movement management of small airports or of movement areas not directly visible for the controller. Ground traffic management is performed from a remotely located control center, e.g. a tower at a different airport. The setup and functions of the high resolution video panorama system with integrated pan-tilt zoom camera (PTZ) at Braunschweig research airport is described. It provides the framework for integrating SMGCS traffic data, real time image processing and replay possibility of the complete 180° - panorama. Furthermore we describe how the results of a formal tower work analysis are transferred into a cognitive operator model for simulating the RTO controller work processes to be integrated into the conventional tower work environment.**

Index Terms—**Airport tower, remote operation, video panorama, augmented vision, work analysis, cognitive modelling**

I. Introduction

Remote Tower Operation (RTO) describes the goal of remote control of small airports and of movement areas of large airports which are not directly visible from the tower, by means of an augmented vision video panorama that replaces the direct far view out of the tower windows. In 2005 the DLR project RapTOr was started in order to realize an RTO experimental system as extension of the Advanced Surface Movement Guidance and Control System (A-SMGCS) at the Braunschweig research airport. Analysis and simulation of the tower work procedures support the design and development of the demonstrator. RTO is the first step on the way to the Virtual Tower (ViTo) as long term goal [1][2][3].

The direct view out of the tower windows is of central importance for surface traffic control under the present day working conditions of tower and apron controllers. That is why each attempt of surface movement management without direct view out of tower windows represents a revolution for tower controller working conditions. Consequently it is assumed that under the guideline of human centered automation, the reconstruction of the direct far view in future A-SMGCS, in addition to the abstract (birds view) ground movement situation display showing the vehicle positions, will greatly improve the transition process to the new work environment and make it acceptable to the user. The design process for a RTO (and future ViTo) work environment relies on support from domain experts (controllers) of the German Air Traffic Control Organisation (DFS), in particular with respect to the work analysis [8]. Initial design goal is the integration of the RTO workplace into an existing tower work environment of a medium size airport in order to simultaneously control one or more neighbouring small airports.

The project encompasses two major research and development goals: 1. Simulation of operator decision making within the RTO augmented video panorama workplace environment in order to support the development process and the investigation of design alternatives, without the necessity of actually realizing each of the alternative solutions; 2. realization of an RTO – experimental environment and a demonstrator for augmented video panorama based remote airport control, including tests and concept validation with professional tower- / apron controllers. The following section 2 describes initial results of a tower work and task analysis following a systematic procedure developed by K. Vicente [4] and the development of a formal colored Petri net model, realized with CPN-Tools [5]. The model serves for simulating the operator's decision making processes with the work analysis data as input. In section 3 the panoramic far view reconstruction as visual basis of the RTO experimental environment at Braunschweig research airport and its basic functions are described. In section 4 we review initial experiments on augmented tower vision (ATV) to be integrated into the video panorama. Section 5 provides a conclusion and outlook with regard to the final demonstrator and the Virtual Tower as long term goal.

II. Cognitive Work Analysis and Simulations Using Colored Petri Nets

The design and development of the experimental RTO HMI

Manuscript received February 20, 2006.

M. Schmidt is with the Institute of Flight Guidance, German Aerospace Center, D-38022 Braunschweig, Germany e-mail: markus.schmidt@ dlr.de).

M. Rudolph is with the Institute of Flight Guidance, German Aerospace Center, D-38022 Braunschweig, Germany (e-mail: michael.rudolph@ dlr.de).

B. Werther is with the Institute of Flight Guidance, German Aerospace Center, D-38022 Braunschweig, Germany (e-mail: bernd.werther@ dlr.de).

N. Fürstenau is with the Institute of Flight Guidance, German Aerospace Center, D-38022 Braunschweig, Germany (corresponding author to provide phone: +49 531 295 2549; fax: +49 331 295 2550; e-mail: norbert.fuerstenau@ dlr.de).

is supported by a cognitive work and task analysis (CWA) of the presently existing work environment and decision processes. The formalized results serve as input data for a computer simulation of controller decision making in a tower workplace design which is modified by an additional RTO HMI for controlling one or more remote small airports or movement areas. The CWA is based on a formal procedure suggested by Vicente [4], separating the analysis into five areas: work domain analysis, control task analysis, strategy analysis, analysis of social organisation and cooperation, and operator competency analysis, the latter however not being considered in this phase. A CWA at Leipzig airport with potential small airports Altenburg and Erfurt for remote control is performed with support of the German air traffic control organization (DFS). The formal mathematical background of the simulation is a colored Petri net model of the work processes. Details of the CWA as well as the cognitive modelling and simulation using colored Petri nets are described in [6], [7] and [8]. In what follows only a brief review will be presented. Fig. 1 provides a block diagram of the CWA with special focus on the analysis of Tower Leipzig processes [8].

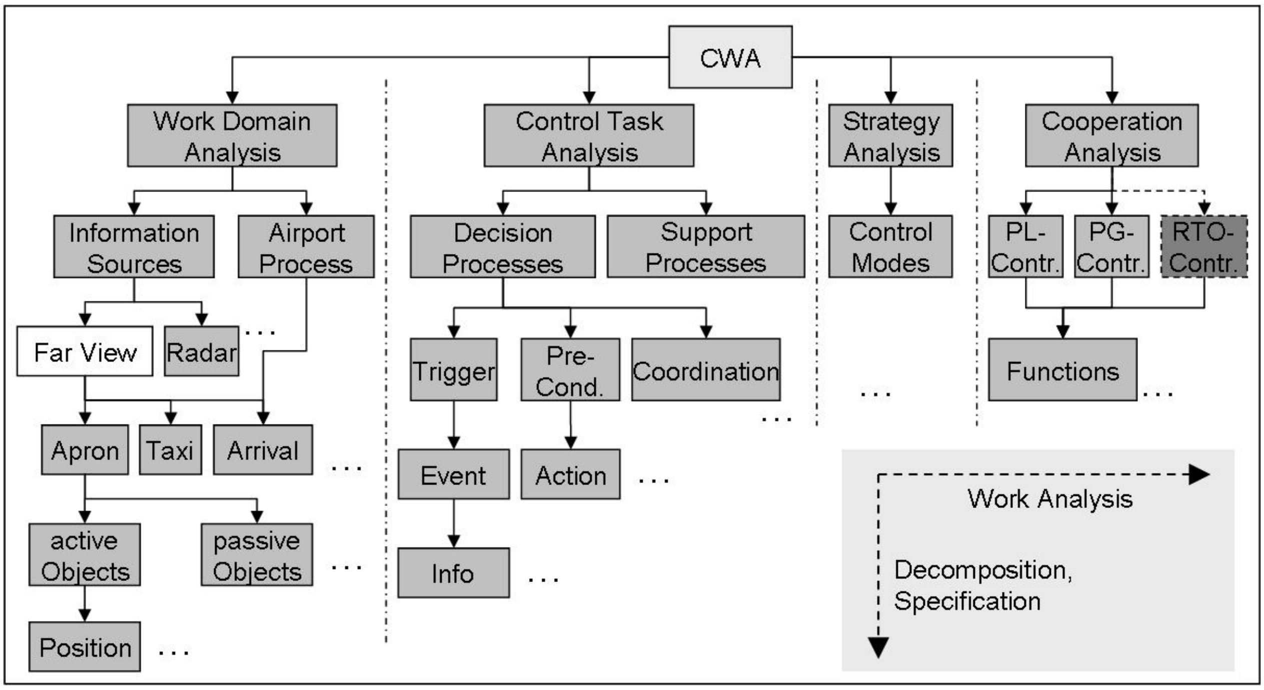


Fig. 1: Decomposition and specification of the CWA areas for a two – controller (PL, PG) tower with additional RTO workplace.

The work domain analysis aims at analysing the aircraft movements. For this purpose the air-to-air process which describes the complete movements from arrival to departure is treated separately for the different control areas (e.g. approach, runway, taxi, apron). Acquired information and possible actions are attributed to corresponding control areas. The analysis of accessible information from the different sources (e.g. far view from the tower windows, approach and ground radar) and possible actions via the corresponding interaction devices (e.g. radio, telephone) at this step is performed without considering the controllers tasks.

In the control task analysis as second step the tasks are identified which have to be completed. Here decision and support processes are treated separately. The task description follows a well defined structure which covers the triggering event, the preconditions, the task containing coordination, and the post-condition.

The strategy analysis is the most laborious step. This is because controllers to a large extent use implicit knowledge which is hard to extract. In empirical studies (e.g. [9]) strategy differences, e.g. dependent on work load were detected. The development of strategies to a large extent depends on the handling of goals under restricted cognitive resources [10]. This is one important motivation for the use of the resource based Petri net modelling technique. An important aspect of multiple task situations as typical for controllers is the relative weighting of different simultaneous goals with respect to each other. Action strategies evolve due to limited human processing capacity.

The analysis of cooperation and social organisation yields a clear tasks and functions distribution between the two controllers (ground controller (PG), runway controller (PL)) in the Leipzig tower example. The future RTO workplace, however, represents a significant change of this situation. On the one hand the augmented vision video panorama offers revolutionary new possibilities for the support of air traffic controllers. On the other hand the integration of remotely located control areas within the present day tower environment represents a completely new work condition.

Colored Petri nets (CPN) represent a compact and transparent method for modelling in discrete time intervals the evolution of complex dynamical systems with limited resources. Based on graph theory and linear algebra Petri nets provide an inherent formal method for analysing the reachability of system states. Within the present context the hierarchical structuring potential of colored Petri nets is fully utilized. The results of the single steps of the CWA are fed separately into respective levels of the hierarchy structure as indicated in Fig. 2.

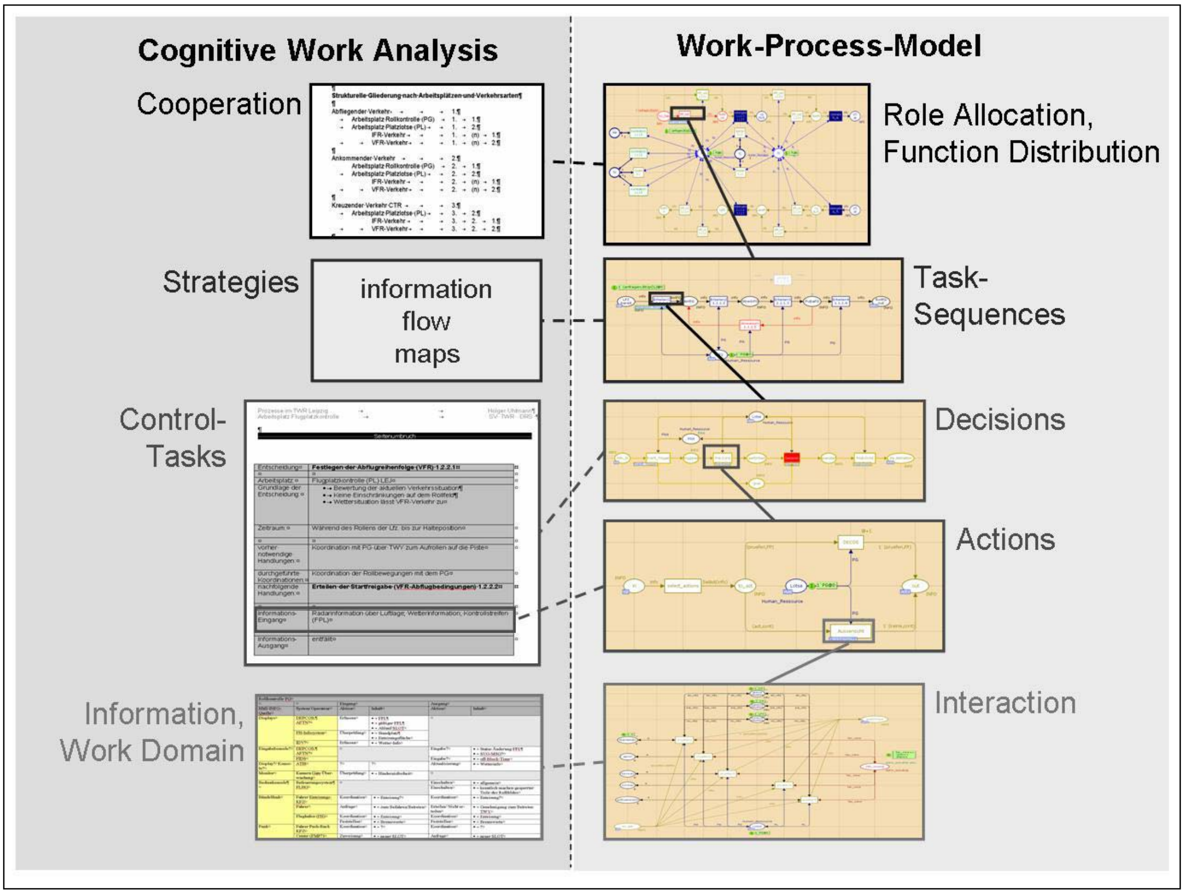


Fig. 2: Transfer of CWA results into the formal mathematical structure of Colored Petri Nets.

The results of the fourth CWA step (cooperation) are fed into the highest hierarchy level of the CPN structure. Here the distribution of roles and functions among the different human operators and their technical support systems is performed. On this level the work process is described in a holistic manner whereas on the lower levels focus is put on the single work positions.

In the next lower level of the Petri net architecture (task sequences) the goal driven actions for fulfilling the respective tasks of each operator are implemented. The results of the strategy analysis (CWA step 3) are fed into this level.

The identified control tasks in CWA step two are modelled with regard to the actions to be performed and the required and created informations. Actions are combined to decision and support processes respectively. Actions are separated into those representing preconditions and others which coordinate the task (decision process, support processes) and those actions which complete the task as post condition. The Petri net realizes this description by means of two net levels (decisions, actions).

The lowest level of the Petri net hierarchy models the results of the work domain analysis (1st step of the CWA). Here the interface is realized between the operators and the traffic processes to be controlled. The analysis of the information sources and interaction systems as basis of the work processes provides the data for the corresponding Petri net model of the HMI. Of primary interest for RTO is the far view out of the tower windows as basic information source which will be replaced by the augmented video panorama. Augmented tower vision (ATV), i.e. the superposition of additional information on the far view like weather data and aircraft labels with object tracking rises questions adressing controllers attention and perception processes. Problems like attentional tunneling and involuntary switching of perception may require separate modelling approaches [13][14][15].

With the homogeneous description of the human machine system based on colored Petri nets it is possible to investigate the consistency of the human and the process model based on formal analysis as suggested by Degani & Heyman [11]. Critical work situations can be detected and analyzed in an early phase of the system design and alternative solutions can be investigated by means of model based simulations. The transfer of the CWA results into the formal operator model allows for monitoring psychological parameters of the operators (e.g. work load), for deriving the operator requirements and for uncovering missing situational awareness by means of simulations and reachability analysis [12]. Furthermore the graphically represented formal work process model provides a valuable support for the communication between domain experts and system developers.

III. Augmented Vision Video Panorama Experimental Environment

A high resolution video panorama system is presently being set up at Brauschweig research airport as experimental environment for investigation of different aspects of the video based RTO – HMI and development of a demonstrator. A schematic showing the main concept and the integration into the local A-SMGCS is depicted in Figure 3.

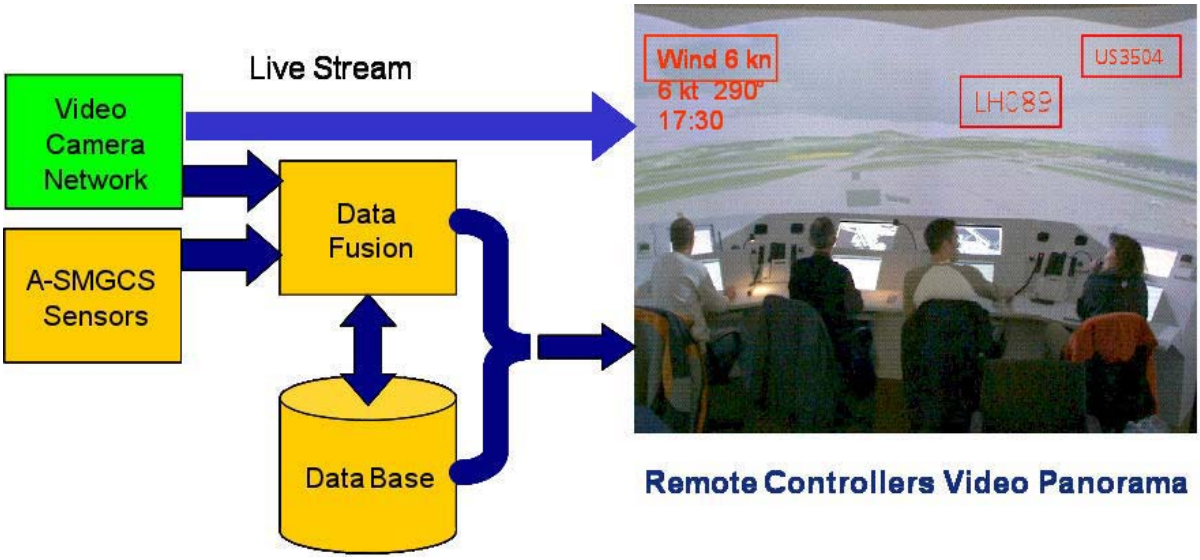


Fig. 3: Conceptual Sketch of RTO – Experimental System at Braunschweig Research Airport.

A more detailed block diagram of the RTO system with augmented vision video panorama is depicted in Figure 4.

The panorama camera system represents an additional sensor system of the local experimental A-SMGCS including a multilateration system and D-GPS. The reconstructed far view within the video panorama will be augmented by means of traffic and environmental data. While traffic data like aircraft labels move coherently with the corresponding objects other information like weather data is statically integrated into the reconstructed far view.

The present sensor system consists of four high resolution (1600 x 1200) high dynamic range (14 bit/pixel) CCD cameras (P_1 ... P_4) covering the Braunschweig airport (runway length 1.6 km, extending in E – W direction, 400 m distance from viewing position) within 180° and a remotely controlled pan-tilt – zoom camera (PTZ). The system is positioned ca. 20 m above the airport surface on top of a building at the southern boundary of the airport with 100 m distance to Braunschweig tower and central view looking north with horizontal alignment. Figure 5 shows a photo of the system.

The vertical aperture angle of about 20° (half angle with respect to the horizontal line of sight) allows for a closest surveillance distance of about 60 m. The object resolution per pixel in 500 m distance is about 0.25 m vertical and 1 m along the line of sight. Camera signals with 25 frames/s are split into two outputs for real time image processing and simultaneous data compression for transmission to the remote RTO – HMI. A GBit ethernet switch feeds the images from the five sensors into a single mode fiber – optic data link which transfers the typically 100 MBit/s data (depending on degree of compression, brightness and image content) of the panorama system and PTZ over a distance of 450 m to the DLR Advanced Control Center Simulator (ACCES). A second GBit ethernet switch splits the incoming data into five output channels for decompression with one PC per camera. Each camera is remotely controlled with respect to aperture and γ – correction. The PTZ camera is controlled with respect to azimuth and vertical angle as well as zoom (23-fold, focal width 3.6 mm – 82.8 mm, corresponding to 54° - 2.5° visual angle).

In addition to the visual information digitized acoustic signals of a microphone and weather data (temperature, wind speed, static pressure) from a weather station at the camera position are transmitted to the control room via the same fiber optic link.

The panorama visualization in the present state of development (Fig. 6) is realized with one row of a high resolution tiled wide angle backprojection system (WAP, 4 x 2 single 1280 x 1024 SXGA projectors) and in parallel via four conventional SXGA monitors arranged as a cylinder section for reproducing the azimuth angles of the cameras. Both options are realized within ACCES, based on a PC cluster with central workstation for controlling the tiled projection wall and interaction. Figure 6 depicts the live video panorama in the top row, with the zoomed apron area below and a birds view of Braunschweig airport with camera sectors indicated in the two lower right projection tiles. The monitor based live panorama with separate PTZ monitor can be seen below the projection wall.

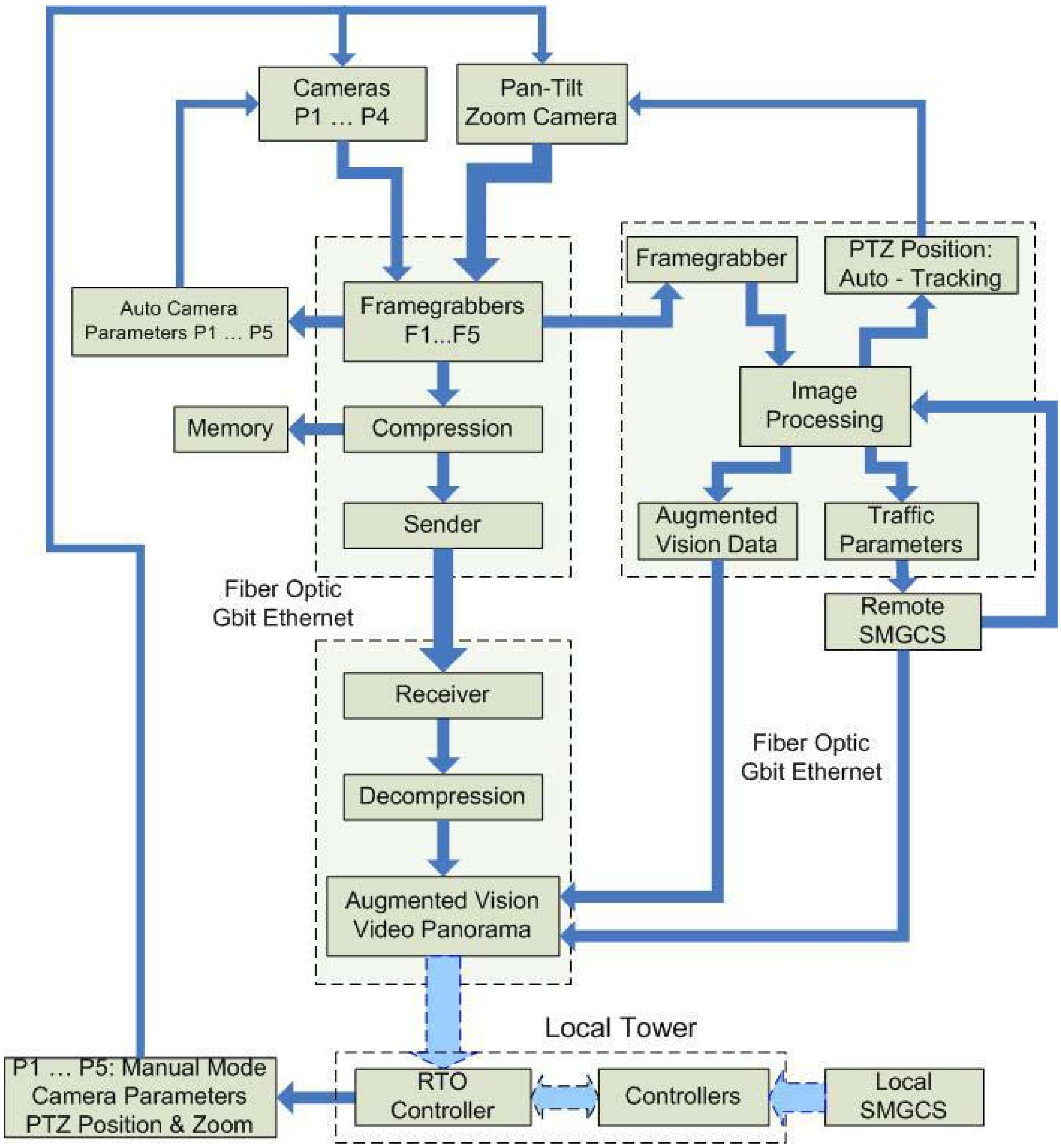


Fig. 4: Schematic block diagram of augmented vision video panorama system.

Interaction of the operator with the panoramic system (cameras P1 ... P4, PTZ, weather station, microphone) in the present development stage is performed via keybord and mouse for modifying contrast, brightness and gamma correction of cameras. Contours of the movement areas filled with transparent coloring can be superimposed on the reconstructed panorama. The latter kind of image augmentation is particularly helpful for guiding the operators attention during darkness or bad weather conditions to those areas where moving vehicles are expected.

For positioning the PTZ camera the target can be defined manually or by automatic movement detection. In the manual mode a rectangular contour is positioned at the required point of the panorama, defining the area to be enlarged. The center of this area defines the new line of sight of the zoom camera as position command value. With the tracking mode turned on

Fig. 5: High resolution panorama camera setup and pan-tilt zoom camera of the video panorama system. Braunschweig tower in the background (Photo DLR).

Fig. 6: High resolution wide angle projection with Braunschweig airport video panorama in the Advanced Control Center Simulator (ACCES) of the DLR Inst. for Flight Guidance (Photo DLR)

the square moves coherently with the corresponding object. An algorithm for real time movement detection is running on a separate parallel processor of the image compression PC of each camera. An overall latency time between image acquisition and panorama visualization of 230 ms – 270 ms was measured by means of a special shuttered laser arrangement.

The five recording PC's with the compression software at the camera position allow for storing panorama and zoom data which amounts to typically 140 MByte / min / camera (i.e.

roughly 40 GByte of data per hour). This feature provides the possibility of complete panorama replay. Besides future operational applications it is of great value for the augmented vision HMI development and validation because, e.g. real traffic scenarios can be repeatedly presented to operators with modified augmented vision content.

IV. AUGMENTED TOWER VISION: INITIAL EXPERIMENTS

One important advantage of the digital video panorama is the easy integration of augmented vision features which was proposed also for enhancing the real view [19]. Initial field trials as well as laboratory experiments with superimposed information on the far view have been performed in order to address the question of acceptance of augmented vision systems by controllers and the aspect of head down time reduction by using transparent displays for reducing the number of monitors.

A field trial was performed at the (old) tower of Dresden airport with a transparent holographic back projection display (HOPS). The photo in Fig. 7 shows the HOPS between the controllers workplace and the tower window. With respect to the see – through characteristics the HOPS function is similar to the head – up display (HUD) frequently used in military aircraft. In contrast to HUD's however, the (relatively cheap) HOPS system provides no collimated view, i. e. instead of a virtual image in the far distance a real image is created on the pane. The goal of the system is to provide a very bright image, which is of use under daylight conditions. Technically a holographic foil between two glass panes collimates the light from the projector (beamer) positioned on the ceiling or floor behind the screen under an angle of 36° with respect to the normal of the projection plane. Typical resolution is 100 lines/mm; visibility ranges are ±10° vertical and ±30° horizontal. The main effect is the concentration of the light from the beamer to a limited range on the opposite (operator's) side of the screen and the suppression of ambient light which does not originate from the projector. As can be seen in Fig.7, the brightness of the white symbols as obtained with a 4100 lm – DLP projector (XGA resolution) is sufficient for recognition of the text. It is evident that colors have to be chosen according to the see through condition and are different as compared to the standard monitor situation: the background has to be black (no color); good results concerning readability of the symbols are obtained with white, yellow or bright red.

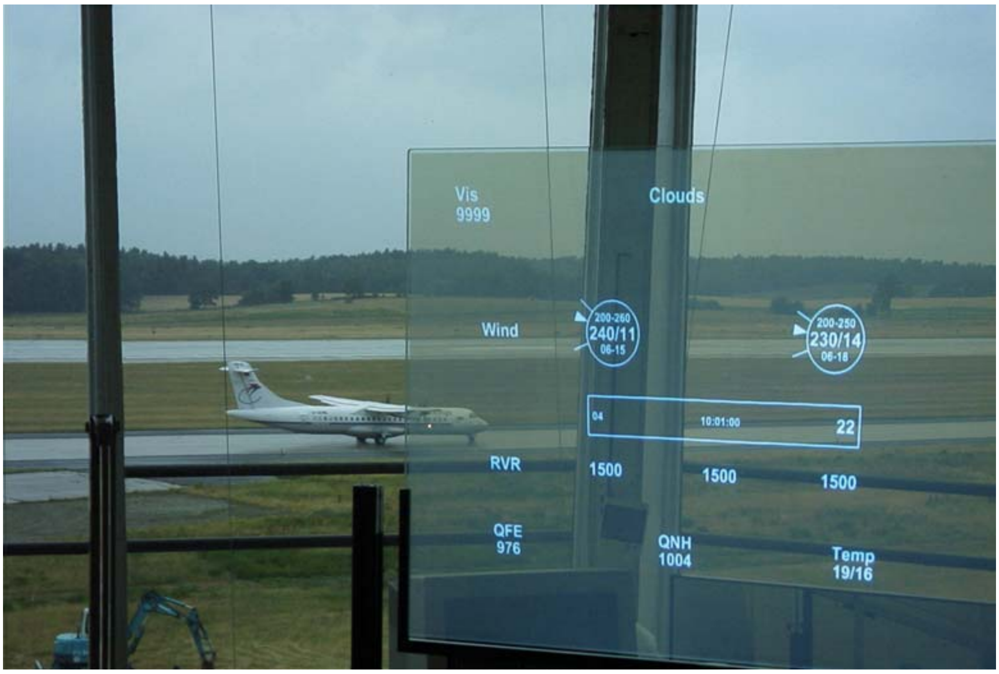


Fig. 7: Field Trials withTransparent Holographic Backprojection Display at Dresden Tower Showing Real Time Weather Data (Photo DLR).

Augmented vision systems are expected to reduce head–down times of controllers, i.e. time intervals needed to read data from displays instead of controlling traffic by means of the far view out of the windows. As a rule of thumb the head –down time should not significantly exceed more than 30 % of the working time.

A large number of experimental results have been obtained in recent years concerning attention and perception effects with augmented vision using head – mounted (HMD) and head-up displays (HUD) which are used in cockpits of military aircraft and helicopters (e.g. [13][16] [17] [18]). It was found that the expected advantage of reduced visual scanning between the head-down instruments and the far domain is not stable under all circumstances and may be counteracted by effects such as cognitive tunneling and visual interference.

In order to obtain initial quantitative data on the expected head-down time reduction by means of augmented vision systems we performed a schematic laboratory experiment with the HOPS system of Fig. 7 [19]. The setup consisted of a wide angle background projection with the transparent HOPS projection screen in the foreground. The latter one superimposed a list with numerical information on the background scene. Subjects (18 participants) had to perform a visual search task and response times (RT) were measured. Sets of abstract objects (geometric shapes: circle, square, triangle) had to be scanned in the background in order to decide if a specified one was within the set. The alphanumeric foreground information consisted of a list potential target objects together with individual codes. The specific code for the target object to be searched in the background was shown in the background together with the set of objects. The subject's search task consisted of three steps: find the code between the background objects, associate code and target object on the superimposed numerical list, decide if the target is present within the set of background objects. The same visual search task was repeated with four different modes of presenting the foreground information: 1. conventional display positioned below the line of sight to the background (requiring no head movement), and 2. 45° to the side of the line of sight, 3. HOPS display superimposing information, and 4. alphanumeric list integrated into the background (corresponding to the case of augmented vision information integrated into the panorama of Fig. 6). In these initial experiments addressing attention and perception aspects of augmented tower vision only condition 2 showed significantly longer RT (ca. 400 ms) as compared to all other conditions with similar RT's.

V. CONCLUSION AND OUTLOOK

The formal cognitive modelling and simulation of work processes based on a detailed work analysis in combination with the technical development of a revolutionary new work environment appears to be a promising approach for introducing new concepts into the conservative and safety critical ATM field. For remote control of small airports and movement areas from a local tower at a larger airport the present work conditions have to be modified in order to integrate the RTO controllers workplace in the tower environment. One important question to be investigated by simulating the work processes refers to the number of operators: can the additional control of the remote airport be performed by the existing controller team or is an additional RTO controller required? This question of course is closely related to the design of the new RTO workplace and the integration into the existing tower environment as well as work organisation and training procedures. It is expected that the utilization of advanced augmented vision and virtual reality technologies may help improving the integration of functions and collaboration of operators. The stability of work processes under all conditions has to be proven by realistic simulations based on the formal work process model. Problems associated with augmented tower vision systems may require additional experimental and modelling approaches as indicated in section 4.

Compared to the human factors related questions of the RTO work environment the technical problems appear to be a minor problem. Nevertheless the usage of most advanced high resolution and high dynamic range cameras in combination with automatic zoom and tracking functions and the augmented vision wide-angle video panorama projection represents also a new challenge within the ATM technology field which has to be validated within the RTO experimental environment. Systematic tests in the (tower) simulator environment with active controllers are planned for the final project phase, by utilizing the panorama replay function.

ACKNOWLEDGMENT

We are indebted to Detlef Schulz-Rückert, Holger Uhlmann and Dieter Bensch of German Air Traffic Control (DFS) for continuous support of the Virtual Tower and Remote Tower Operation projects during the recent years, in particular for providing detailed domain expert knowledge for the tower work analysis.

REFERENCES

[1] N. Fürstenau, "Virtual Reality for Integration", in *Proc. 12th Scientific Seminar,* in: DLR – Mitteilungen, 2003-03, DLR, Inst. of Flight Guidance, Braunschweig, 2003.
[2] N. Fürstenau, M. Rudolph, M. Schmidt, B. Werther, "Virtual Tower", in *Wettbewerb der Visionen 2001 – 2004*, DLR-report, 2004, pp. 16 – 21.
[3] N. Fürstenau, M.Rudolph, M. Schmidt, B.Werther, 2005, "Virtual Tower", in: *5th ATM R&D Symposium: Airport – Bottleneck or Booster for Future ATM,* Braunschweig 2005. Available: http://atmsymposium.dlr.de
[4] K.J. Vicente, *Cognitive Work Analysis,* Mahwah/NJ: Lawrence Erlbaum Associates, 1999.
[5] K.Jensen, *Coloured Petri Nets.* Berlin: Springer, 1997.
[6] B. Werther, B. Lorenz, "Menschzentrierte Modellierung von Automatisierungssystemen", in: *Automatisierungs – und Assistenzsysteme für Transportmittel, Fortschritt-Berichte VDI,* vol. 12 no. 525, 2003, pp.63 – 79.
[7] B.Werther, "Kognitive Modellierung mit farbigen Petrinetzen zur Analyse menschlichen Verhaltens", PhD Dissertation, Technische Universität Braunschweig, 2006.
[8] B.Werther, H. Uhlmann, "Ansatz zur modellbasierten Entwicklung eines Lotsenarbeitsplatzes", in: *Zustandserkennung und Systemgestaltung, Fortschritt Berichte VDI*, vol. 22, 2005 pp. 291 -294
[9] J.C. Sperandio, "The regulation of working methods as a function of work-load among air traffic controllers", *Ergonomics,* vol. 21 (3), 1978, pp. 195-202.
[10] D. Dörner, 1989, *Die Logik des Misslingens. Strategisches Denken in komplexen Situationen.* Reinbek bei Hamburg: Rowohlt Verlag.
[11] A. Degani, M. Heymann, "Formal verification of human – automation interaction", *Human Factors*, vol. 44, 2002, pp. 28-43.
[12] B. Werther, E. Schnieder, "Formal Cognitive Resource Model: Modeling of human behavior in complex work environments", in: *Proc. Int. Conf. Computational Intelligence for Modelling, Control & Automation (CIMCA 2005),* Wien: 2005, pp. 606 – 611.
[13] R. Martin-Emmerson, C. D.Wickens, "Superimposition, Symbology, Visual Attention, and the Head-Up Display", *Human Factors,* vol. 39, 1997, 581-601.
[14] J.S. McCarley, C.D. Wickens, J. Goh, W.J. Horrey, "A computational model of attention/situation awareness", in: *Proc. 46th Annual Meeting of the Human Factors and Ergonomics Scociety,* Santa Monica: 2002, pp. 1669 - 1673

[15] N. Fürstenau, "A nonlinear dynamics phase oscillator model for the simulation of multistable perception", *in: Proceedings 7th Int. Conf. on Cognitive Modeling (ICCM 7),* Trieste: Edizioni Goliardiche, 2006, pp. 106 – 111.
[16] G.W. Stuart, K.I. McAnally, J.W. Meehan, "Head-up displays and visual attention: Integrating Data and theory", *Human Factors and Aerospace Safety,* vol.1, 2001, pp. 103 – 124.
[17] E. Peli, "Visual issues in the use of a head-mounted monocular display", *Optical Engineering,* vol. 29, 1990, pp. 883-892
[18] R.S. Laramee, C. Ware, "Rivalry and Interference with a Head Mounted Display", *ACM Transactions on Computer-Human Interactions,* vol. 9, 2002, pp. 238-251.
[19] N. Fürstenau, M. Rudolph, M. Schmidt, B. Lorenz, T. Albrecht, "On the use of transparent rear projection screens to reduce head – down time in the air – traffic control tower", in: *Proc. Human Performance, Situation Awareness and Automation Technology (HAPSA II),* Mahwah/NJ: Lawrence Erlbaum Publishers Inc., 2004, pp. 195 – 200

Markus Schmidt studied mechanical engineering / micromechanics at the Technical Univ. of Braunschweig and received the Diplom Ingenieur degree in 1998 with a work in fiber – optic sensor technology. In the same year he joined the sensor technology group of the DLR Inst. for Flight Guidance where he was involved in the development of microinterferometer strain gauges and phase demodulation systems for sensor readout. Together with his co-authors he joined the human factors department in 2000 where he focused on research in new HMI technologies and was involved in the Virtual Tower concept study (2002 – 2004). He is presently deputy project leader of the RapTOr project and responsible for the realization of the RTO experimental system. His research interest is in the system aspects of high resolution video and projection technologies and data fusion.

Michael Rudolph was educated as Mathematical Assistant at Airbus Industries in Bremen and studied Informatics at the Technical University of Braunschweig, After receiving the Diplom Informatiker degree in 2002, he joined the Human Factors Department of the Inst. for Flight Guidance in the same year. Presently he is responsible for the realization of the video panorama and augmented vision software of the RTO experimental system.

Bernd Werther studied electrical engineering at the Technical University of Braunschweig and received the Diplom Ingenieur degree in 1999 with a thesis on fiber optic sensor technology, performed at the German Aerospace Center (DLR). In the same year he joined the sensor technology group of the DLR Inst. for Flight Guidance. He joined the Human Factors department in 2000 and in 2001 started PhD research in the field of cognitive modelling. He finished his dissertation in 2005 with a work on cognitive modeling of human decision making using colored Petri nets. Within the RapTOr project he is presently responsible for the work analysis, modelling and simulation of tower work processes for RTO - HMI design.

Norbert Fürstenau studied physics at the universities of Braunschweig, Darmstadt and Frankfurt. He received the Diplom Physiker degree from the Inst. of Nuclear Physics of Techn. Univ. of Darmstadt in 1977 and Dr. phil.-nat. degree from Frankfurt University in 1981 with a work in applied Laser physics and physical chemistry of molecular clusters. He joined the DLR Inst. for Flight Guidance in 1981 and was leader of the Fiber Optic Sensors group until 2000, when he joined the Human Factors Department together with two of his co-authors. He initialized the Virtual Tower (ViTo) concept study (2002 – 2004) and is presently project leader of the Remote Tower Operations (RapTOr) project. His present research interest is in cognitive modelling of visual perception and attention using nonlinear dynamical systems theory.

Human Performance with Simulated Collimation in Transparent Projection Screens

Stephen Peterson and Ella Pinska

***Abstract*—AR displays might prove a valuable tool in the controlling tasks at future Air Traffic Control towers. A non-intrusive alternative to the traditional AR display device, the Head Mounted Display, is the spatial display category. It involves displays fixed in front of the user, where one type is the transparent projection screen. They are in some ways similar to head-up displays (HUDs) used in aircraft cockpits, since they both superimpose graphics on the background view. However, they also have fundamental differences, like their abilities to spatially align objects in the displayed image with background objects. The concept of making a displayed image appear to be coming from a distance much further than the display surface is known as collimation. The ability to provide a collimated image varies form display to display. When true collimation cannot be naturally provided, it can be partially achieved, or simulated.**

This paper reports on the findings from an experiment on human performance in different collimation levels of a transparent projection screen. The main conclusion of this experiment is that simulated collimation in transparent projection screens shows significantly lower response times when performing a visual search task, compared to a non-collimated display.

***Index Terms*—Augmented Reality, Air Traffic Control, Collimation, Head-Up Display, Head-down Time.**

I. INTRODUCTION

AS the world's air traffic continues to grow, more demand is put on infrastructure including airspace and airports. One component of the airport that is crucial to the controlled flow of aircraft is the Air Traffic Control Tower (ATCT), in which air traffic controllers (ATCOs) perform specific controlling and monitoring tasks. For many tasks the visual input from the airport, the "out-of-the-window" reference, is important. According to ICAO procedures [10], local Air Traffic Service authorities shall under low visibility conditions specify the appropriate separation on the airport surface areas, such as taxiways, possibly affecting the smooth traffic flow.

Stephen Peterson is a PhD candidate at the Department of Science and Technology, Linköping University, Norrköping, Sweden. He is a doctoral researcher at EUROCONTROL Experimental Centre, Brétigny-sur-Orge, France. (e-mail: stephen.peterson@eurocontrol.int)

Ella Pinska is a PhD candidate at the École Pratique des Hautes Etudes, Paris, France. She is a doctoral researcher at EUROCONTROL Experimental Centre, Brétigny-sur-Orge, France. (e-mail: ella.pinska@eurocontrol.int)

If the controller's visual input level was held constant at all times, and also enhanced with valuable information for the task, supplementary input devices like radar screens would become less important. Less head-up/head-down movements could significantly decrease the response time to critical signals [9], consequently leading to increased performance in these visually demanding tasks.

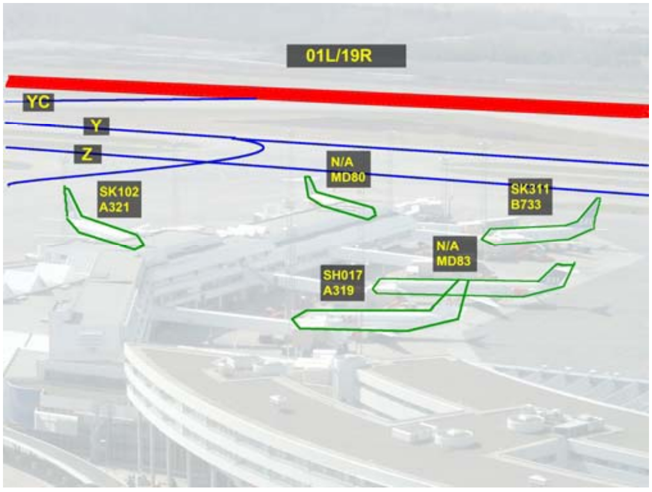


Figure 1: An example of AR user interface in an ATCT environment.

Such an enhanced view could be provided through an Augmented Reality (AR) display system. AR displays add a data layer in the visual field through an (almost) transparent display device situated in the ATCT between the observer (ATCO) and the outside world (airport). The graphics displayed in an AR display is spatially aligned with objects in the background, essentially making it look the two visual stimuli (display/background) are integrated. This alignment process (know as "registration") requires exact knowledge about the positions of the objects, user and display in order to compute correct perspectives. Figure 1 shows a prototype example of an AR user interface in an ATCT environment.

The most widespread AR display device is the Head-Mounted Display (HMD) which is a miniature display worn on the user's head. While it has some advantages, it is still a relatively intrusive device since it places a substantial weight to the user's head. Moreover, such a display requires signal and power cables (or batteries) to function.

An alternative to HMDs is the spatial display, i.e. a display that is fixed in space and not attached to the user. A similar system is the Head-Up Display (HUD) used in aircraft cockpits. The main difference between a HUD system and a spatial AR display system is that the latter registers the displayed graphics with the background. The HUD could of course be a part of a spatial AR display system, but normally it would require tracking of the user position in order to provide correct alignment of the graphics with the background.

This paper reports on the findings from an experiment on human performance in one type of spatial display, the transparent projection screen. Before detailing the experiment and results, some theoretical background and related work is reviewed.

II. BACKGROUND

A. Depth Cueing in Spatial AR Displays

In an AR display the aim is to spatially align the displayed graphics with the background. Provided that accurate positional data is provided, the display must be able to output the data at the correct spatial location, including depth (z-direction). Displays that can present graphics in different depth planes are called stereoscopic or "three dimensional". These displays address some of the processes (depth cues) that are used by the human vision system to extract depth information from an observed scene in order to build a three-dimensional mental representation of it. Cutting [1] provides an extensive review of the different depth cues and their specific importance.

One of the most important cues is *binocular disparity*, which utilizes the fact that we have two eyes at separate physical locations. Each eye is provided a slightly different perspective of an observed object, and this difference is processed in the brain to extract depth knowledge. (See figure 2)

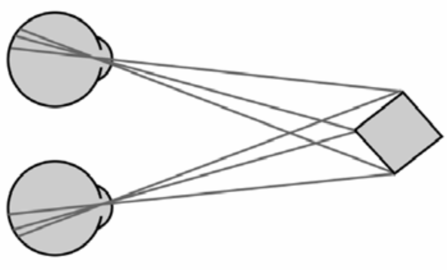


Figure 2: Binocular disparity. Each eye has a slightly different perspective on an object.

If two objects are observed, binocular disparity provides information about their spatial location by correlating the images projected on the two retinas. (See figure 3) As the object moves away from the observer the angle between the light rays reaching the two eyes approaches zero. After about 6 m the rays are so close to parallel that the eyes normally cannot tell the difference.

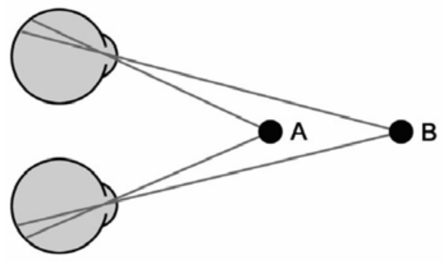


Figure 3: Binocular disparity. The image of the close object (A) falls further from the center of the retina than the far object (B).

Two other important cues are *accommodation* and *vergence*. Accommodation is the physical change in the shape of the eye's lens in order to focus incoming light rays on the retina and make the observed object look sharp. (See figure 4a) Vergence is the change in angle between the optical axes of the two eyes. (See figure 4b) The closer an observed object is, the more the eyes converge in order to make the incoming light rays fall on the retina. However, as with binocular disparity, these cues also have a limited working range. Alone they are effective in determining depth at up to 2 m distance; combined up to 3 m. [1] A point called optical infinity is defined as 6 m from the eyes. [2] When objects beyond this distance are observed the eyes accommodation and vergence are at their maximum regardless of actual distance to the object.

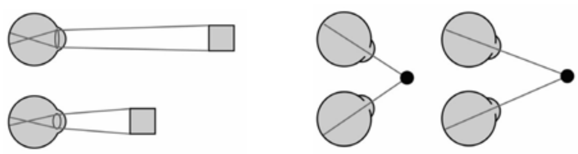


Figure 4: a) Accommodation and b) vergence.

An ATCT involves much larger distances than the maximum working range of these three depth cues. A full integration between the augmenting data layer and the background objects can only be achieved if the data layer is placed beyond the working range of these depth cues, i.e. at or beyond optical infinity. Beyond optical infinity the augmenting data layer will match in depth with all objects behind it, regardless if they are 10, 100 or 1000 m away.

A display device that produces its image at optical infinity is referred to as being collimated at infinity, or simply as a "collimated display" (collimate = make parallel). The light rays from the image in such a display are roughly parallel and the image looks like it is 6 m away or more. Commonly this is performed through collimating optics (lenses and curved mirrors). The resulting image is finally displayed to the user by reflection with a half-silvered mirror, a beam-splitter, which also provides a view of what is behind. (This is usually how an aircraft HUD works.)

Another option is to use a transparent display device, for example a transparent projection screen, which produces the actual image on the screen surface. There are some advantages and disadvantages of this approach. The main disadvantage is that the screen should be placed at optical infinity or beyond to have true collimation. The projection screen must also provide a biocular overlay since the (almost parallel) light rays from a single point in the environment will intersect the display surface at two points before reaching the two eyes. Since the image must be aligned with the observed objects in both eyes, two images, separated by the observer's interpupilary distance (IPD), must be displayed. However, as a projection display surface is positioned in front of the user, both eyes can actually see it. Therefore, the left eye image in the biocular display must be blocked for the right eye and vice versa. In projection screens this can be performed through different multiplexing techniques [3].

To have true collimation in a transparent projection screen, the user-screen distance should be 6 m or more. This is in most cases unacceptably large for an AR installation in general, and the for an ATCT environment in particular. In this experiment we test how users perform in a system with a decreased user-screen distance (2.1 m), while maintaining the correct disparity in the screen. In effect, the binocular disparity and vergence cues tells the brain that the screen (data layer) is located at or beyond optical infinity, while the accommodation cue would say that the screen was 2.1 m away. If the user tries to see both the data layer and the background at the same time, the data layer would seem to be at the same depth as the background but slightly out of focus.

Essentially, what we are doing is decoupling the accommodation and the vergence cues, and thereby we get the accommodation-vergence conflict inherent to most stereoscopic display systems. [4] The negative effect of this conflict is however decreased as the user-screen distance increases [5]. We call this "simulated collimation" and this paper reports on the experimental findings of human performance in such a display condition. The simulated collimation mode is compared to a second mode, the non-collimated mode, where no disparity is present in the transparent projection screen. It is also compared to a third mode called "true collimation" which represents how a perfectly collimated display would perform.

B. Related Work

Fürstenau et al. [6] evaluated the potential performance benefits for a transparent projection screen as a HUD in comparison with traditional computer monitors (head-down displays, HDD) in air traffic control towers. The transparent projection screen was non-collimated, but collimation was mimicked in one case when data was displayed directly on the background. An abstract but complex visual search task was performed, requiring the subjects to switch focus between the different display locations (the "inside" view) and a more distant background plane (the "outside" view). The result was that non-collimated HUD, "collimated" HUD and HDD directly below the background showed similar performance, while a HDD below and horizontally displaced showed poorer performance. Display arrangements that minimized head movements, especially horizontally, were therefore concluded sufficient for the chosen task. A proposed follow-up experiment would involve a higher integration of inside and outside information. This would require parallel information processing instead of successive sub-tasks in the different display locations.

C. Hypothesis

1) Simulated collimation in a transparent projection screen decreases a) reaction time and b) error rate compared to a non-collimated display.

2) Simulated collimation in a transparent projection screen increases a) reaction time and b) error rate compared to a true collimated display.

III. METHOD

A. Experimental Setup

The experiment involved three different display modes (simulated collimation, no collimation and true collimation) in a transparent projection screen. Such a transparent projection screen ("the foreground") was placed 2.1 m in front of the user. 12 m in front of the user, directly behind the transparent projection screen, a traditional opaque projection screen ("the background") was placed. Figure 5 below shows the experimental setup.

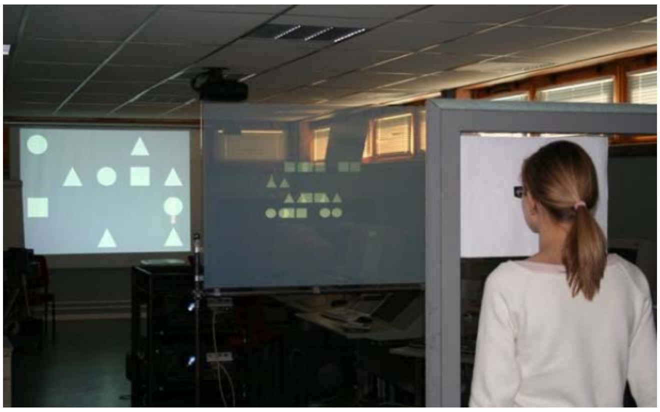


Figure 5: The experimental setup showing (from left to right) the opaque projection screen ("the background"), the transparent projection screen ("the foreground"), and a test subject.

1) Simulated collimation mode

In the simulated collimation mode, data was presented on both the background surface and on the transparent projection screen. The disparity in the projection screen was provided through polarization multiplexing (passive stereo), where the disparate images had linearly perpendicular polarization. The user looked through a pair of similarly polarized glasses (mounted on a stand), which ensured only the correct image was provided to each eye. Thus the user did not have to wear any glasses while performing the experiment.

The disparity had to be calibrated for every test subject's IPD. Calibration was performed using a simple bore-sight test where two disparate vertical lines (one for each eye) in the foreground had to be aligned with a vertical line in the background. The disparity was corrected manually according to user feedback.

No tracking of the subject was performed, so to ensure correct alignment of objects in the foreground and background display planes, the position of the transparent projection screen and the glasses on the mount was fixed throughout the experiment.

2) Non-collimated mode

In the non-collimated mode, no disparity was provided in the transparent projection screen. Only one image was presented in the transparent projection screen and therefore the user did not need to look through polarized glasses. However, in order to maintain alignment as described in the previous mode, the subject looked through two holes (one for each eye) fixed on a mount.

3) True collimation mode

In the true collimation mode, the transparent projection screen was not used at all. Instead, the foreground data was presented together with the background data on the opaque projection screen. This is how true collimation would perform optimally, and onwards we therefore refer to this mode as the "reference mode".

B. Task

Patterns with 20 shapes (triangles, circles and squares) were shown to the participants. The subjects were asked to count the total number of a specific target shape in the presented patterns. The target shape was presented separately before showing the full pattern. The subject was asked to perform the counting as quickly as possible while maintaining reasonable accuracy (i.e. time was prioritized over accuracy).

The previous study from Fürstenau [6] revealed that the actual target shape affected performance. Response times were significantly lower for circles than for triangles and squares, but no difference was found between square and triangle. Additionally the errors occurred more frequently when a square was used as a target shape. To remove any such effects, all patterns had a counterbalanced number of shapes. Also the target shape was balanced within the ten trials.

In the simulated collimation and non-collimated modes, half of the shapes were shown on the transparent projection screen and half on the opaque projection screen (see figure 6a). In the reference mode all 20 shapes were shown on the opaque projection screen (see figure 6b).

All the shapes were white and had the same size. They were shown in a grid layout with dimensions 5 (horizontally) x 4 (vertically). The background color of the opaque projection screen was 50% gray.

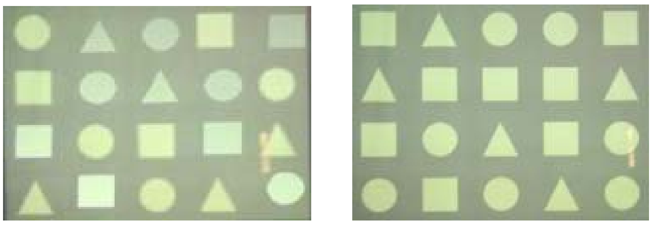


Figure 6: A pattern in a) simulated collimation and non-collimated modes (left) and b) in reference mode (right).

There were 10 unique patterns dedicated for the experimental run and 3 patterns for test trials. All patterns maintained the same complexity regarding the number of each shape, location of target shapes and spatial balance.

The distribution of triangles, squares and circles as target shapes was random. Every pattern consisted of 5 to 9 triangles/squares/circles, thus for each pattern the correct answer was always in the range between 5 and 9.

Since the same 10 patterns were re-used for the three display modes, we needed to eliminate any learning effect. This was accomplished by having different target shapes for the patterns in the different display modes, effectively rendering 30 unique tasks. Moreover, the division of shapes in the foreground and background was changed in the simulated collimation and non-collimated modes to further avoid any learning effect. The order of the patterns was also changed for each display mode.

C. Procedure

The participants were introduced to the experiment by reading an instruction in English explaining the procedure. As described previously, a calibration of the screen's stereo separation (according to the IPD) was performed before the test trials for every subject. The test trials consisted of 3 exercises in simulated collimation mode and 3 exercises in non-collimated mode. The participants were free to ask questions during the test trials. Following the test trials, a set of 10 experimental trials for simulated collimation, non-collimated and reference modes were performed. In total the subjects performed 30 experimental trials.

The order of display mode was counterbalanced between the subjects to minimize any impact of potential learning effects.

The trials (both test and experimental) started with the presentation of a single target shape. Subsequently, the trial was initialized with the user pressing the space bar, whereby the full pattern of shapes was shown. The number of target shapes was counted and after pressing the space bar again, the pattern was removed and the subject reported the answer. The time during which the full pattern was shown was registered. After reporting the answer the next trial was started by pressing the space bar again.

Also the participants were asked to complete a personal data questionnaire and a post-exercise questionnaire, in which they reported their comments about the experiment.

D. *Participants*

There were 26 participants aged 23-56 years. Of these, eight were female and 18 male. All participants were working at the EUROCONTROL Experimental Centre (EEC) and had a university degree. The participants were of several nationalities; four of the participants were native English speakers but all reported fluent English level.

13 participants reported correct vision. The other 13 either wore correction glasses or reported seeing dysfunction (but preferred not to wear glasses during the experiment).

IV. Results

We collected data for 26 subjects. After the initial analysis of the distribution, results of 2 participants were removed due to their exceptional characteristics. The first had average error rates exceeding 400% that of remaining participants in all display modes. The second had average response times exceeding 250% that of remaining participants in all display modes.

In the following tables and graphs simulated collimation mode is abbreviated “Sim Col”, non-collimated mode is “Non Col” and reference mode is “Ref”.

A. *Reaction Time (RT)*

Table 1 below presents the basic distribution characteristic: mean, standard deviation and minimum and maximum values of the results. The highest variation of responses was found in the non-collimated mode.

Table 1
Reaction time – descriptive statistics

	n	Mean	Std. Dev	Min	Max
Sim Col	240	4.349	1.265	2.343	11.453
Non Col	240	4.915	1.718	1.578	14.031
Ref	240	3.694	1.295	1.687	13.921

The results for average reaction time are shown in figure 7 below. The average reaction time for simulated collimation mode was 4.35 s, for non-collimated mode 4.92 s, and for reference mode 3.69 s.

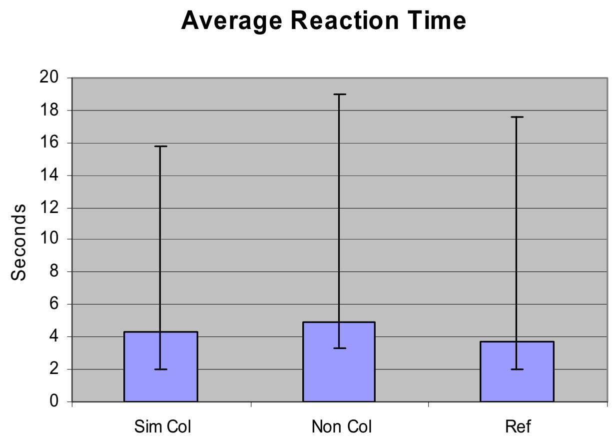


Figure 7: Average reaction time for each display mode. The error bars indicate the maximum and minimum recorded values.

The distribution of data was not satisfying the requirements for parametric tests. The scores were not approximating normal distribution, and an additional Levene test revealed that the variances of results within the three modes were not homogenous (Levene statistic 11.280; p=.000). Therefore the non-parametric Friedman test was used to analyse the effect of display mode.

The mean rank for collimated mode was 2.11, for non collimated mode 2.43 and for reference mode 1.46 (Friedman chi-square 119.252; p=.000, n=240, df=2). These results allowed us to conduct the pair-wise comparison between each mode.

The Wilcoxon signed rank test was applied to compare the results of each mode. The results are presented in table 2 below.

Table 2
Wilcoxon Sign rank test results

	Non Col-Sim Col	Ref-Sim Col	Ref -Non Col
Z	-4.842(a)	-7.773(b)	-10.459(b)
Asymp. Sig. (2-tail)	.000	.000	.000

a) Based on negative ranks
b) Based on positive ranks

The results confirmed that there are significant differences between display modes. Analysis of the means (table 1) shows that the reaction time in the simulated collimation mode was lower than in the non-collimated mode. Furthermore the reaction time in reference mode was lower than in reference mode. These results allowed us to accept hypotheses 1a and 2a.

We tested the impact of wearing glasses (or reporting a seeing dysfunction) on reaction time with the Kruskal-Wallis test, but the results were not significant (chi-square 2.653; df=1, p=0).

Performing the experimental task required the subjects to accommodate their eye lens to different distances. The eye's accommodation abilities decrease after the age of 40 [7]. Also, the visual acuity that contribute to pattern recognition starts to decline at age 30 [8]. To investigate any impact of age in our study, participants were divided into three age groups: less than 29 years (9 subjects), 30-39 years (8 subjects), and over 40 years (7 subjects). The variances within these three groups were not homogenous (Levene statistic 7.227; p=.001), therefore we used the Kruskal-Wallis test. The analysis showed that a significantly lower reaction time was found within the group with subjects over age 40 (chi-square 14.667, n=2, p=.001).

B. Accuracy

The response accuracy was determined by calculating the errors in each individual response given by the subject. The error was defined as the total difference between the correct number of target shapes shown and the number reported by the subject (e.g. if the subject said 5 and the correct number was 7, then the error for that exercise was 2).

The following table presents the mean, standard deviation, minimum and maximum value for each sequence of 10 patterns shown to the subject in each display mode.

TABLE 3
ERROR RATE - DESCRIPTIVE STATISTIC

	Mean	Std. Dev	Min	Max
Sim Col	1.5	1.694	0	6
Non Col	2.0	2.085	0	8
Ref	.96	1.160	0	3

The mean value (average error) for each display mode is shown in figure 8 below.

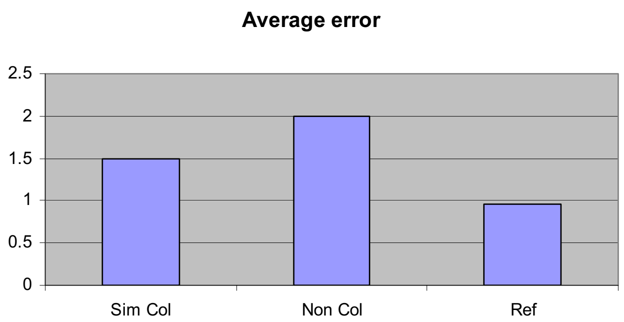


Figure 8: Average error for each display mode.

The highest error rate was shown in the non-collimated mode (2.0), followed by simulated collimation mode (1.5) and reference mode (0.96).

The distribution of errors was not approximating a normal distribution; therefore the non-parametric Friedman Test was applied. We compared the distribution of errors for simulated collimation, non-collimated and reference modes.

The mean rank for collimated mode was 2.02, for non-collimated mode 2.27 and for reference mode 1.71 (chi-square 4.75; n=24, df=2, p=.093). The results were not significant; therefore hypotheses 1b and 2 b are rejected.

For more extensive exploration of the data, we performed additional tests to investigate if error rates were higher for subjects wearing glasses or reporting a seeing dysfunction. We found that the error rate was higher for participants claiming a seeing dysfunction than for those with correct vision only in the reference mode (Kruskal-Wallis chi-square 3.842; df=1, p=.050), while no significant difference was found in simulated collimation or non-collimated modes.

We also performed an analysis of impact of age on error rate. Subjects were divided into there groups, less than 29, 30-39 and over 40 years old, but the results were not significant.

C. Difficulty Level

In the post-exercise questionnaire, the participants were asked to report the relative difficulty for each display mode. The results are presented in figure 9.

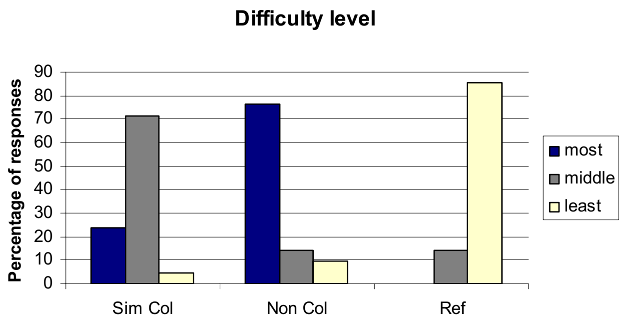


Figure 9: Reported difficulty level for each display mode.

The non-collimated mode was reported to be the most difficult (76%) while simulated collimation mode was reported second difficult (71%). Reference mode was reported as the least difficult mode (85%).

D. Search Strategy

In the post-exercise questionnaire we obtained information about search strategy from 20 subjects. Generally subjects claimed that the strategy depended on display mode.

In simulated collimation and non-collimated mode, most subjects counted shapes combining both display planes, row-by-row or column-by-column. However, some subjects claimed that they changed strategy during the series and counting shapes on the opaque projection screen and the transparent projection screen separately.

In the reference mode subjects preferred to count row-by-row or in clusters, depending on the pattern.

E. *Discomfort*

Stereoscopic vision through polarized glasses was provided in simulated collimation mode. Especially stereoscopic display systems that decouple the accommodation and vergence cues have been reported to have discomfort for the user. Therefore participants were asked for any symptoms of discomfort; eyestrain, headache, sweating, disorientation and dizziness. The only two symptoms claimed were eyestrain (four cases) and disorientation (three cases).

V. DISCUSSION

The differences in reaction time were, as expected, significant between the display modes, and thus confirmed the experimental hypotheses 1a and 2a.

In our additional investigations, it was found that the use of glasses or reporting seeing dysfunction did not have a significant impact on reaction time.

Furthermore, it was found that subjects above 40 years of age had significantly lower reaction times than those in the other age groups. Surprising at first glance, but it might also be related to search strategy. Slower adaptation of eye lens due to age could lead to optimized counting through e.g. clustering, which could render an overall faster response time.

The analysis of response accuracy did not show a significant difference between display modes. The experimental hypotheses 1b and 2b were thus rejected. It might be explained by the fact that the reaction time was prioritized over accuracy.

In our additional investigations, it was found that the use of glasses or claiming a seeing dysfunction had a significant impact on accuracy only in reference mode. This unexpected result requires further investigation with a more detailed analysis addressing the subject's visual system. Furthermore, it was found that age had no significant impact on accuracy.

Levels of discomfort could be decreased further in simulated collimation mode by performing even more accurate calibration of the stereo separation in the transparent projection screen. Further attention could also be given to more accurately matching the light intensities and colors of the two display surfaces.

VI. CONCLUSION AND FUTURE WORK

The main conclusion of this experiment is that simulated collimation in transparent projection screens shows significantly lower response times when performing a visual search task, compared to a non-collimated screen.

The focus switch between near (display) and far (airport) image planes in the ATC tower is a main part of the head down time problem [9]. Even though see-through display devices introduced in ATC towers could minimize head-up/head-down movements, they should also minimize focus switching. As this requires a collimated display, (currently) not possible with transparent projection screens, other approaches are needed. The results in this paper indicate that our approach of simulating collimation could lead to an increased controller performance in AR systems using transparent projection screens.

Future work could involve a visual search task with a higher level of integration between the objects and display planes. This could potentially decrease the importance of, or completely eliminate, any search strategies which involves separating the displayed objects. Intuitively a higher level of integration would further improve response times and accuracy in displays with simulated collimation compared to those with no collimation.

Other future work could involve evaluations of user performance in collimated spatial display devices like the HUD, and comparisons to the traditional HMD.

ACKNOWLEDGMENT

The authors would like to thank Dr. Monica Tavanti for her assistance in experimental design.

The authors would also like to thank Marc Bourgois for his comments on the experiment and for the review of this paper.

This work was funded by PhD scholarships from the EUROCONTROL Experimental Centre, Brétigny-sur-Orge, France.

REFERENCES

[1] J. E. Cutting, "How the eye measures reality and virtual reality", Behavior Research Methods, Instrumentation, and Computers, 29, 29-36, 1997.

[2] N. M. Simonelli, "Apparent size and visual accommodation under day and night conditions", Proceedings of the Human Factors and Ergonomics Society, TX, 23, 374-378, 1979.

[3] D. F. McAllister, "Display Technology: Stereo & 3D Display Technologies", Wiley Encyclopedia on Imaging, Jan volume, 1327-1344, 2002.

[4] D. Drascic and P. Milgram, "Perceptual issues in augmented reality", In M. T. Bolas, S. S. Fisher, and J. O. Merritt, editors, SPIE Volume 2653: Stereoscopic Displays and Virtual Reality Systems III, pages 123–134, January/February 1996.

[5] N. A. Valyus, Stereoscopy, The Focal Press, London and New York, 1966.

[6] N. Fürstenau, M. Rudolph, M. Schmidt, B. Lorenz and T. Albrecht, "On the use of transparent projection screens to reduce head down time in the air traffic control tower", Proc. Human Performance, Situation Awareness and Automation Technology Conference (HPSAA II), Daytona Beach, FL, 2004.

[7] B. Bell, E. Wolf and C. D. Bernholtz, "Depth perception as a function of age", Aging and Human Development, 3, 77-81, 1972.

[8] R. Sekuler, D. Kline and K. Dismukes (Eds.), "Aging and human visual function", New York, NY: Alan R. Liss, Inc, 1982.

[9] B. Hilburn, "Head down time in aerodrome operations: a scope study", Center for Human Performance Research, 2004.

[10] ICAO, "Procedures for Air Navigation Service – Air Traffic Management (PANS-ATM)", Document 4444 ATM/501, International Civil Aviation Organization (ICAO), 2000.

Aeronautical Voice Radio Channel Modelling and Simulation – A Tutorial Review

Konrad Hofbauer[1,2] and Gernot Kubin[1]

[1] Graz University of Technology
SPSC – Signal Processing
and Speech Communication Laboratory
8010 Graz, Austria
konrad.hofbauer@tugraz.at
gernot.kubin@tugraz.at

[2] Eurocontrol Experimental Centre
INO – Innovative Research
91222 Brétigny-sur-Orge, France

***Abstract*— The basic concepts in the modelling and simulation of the mobile radio channel are reviewed. The time-variant channel is dominated by multipath propagation, Doppler effect, path loss and additive noise. Stochastic reference models in the equivalent complex baseband facilitate a compact mathematical description of the channel's input-output relationship. The realisation of these reference models as filtered Gaussian processes leads to practical implementations of frequency selective and frequency nonselective channel models.**

Three different small-scale area simulations of the aeronautical voice radio channel are presented. We demonstrate the practical implementation of a frequency flat fading channel. Based on a scenario in air/ground communication the parameters for the readily available simulators are derived. The resulting outputs give insight into the characteristics of the channel and serve as a basis for the design of digital transmission and measurement techniques.

I. Introduction

Air traffic control (ATC) has relied on the voice radio for communication between aircraft pilots and air traffic control operators since its beginning. The amplitude-modulation (AM) radio, which is in operation worldwide, has basically remained unchanged for decades. Given the aeronautical life cycle constraints, it is expected that the analogue radio will remain in use for ATC voice in Europe well beyond 2020 [1].

Eurocontrol Experimental Centre (France) and Graz University of Technology (Austria) are currently working on embedding supplementary digital data, such as the call-sign of the aircraft, into the voice signal of the analogue air/ground communication (see [2], [3], and Fig. 1). The radio transmission channel has a strong impact on the performance of such a speech watermarking system in terms of data rate and robustness. The degradation of the transmitted signal is studied for the purpose of system design and evaluation. This paper reviews the general concepts of radio channel modelling and demonstrates the use of simulators for the aeronautical voice channel.

Radio channel modelling has a long history and is a very active area of research. This is especially the case with respect to terrestrial mobile radio communications and wide-band data communications due to commercial interest.

However, the results are not always transferable to the aeronautical domain. A comprehensive up-to-date literature review on channel modelling and simulation with the aeronautical radio in mind is provided in [4]. It is highly recommended as a pointer for further reading and its content is not repeated here.

II. Basic Concepts

This and the following section are based on the work of Pätzold [5] and provide a summary of the basic characteristics, the modelling, and the simulation of the mobile radio channel. Another comprehensive treatment on this extensive topic is given in [6].

A. Amplitude Modulation and Complex Baseband

The aeronautical voice radio is based on the double-sideband amplitude modulation (DSB-AM, A3E or simply AM) of a sinusoidal, unsuppressed carrier [7]. An analogue baseband voice signal $x(t)$ which is band-limited to a bandwidth f_m modulates the amplitude of a sinusoidal carrier with amplitude A_0, carrier frequency f_c and initial phase φ_0. The modulated high frequency (HF) signal $x_{AM}(t)$ is defined as

$$x_{AM}(t) = (A_0 + kx(t))\cos(2\pi f_c t + \varphi_0)$$

with the modulation depth

$$m = \frac{|kx(t)|_{max}}{A_0} \leq 1$$

The real-valued HF signal can be equivalently written using complex notation and $\omega_c = 2\pi f_c$ as

$$x_{AM}(t) = Re\left\{(A_0 + kx(t))e^{j\omega_c t}e^{j\varphi_0}\right\} \tag{1}$$

Under the assumption that $f_c \gg f_m$ the HF signal can be demodulated and the input signal $x(t)$ reconstructed by detecting the envelope of the modulated sine wave. The absolute value is low-pass filtered and the original amplitude of the carrier is subtracted.

$$x(t) = \frac{1}{k}([|x_{AM}(t)|]_{LP} - A_0)$$

Fig. 2 shows the spectra of the baseband signal and the corresponding HF signal. Since the baseband signal is, by definition, low-pass filtered, the HF signal is a bandpass signal

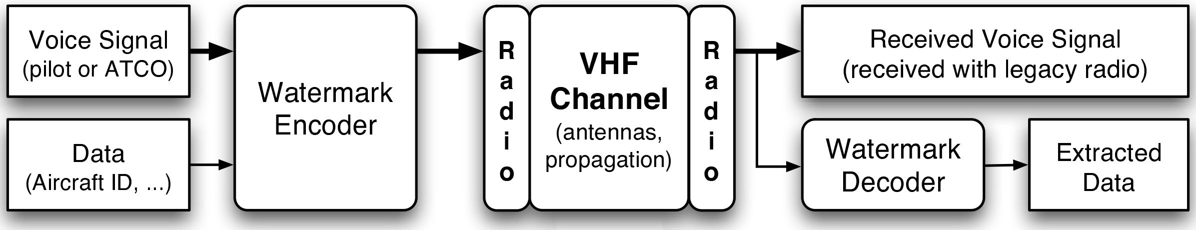


Fig. 1. General structure of a speech watermarking system for the aeronautical voice radio.

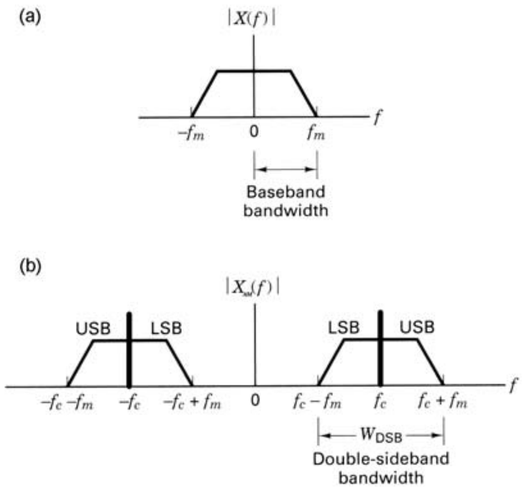


Fig. 2. Signal spectra of (a) the baseband signal $x(t)$ and (b) the HF signal $x_{AM}(t)$ with a carrier at f_c and symmetric upper and lower sidebands. [8, with modification]

and contains energy only around the carrier frequency and the lower and upper sidebands LSB and USB.

In general, any real bandpass signal $s(t)$ can be represented as the real part of a modulated complex signal,

$$s(t) = Re\left\{g(t)e^{j\omega_c t}\right\} \tag{2}$$

where $g(t)$ is called the *equivalent complex baseband* or *complex envelope* of $s(t)$ [8]. The complex envelope $g(t)$ is obtained by downconversion of the real passband signal $s(t)$, namely

$$g(t) = (s(t) + j\hat{s}(t))e^{-j\omega_C t}$$

with $\hat{s}(t)$ being the Hilbert transform of $s(t)$. The Hilbert transform removes the negative frequency component of $s(t)$ before downconversion [9]. A comparison of Eq. 1 and Eq. 2 reveals that the complex envelope $g_{AM}(t)$ of the amplitude modulated HF signal $x_{AM}(t)$ simplifies to

$$g_{AM}(t) = (A_0 + kx(t))e^{j\varphi_0}$$

The signal $x(t)$ is reconstructed from the equivalent complex baseband of the HF signal by demodulation with

$$x(t) = \frac{1}{k}\left(|g_{AM}(t)| - A_0\right) \tag{3}$$

The complex baseband signal can be pictured as a time-varying phasor or vector in a rotating complex plane. The rotating plane can be seen as a coordinate system for the vector, which rotates with the angular velocity ω_c.

In order to represent the HF signal as a *discrete-time* signal, it must be sampled with a frequency of more than twice the carrier frequency. This leads to a large number of samples and thus makes numerical simulation difficult even for very short signal durations. Together with the carrier frequency ω_c the complex envelope $g_{AM}(t)$ fully describes the HF signal $x_{AM}(t)$. The complex envelope $g_{AM}(t)$ has the same bandwidth $[-f_m; f_m]$ as the baseband signal $x(t)$. As a consequence it can be sampled with a much lower sampling frequency, which facilitates efficient numerical simulation without loss of generality. Most of the existing channel simulations are based on the complex baseband signal representation.

B. Mobile Radio Propagation Channel

Proakis [7] defines the communication channel as "... the physical medium that is used to send the signal from the transmitter to the receiver." Radio channel modelling usually also includes the transmitting and receiving antennas in the channel model.

1) Multipath Propagation: The transmitting medium in radio communications is the atmosphere or free space, into which the signal is coupled as electromagnetic energy by an antenna. The received electromagnetic signal can be a superposition of a line-of-sight path signal and multiple waves coming from different directions. This effect is known as *multipath propagation*. Depending on the geometric dimensions and the properties of the objects in a scene, an electromagnetic wave can be reflected, scattered, diffracted or absorbed on its way to the receiver.

From hereon we assume, without loss of generality, that the ground station transmits and the aircraft receives the radio signal. The effects treated in this paper are identical for both directions. As illustrated in Fig. 3, reflected waves have to travel a longer distance to the aircraft and therefore arrive with a time-delay compared to the line-of-sight signal. The received signal is spread in time and the channel is said to be *time dispersive*. The time delays correspond to phase shifts in between the superimposed waves and lead to constructive or destructive interference depending on the position of the aircraft. As both the position and the phase shifts change

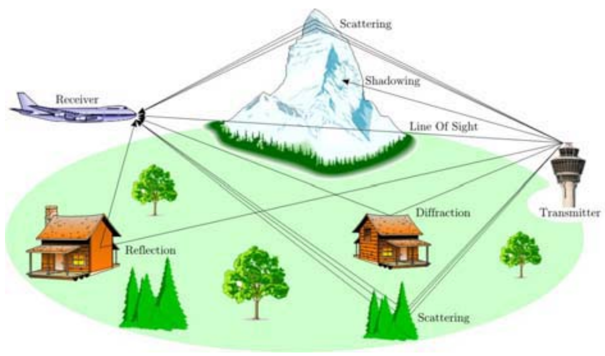


Fig. 3. Multipath propagation in an aeronautical radio scenario. [10]

constantly due to the movement of the aircraft, the signal undergoes strong pseudo-random amplitude fluctuations and the channel becomes a *fading channel*.

The *multipath spread* T_m is the time delay between the arrival of the line-of-sight component and the arrival of the latest scattered component. Its inverse $B_{CB} = \frac{1}{T_M}$ is the *coherence bandwidth* of the channel. If the *frequency bandwidth* W of the transmitted signal is larger than the coherence bandwidth ($W > B_{CB}$), the channel is said to be *frequency selective*. Otherwise, if $W < B_{CB}$, the channel is *frequency nonselective* or *flat fading*. This means that all the frequency components of the received signal are affected by the channel always in the same way [7].

2) Doppler Effect: The so-called *Doppler effect* shifts the frequency content of the received signal due to the movement of the aircraft relative to the transmitter. The Doppler frequency f_D, which is the difference between the transmitted and the received frequency, is dependent on the angle of arrival α of the electromagnetic wave relative to the heading of the aircraft.

$$f_D = f_{D,max}\cos(\alpha)$$

The maximum Doppler frequency $f_{D,max}$, which is the largest possible Doppler shift, is given by

$$f_{D,max} = \frac{v}{c} f_c \tag{4}$$

where v is the aircraft speed, f_c the carrier frequency and $c = 3 \cdot 10^8\ \frac{\text{m}}{\text{s}}$ the speed of light.

The reflected waves arrive not only with different time-delays compared to the line-of-sight signal, but as well from different *directions* relative to the aircraft heading (Fig. 3). As a consequence, they undergo different Doppler shifts. This results in a continuous distribution of frequencies in the spectrum of the signal and leads to the so-called *Doppler power spectral density* or simply *Doppler spectrum.*

3) Channel Attenuation: The signal undergoes significant attenuation during transmission. The *path loss* is dependent on the distance d and the obstacles between transmitter and receiver. It is proportional to $\frac{1}{d^p}$, with the pathloss exponent p in the range of $2 \leq p < 4$. In the optimal case of line-of-sight free space propagation $p = 2$.

4) Additive Noise: During transmission *additive noise* is imposed onto the signal. The noise results, among others, from thermal noise in electronic components, from atmospheric noise or radio channel interference, or from man-made noise such as engine ignition noise.

5) Time Dependency: Most of the parameters described in this section vary over time due to the movement of the aircraft. As a consequence the response of the channel to a transmitted signal also varies, and the channel is said to be *time-variant*.

C. Stochastic Terms and Definitions

The following section recapitulates some basic stochastic terms in order to clarify the nomenclature and notation used here. The reader is encouraged to refer to [5] for exact definitions.

Let the *event* A be a collection of a number of possible *outcomes* s of a random experiment, with the real number $P(A)$ being its *probability* measure. A *random variable* μ is a mapping that assigns a real number $\mu(s)$ to every outcome s. The *cumulative distribution function*

$$F_\mu(x) = P(\mu \leq x) = P(\{s|\mu(s) \leq x\})$$

is the probability that the random variable μ is less or equal to x. The *probability density function* (*pdf*, or simply *density*) $p_\mu(x)$ is the derivative of the cumulative distribution function,

$$p_\mu(x) = \frac{dF_\mu(x)}{dx}$$

The most common probability density functions are the *uniform distribution*, where the density is constant over a certain interval and is zero outside, and the *Gaussian distribution* or *normal distribution* $N(m_\mu, \sigma_\mu^2)$, which is determined by the two parameters expected value m_μ and variance σ_μ^2.

With μ_1 and μ_2 being two statistically independent normally distributed random variables with identical variance σ_0^2, the new random variable $\zeta = \sqrt{\mu_1^2 + \mu_2^2}$ represents a *Rayleigh distributed* random variable (Fig. 4(a)). Given an additional real parameter ρ, the new random variable $\xi = \sqrt{(\mu_1 + \rho)^2 + \mu_2^2}$ is *Rice* or *Rician distributed* (Fig. 4(b)). A random variable $\lambda = e^\mu$ is said to be *lognormally distributed*. A multiplication of a Rayleigh and a lognormally distributed random variable $\eta = \zeta\lambda$ leads to the so-called *Suzuki distribution.*

A *stochastic process* $\mu(t, s)$ is a collection of random variables, which is indexed by a *time* index t. At a fixed time instant $t = t_0$, the value of a random process, $\mu(t_0, s)$, is a random variable. On the other hand, in the case of a fixed outcome $s = s_0$ of a random experiment, the value of the stochastic process $\mu(t, s_0)$ is a time function, or signal, that corresponds to the outcome s_0. As in common practise, the variable s is dropped in the notation for a stochastic process and $\mu(t)$ written instead. With $\mu_1(t)$ and $\mu_2(t)$ being two real-valued stochastic processes, a *complex-valued stochastic process* is defined by $\mu(t) = \mu_1(t) + j\mu_2(t)$. A stochastic process is called stationary if its statistical properties are invariant to a shift in time. The Fourier transform of the autocorrelation function of such a stationary process defines

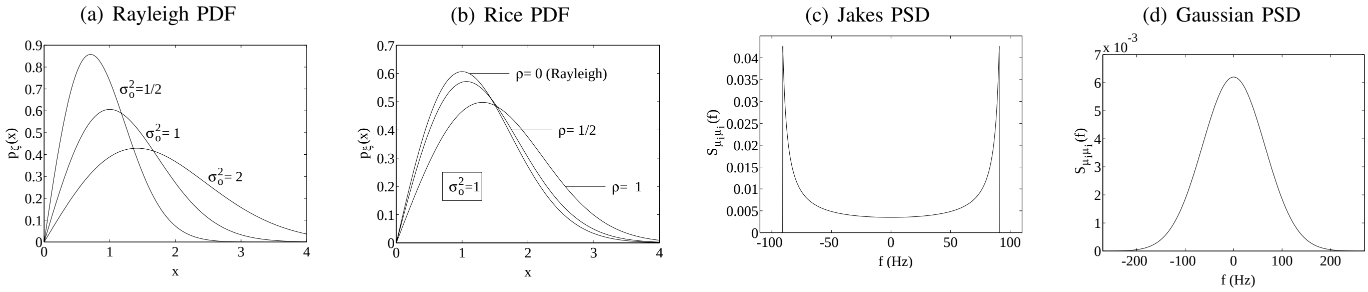


Fig. 4. Probability density functions (PDF) and power spectral densities (PSD, $f_{D,max} = 91\,\mathrm{Hz}$, $\sigma_0^2 = 1$) for Rayleigh and Rice channels. [5]

the *power spectral density* or *power density spectrum* of the stochastic process.

III. Radio Channel Modelling

Sec. II-B.1 illustrated the effect of multipath propagation from a geometrical point of view. However, geometrical *modelling* of the multipath propagation is possible only to a very limited extent. It requires detailed knowledge of the geometry of all objects in the scene and their electromagnetic properties. The resulting simulations are time consuming to set up and computationally expensive, and a number of simplifications have to be made. Furthermore the results are valid for the specific situation only and can not always be generalised. As a consequence, a stochastic description of the channel and its properties is widely used. It focuses on the distribution of parameters over time instead of trying to predict single values. This class of *stochastic channel models* is the subject of the following investigations.

In *large-scale* areas with dimensions larger than tens of wavelengths of the carrier frequency f_c, the local mean of the signal envelope fluctuates mainly due to shadowing and is found to be approximately lognormally distributed. This *slow-term fading* is important for channel availability, handover, and mobile radio network planning.

More important for the design of a digital transmission technique is the fast signal fluctuation, the *fast-term fading,* which occurs within small areas. As a consequence, we focus on models that are valid for *small-scale areas*, where we can assume the path loss and the local mean of the signal envelope due to shading, etc. , to be constant. Furthermore we assume for the moment a frequency nonselective channel and, for mathematical simplicity, the transmission of an unmodulated carrier.

A. Stochastic Mutlipath Reference Models

The sum $\mu(t)$ of all scattered components of the received signals can be assumed to be normally distributed. If we let $\mu_1(t)$ and $\mu_2(t)$ be two zero-mean statistically independent Gaussian processes with variance σ_0^2, then the sum of the scattered components is given in complex baseband representation as a zero-mean complex Gaussian process $\mu(t)$ and is defined by

$$\text{Scatter:} \qquad \mu(t) = \mu_1(t) + j\mu_2(t) \tag{5}$$

The line-of-sight (LOS) signal component $m(t)$ is given by

$$\text{LOS:} \qquad m(t) = A_0 e^{j(2\pi f_D + \varphi_0)} \tag{6}$$

again in complex baseband representation.

The superposition $\mu_m(t)$ of both signals is

$$\text{LOS+Scatter:} \qquad \mu_m(t) = m(t) + \mu(t) \tag{7}$$

Depending on the surroundings of the transmitter and the receiver, the received signal can consists of either the scatter components only or of a superposition of LOS and scatter components. In the first case (Eq. 5) the magnitude of the complex baseband signal $|\mu(t)|$ is Rayleigh distributed. Its phase $\angle(\mu(t))$ is uniformly distributed over the interval $[-\pi; \pi)$. This type of a *Rayleigh fading channel* is predominant in regions where the LOS component is blocked by obstacles, such as in urban areas with high buildings, etc.

In the second case where a LOS component and scatter components are present (Eq. 7), the magnitude of the complex baseband signal $|\mu(t) + m(t)|$ is Rice distributed. The Rice factor k is determined by the ratio of the power of the LOS component and the scatter components, where $k = \frac{A_0^2}{2\sigma_0^2}$. This *Rice fading channel* dominates the aeronautical radio channel.

One can derive the probability density of amplitude and phase of the received signal based on the Rice or Rayleigh distributions. As a further step, it is possible to compute the *level crossing rate* and the *average duration of fades*, which are important measures required for the optimisation of coding systems in order to address burst errors. The exact formulae can be found in [5] and are not reproduced here.

The power spectral density of the complex Gaussian random process in Eq. 7 corresponds to the Doppler power spectral density when considering the power of all components, their angle of arrival and the directivity of the receiving antenna. Assuming a Rayleigh channel with no LOS component, propagation in a two-dimensional plane and uniformly distributed angles of arrival, one obtains the so-called Jakes power spectral density as the resulting Doppler spectrum. Its shape is shown in Fig. 4(c).

However, both theoretical investigations and measurements have shown that the assumption that the angle of arrival of the scattered components is uniformly distributed does in practise not hold for aeronautical channels. This results in a Doppler spectrum which is significantly different from the

Jakes spectrum [11]. The Doppler power spectral density is therefore better approximated by a Gaussian power spectral density, which is plotted in Fig. 4(d). For nonuniformly distributed angles of arrival, as with explicit directional echos, the Gaussian Doppler PSD is unsymmetrical and shifted away from the origin. The characteristic parameters describing this are the *average Doppler shift* (the statistic mean) and the *Doppler spread* (the square root of the second central moment) of the Doppler PSD .

B. Realisation of the Reference Models

The above reference models are based on coloured Gaussian random processes. The *realisation* of these processes is not trivial and leads to the theory of deterministic processes. Mostly two fundamental methods are applied in the literature in order to generate coloured Gaussian processes. In the *filter method*, white Gaussian noise is filtered by an ideal linear time-invariant filter with the desired power spectrum. In the *Rice method* an infinite number of weighted sinusoids with equidistant frequencies and random phases are superimposed. In practise both methods can only approximate the coloured Gaussian process. Neither an *ideal* filter nor an *infinite* number of sinusoids can be realised. A large number of algorithms used to determine the actual parameters of the sinusoids in the Rice method exist. The methods approximate the Gaussian processes with a sum of a limited number of sinusoids, thus considering the computational expense [5]. For the filter method on the other hand, the problem boils down to filter design with its well-understood limitations.

C. Frequency Nonselective Channel Models

In frequency nonselective flat fading channels, all frequency components of the received signal are affected by the channel in the same way. *The channel is modelled by a multiplication of the transmitted signal with a suitable stochastic model process.* The Rice and Rayleigh processes described in Sec. III-A can serve as statistical model processes.

However it has been shown that the Rice and Rayleigh processes often do not provide enough flexibility to adapt to the statistics of real world channels. This has led to the development of more versatile stochastic model processes such as the Suzuki process and its variations (a product of a lognormal distribution for the slow fading and a Rayleigh distribution for the fast fading), the Loo Model with its variations, and the generalised Rice process.

D. Frequency Selective Channel Models

Where channel bandwidth and data rate increase, the propagation delays can no longer be ignored as compared to the symbol interval. The channel is then said to be frequency selective and over time the different frequency components of a signal are affected differently by the channel.

1) Tapped Delay Line Structure: For the modelling of a frequency selective channel, a tapped delay line structure is typically applied as reference model (Fig. 5). The ellipses model of Parsons and Bajwa [6] shows that all reflections and scatterings from objects located on an ellipse, with the transmitter and receiver in the focal points, undergo the same time delay. This leads to a complex Gaussian distribution of the received signal components for a given time delay, assuming a large number of objects with different reflection properties in the scene and applying the the central limit theorem. As a consequence, the tap weights $c_i(t)$ of the single paths are assumed to be uncorrelated complex Gaussian processes. It is shown in Sec. II-C that the amplitudes of the complex tap weights are then either Rayleigh or Rice distributed, depending on the mean of the Gaussian processes. An analytic expression for the phases of the tap weights can be found in [5].

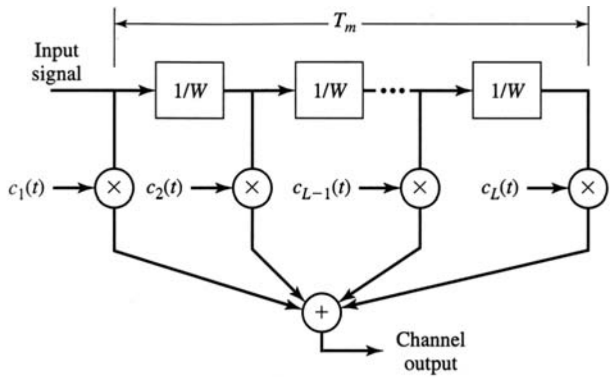


Fig. 5. Tapped delay line structure as a frequency selective and time variant channel model. W is the bandwidth of the transmitted signal, $c_i(t)$ are uncorrelated complex Gaussian processes. [5]

2) Linear Time-Variant System Description: The radio channel can be modelled as a linear time-variant system, with input and output signals in the complex baseband representation. The system can be fully subscribed by its time-variant impulse response $h(t, \tau)$. In order to establish a statistical description of the input/output relation of the above system, the channel is further considered as a stochastic system, with $h(t, \tau)$ as its stochastic system function.

These input/output relations of the stochastic channel can be significantly simplified assuming that the impulse response $h(t, \tau)$ is wide sense stationary[1] (*WSS*), and assuming that scattering components with different propagation delays are statistically uncorrelated (*U*ncorrelated *S*cattering).

Based on these two assumptions, Bello proposed in 1963 the class of *WSSUS* models. They are nowadays widely used and are of great importance in channel modelling. They are based on the tapped delay line structure and allow the computation of all correlation functions, power spectral densities and properties such as Doppler and delay spread, etc., from a given *scattering function*. The scattering function may be obtained by the measurement of real channels, by specification, or both. For example, the European working group 'COST 207' published scattering functions in terms of delay power spectral densities and Doppler power spectral densities for four

[1]Measurements have shown that this assumption is valid for areas smaller than tens of wavelengths of the carrier frequency f_c.

propagation environments which are claimed to be typical for mobile cellular communication.

E. AWGN Channel Model

The noise that is added to the transmitted signal during transmission is typically represented as an additive white Gaussian noise (AWGN) process. The main parameter of the model is the variance σ_0^2 of the Gaussian process, which together with the signal power defines the signal-to-noise ratio (SNR) of the output signal [7]. The AWGN channel is usually included as an additional block after the channel models described above.

IV. Voice Channel Simulation

This section aims to present three different simulators which implement the above radio channel models. We first define a simulation scenario based on which we show the simulators' input parameters and the resulting channel output. We use as example the aeronautical VHF voice radio channel between a fixed ground station and a general aviation aircraft which is flying at its maximum speed.

For all simulators the same pre- and post-processing of the input and output signals is used. It is based on a Matlab implementation of the filtering, amplitude modulation and demodulation in the equivalent complex baseband and receiver gain control. The input and output signals of all the simulators used are represented in the equivalent complex baseband.

A. Simulation Scenario and Parameters

For air-ground voice communication in civil air traffic control, the carrier frequency f_c is within a range from 118 MHz to 137 MHz, the 'very high frequency' (VHF) band. The 760 channels are spaced 25 kHz apart. The channel spacing is reduced to 8.33 kHz in specific regions of Europe in order to increase the number of available channels to a theoretical maximum of 2280. According to specification, the frequency response of the transmitter is required to be flat between 0.3 kHz to 2.5 kHz with a sharp cut-off below and above this frequency range [12].

For the simulation, we assume a carrier with amplitude $A_0 = 1$, frequency $f_c = 120\,\text{MHz}$ and initial phase $\varphi_0 = \frac{\pi}{4}$, a channel spacing of $W = 8.33\,\text{kHz}$, a modulation depth $m = 0.8$ and an input signal which is band-limited to $f_l = 300\,\text{Hz}$ to $f_m = 2.5\,\text{kHz}$. For the illustrations we use a purely sinusoidal baseband input signal $x(t) = \sin(2\pi f_a t)$ with $f_a = 500\,\text{Hz}$, which is sampled with a frequency of $f_{sa} = 8000\,\text{Hz}$ and bandpass filtered according to the above specification. Fig. 6 shows all values that the amplitude modulated signal $x_{AM}(t)$ takes during the observation interval in the equivalent complex baseband representation $g_{AM}(t)$. The white circle represents the *unmodulated* carrier signal, which is a single point in the equivalent complex baseband. A short segment of the magnitude of the signal, $|g_{AM}(t)|$, is also given.

In the propagation model a general aviation aircraft with a speed of $v = 60\,\frac{\text{m}}{\text{s}}$ is used. This results in a maximum Doppler frequency of $f_{D,max} = 24\,\text{Hz}$ (Eq. 4). Given the

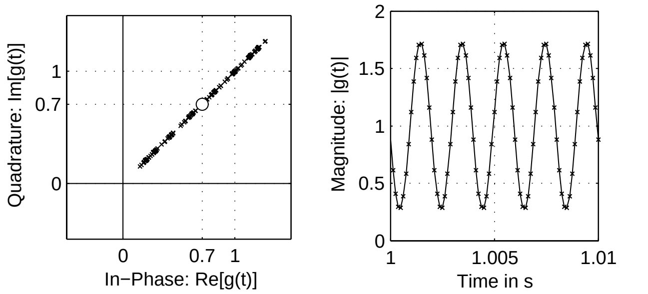


Fig. 6. Sinusoidal AM signal in equivalent complex baseband. Left: In-phase and quadrature components. The white circle indicates an unmodulated carrier. Right: The magnitude of $g(t)$.

carrier frequency f_c, the wavelength is $\lambda = \frac{c}{f_c} = 2.5\,\text{m}$. This distance λ is covered in $t_\lambda = 0.0417\,\text{s}$. Furthermore, we assume that the aircraft flies at a height of $h_2 = 3000\,\text{m}$ and at a distance of $d = 10\,\text{km}$ from the ground station. The ground antenna is taken to be mounted at a height of $h_1 = 20\,\text{m}$. The geometric path length difference Δl between the line-of-sight path and the dominant reflection along the vertical plane on the horizontal flat ground evaluates to

$$\Delta l = \sqrt{h_1^2 + \left(\frac{h_1 d}{h_1 + h_2}\right)^2} + \sqrt{h_2^2 + \left(d - \frac{h_1 d}{h_1 + h_2}\right)^2} - \sqrt{d^2 + (h_2 - h_1)^2} = 11.5\,\text{m}$$

which corresponds to a path delay of $\Delta\tau = 38.3\,\text{ns}$. In the worst case scenario of $T_m = 10\Delta\tau$, the coherence bandwidth is $B_{CB} = 2.6\,\text{MHz}$. With $B_{CB} \gg W$, according to Sec. II-B.1 the channel is surely frequency nonselective. Worst-case multipath spreads of $T_m = 200\,\mu\text{s}$ as reported in [11] cannot be explained with a reflection in the vertical plane, but only with a reflection on far-away steep slopes. In these rare cases, the resulting coherence bandwidth is in the same order of magnitude as the channel bandwidth.

We cannot confirm the rule of thumb given in [11] where $\Delta l \approx h_2$ given $d \gg h_2$. For example, a typical case for commercial aviation where $h_1 = 30\,\text{m}$, $h_2 = 8000\,\text{m}$ and $d = 100\,\text{km}$ results in a path difference of $\Delta l = 4.78\,\text{m}$. In the special case of a non-elevated ground antenna with $h_1 \approx 0$ the path delay vanishes. The Rician factor k is assumed to be $k = 12\,\text{dB}$, which corresponds to a fairly strong line-of-sight signal [11].

B. Mathworks Communications Toolbox Model

The Mathworks Communications Toolbox for Matlab [13] implements a multipath fading channel model. The simulator supports multiple fading paths, of which the first is Rice or Rayleigh distributed and the subsequent paths are Rayleigh distributed. The Doppler spectrum is approximated by the Jakes spectrum. As shown in Sec. III-A, the Jakes Doppler spectrum is not suitable for the aeronautical channel. The preferable Gaussian Doppler spectrum is unfortunately not supported by the simulator. The toolbox provides a convenient a tool for the visualisation of impulse and frequency response, gain and phasor of the multipath components, and the evolution of these quantities over time.

In terms of implementation, the toolbox models the channel as a time-variant linear FIR filter. Its tap-weights $g(m)$ are given by a sampled and truncated sum of shifted sinc functions. They are shifted by the path delays τ_k of the k^{th} path, weighted by the average power gain p_k of the corresponding path, and weighted by a random process $h_k(n)$. The uncorrelated random processes $h_k(n)$ are filtered Gaussian random processes with a Jakes power spectral density.

$$g(m) = \sum_k \operatorname{sinc}\left(\frac{\tau_k}{1/f_{sa}} - m\right) h_k(n) p_k$$

The equation shows once again that when all path delays are small as compared to the sample period, the sinc terms coincide. This result in a filter with only one tap and consequently in a frequency nonselective channel.

In our scenario the channel is frequency-flat, and a model according to Sec. III-C with one Rician path is appropriate. The only necessary input parameters for the channel model are f_{sa}, $f_{D,max}$ and k.

The output of the channel for the sinusoidal input signal as defined above is shown in Fig. 7(a). The demodulated signal (see Eq. 3 and Fig. 7(b)) reveals the amplitude fading of the channel due to the Rician distribution of the signal amplitude. It is worthwhile noticing that the distance between two maxima is roughly one wavelength λ. This fast-term fading results from the superposition of the line-of-sight component and the multitude of scattered components with Gaussian distribution. As mentioned above, this is under the small-scale area assumption where path loss and shading are assumed to be constant.

Fig. 7(c) and 7(d) show the demodulated signal after bandpass filtering and after automatic gain control, respectively. Amplitude modulations of the carrier wave with a frequency of less than $f_l = 0.3\,\text{kHz}$ are not caused by the input signal, as it is band-limited, but by the channel. These modulations scale the *entire* amplitude modulated carrier signal due to the frequency nonselectiveness of the channel. The automatic gain control can therefore detect these low-frequency modulations and compensate for them. This eliminates signal fading with frequencies of up to $f_l = 300\,\text{Hz}$.

The toolbox also allows a model structure with several discrete paths similar to Fig. 5. One can specify the delay and the average power of each path. A scenario similar to the first one with *two* distinct paths is shown for comparison. We define one Rician path with a Rician factor of $k = 200$. This means that it contains only the line of sight signal and no scattering. We furthermore define one Rayleigh path with a relative power of -12 dB and a time delay of $\Delta\tau = 38.3\,\text{ns}$, both relative to the LOS path.

Due to the small bandwidth of our channel, the results are equivalent to the first scenario. Fig. 8 shows the time-variation of the *power* of the two components, with the total power being normalised to 0 dB.

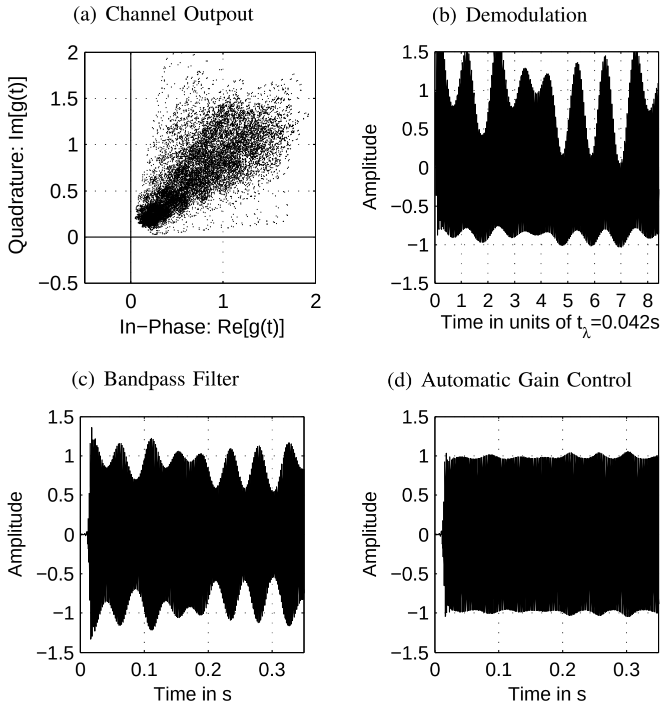


Fig. 7. Received signal at different processing stages. Received signal (channel output of Mathworks Communication Toolbox and an observation interval of 2 s) in equivalent complex baseband (a), after demodulation (b), after bandpass filtering (c), and after automatic gain control (d).

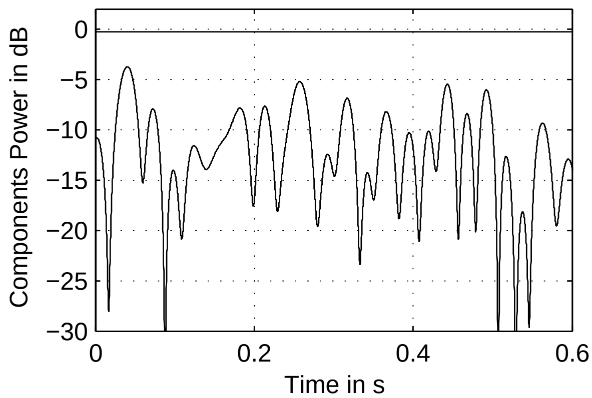


Fig. 8. Power of the line-of-sight component (top) and the Rayleigh distributed scattered components (bottom).

C. The Generic Channel Simulator

The Generic Channel Simulator (GCS) is a radio channel simulator which was developed between 1994 and 1998 under contract of the American Federal Aviation Administration (FAA). Its source code and documentation is provided in [14]. The GCS written in ANSI C and provides a graphical MOTIF-based interface and a command line interface to enter the model parameters. Data input and output files are in a binary floating point format and contain the signal in equivalent complex baseband representation. The last publicly available version of the software dates back to 1998. This version requires a fairly complex installation procedure and a number of adjustments in order to enable compiling of the source code on current operating systems. We provide some advice on how to install the software on the Mac OS X operating system and how to interface the simulator with Matlab [15].

The GCS allows the simulation of various types of mobile radio channels, the VHF air/ground channel among others. Similar to the Mathworks toolbox, the GCS simulates the radio channel by a time-variant IIR filter. The scatter path delay power spectrum shape is approximated by a decaying exponential multiplying a zeroth order modified Bessel function, the Doppler power spectrum is assumed to have a Gaussian shape.

In the following example we use a similar setup as in the second scenario in Sec. IV-B, a discrete line-of-sight signal and a randomly distributed scatter path with a power of -12 dB. With the same geometric configuration as above, the GCS software confirms our computed time delay of $\Delta\tau = 38.3$ ns.

Using the same parameters for speed, geometry, frequency, etc., as in the scenarios described above, we obtain a channel output as shown in Fig. 9.

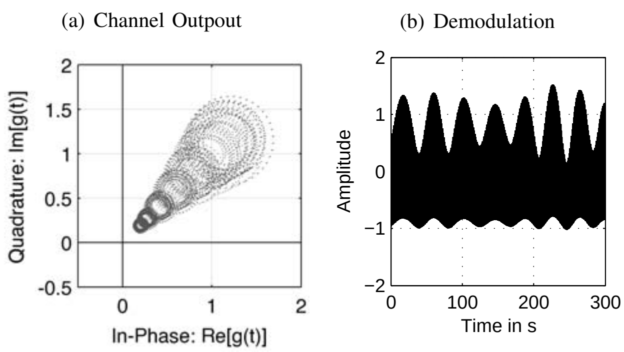


Fig. 9. Generic Channel Simulator: Received signal in an observation interval of 300 s (channel output) in equivalent complex baseband (a) and after demodulation (b).

The time axis in the plot of the demodulated signal is very different as compared to the one in Fig. 7(c). The amplitude fading of the signal has a similar shape as before but is in the order of *three magnitudes* slower than observed in Sec. IV-B. This contradicting result cannot be explained by the differing model assumptions, nor does it coincide with first informal channel measurements that we pursued.

We believe that the out-dated source code of the GCS has legacy issues which lead to problems with current operating system and compiler versions.

D. Direct Reference Model Implementation

The third channel simulation is based on a direct implementation of the concept described in Sec. III-C with the Matlab routines provided in [5]. We model the channel by multiplying the input signal with the complex random process $\mu_m(t)$ as given in Eq. 7. The random processes are generated with the sum of sinusoids approach as discussed in Sec. III-B. Fig. 10 shows the channel output using a Jakes Doppler PSD and a Rician reference model. The result is very similar to the one obtained with the Mathworks toolbox.

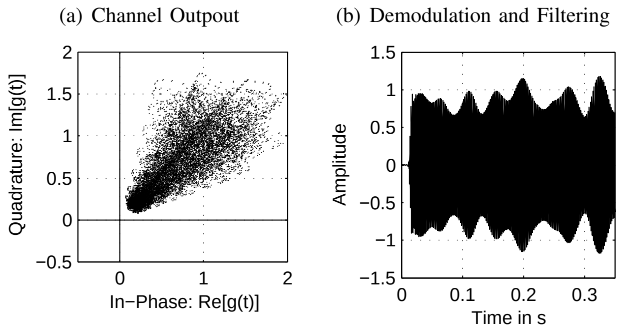


Fig. 10. Reference channel model with Jakes PSD: Received signal (channel output) in equivalent complex baseband (a) and after demodulation and bandpass filtering (b).

In a second example, a Gaussian distribution is used for the Doppler power spectral density instead of the Jakes model. The additional parameter $f_{D,co}$ describes the width of the Gaussian PSD by its 3 dB cut-off frequency. The value is arbitrarily set to $f_{D,co} = 0.3 f_{D,max}$. This corresponds to a fairly small Doppler spread of $B = 6.11$ Hz, compared to the Jakes PSD with $B = 17$ Hz. Fig. 11 shows the resulting discrete Gaussian PSD. The channel output in Fig. 12 confirms the smaller Doppler spread by a narrower lobe in the scatter plot. The amplitude fading is by a factor of two slower than with the Jakes PSD.

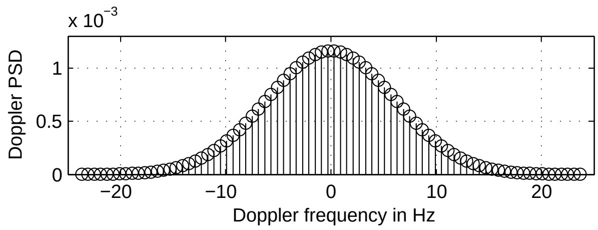


Fig. 11. Discrete Gaussian Doppler PSD with $f_{D,max} = 24$ Hz and a 3-dB-cut-off frequency of $f_{D,co} = 7.2$ Hz.

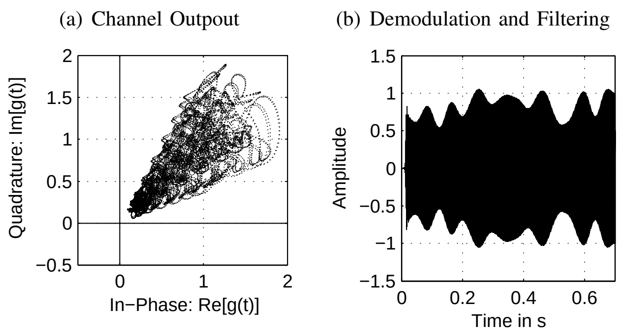


Fig. 12. Reference channel model with Gaussian PSD: Received signal (channel output) in equivalent complex baseband (a) and after demodulation and bandpass filtering (b).

V. Conclusion

The channel simulations presented here serve as a basis for the design of a measurement system to be used in upcoming tests on the aeronautical voice radio channel [16], [17]. The following example illustrates the usefulness of the simulations.

Fig. 13(a) shows the path gain of the frequency nonselective channel from the simulator in Sec. IV-B. Since the channel is flat fading, the path gain corresponds to the amplitude fading of the received signal. The magnitude spectrum of the path gain in Fig. 13(b) reveals that the maximum rate of change of the path gain approximately coincides with the maximum Doppler frequency $f_{D,max} = \frac{v}{c} f_c$ of the channel. The amplitude fading is band-limited to the maximum Doppler frequency.

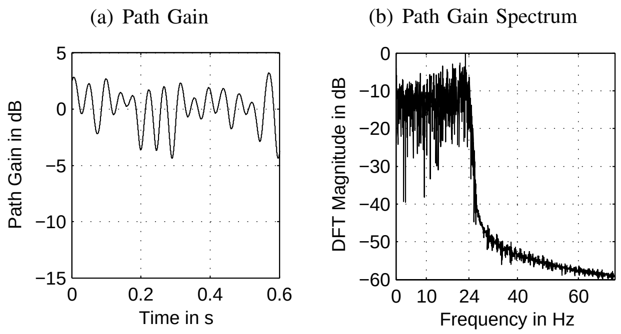


Fig. 13. Time-variant path gain and its magnitude spectrum for a flat fading channel with $f_{D,max} = 24\,\text{Hz}$.

This fact can also be explained with the concept of *beat* as known in acoustics [18]. The superposition of two sinusoidal waves with slightly different frequencies f_1 and f_2 leads to a special type of interference, where the *envelope* of the resulting wave modulates with a frequency of $f_1 - f_2$.

The maximum frequency difference $f_1 - f_2$ between a scatter component and the carrier frequency f_c with which we demodulate in the simulation is given by the maximum Doppler frequency. This explains the band-limitation of the amplitude fading.

However, in a real world system a coherent receiver possibly demodulates the HF signal with a reconstructed carrier frequency $\hat{f}_c$ which is *already* Doppler shifted. In this case, the maximum frequency difference between Doppler shifted carrier and Doppler shifted scatter component is $2f_{D,max}$. This maximum occurs when the aircraft points towards the ground station so that the LOS signal arrives from in front of the aircraft, and when at the same time a scatter component arrives from the the back of the aircraft [11]. We can thereof conclude that the amplitude fading of the frequency nonselective aeronautical radio channel is band-limited to twice the maximum Doppler frequency $f_{D,max}$.

For a measurement system this now implies that the amplitude fading of the channel has to be sampled with at least double the frequency to avoid aliasing, so with a sampling frequency of $f_{ms} \geq 4f_{D,max}$. With the parameters used throughout this paper this means that the amplitude scaling has to sampled with a frequency of $f_{ms} = 96\,\text{Hz}$ or, in terms of a signal sampling rate $f_{sa} = 8000\,\text{Hz}$, every 83 samples. For a measurement system based on maximum length sequences (MLS, see e.g. [8]) this means that the MLS length should be no longer than $L = 2^n - 1 = 63$ samples.

References

[1] D. van Roosbroek, "EATMP communications strategy," Eurocontrol, Technical Description Vol. 2 (Ed. 6.0), 2006.
[2] M. Hagmüller, H. Hering, A. Kröpfl, and G. Kubin, "Speech watermarking for air traffic control," in *Proceedings of the 12th European Signal Processing Conference (EUSIPCO'04)*, Vienna, Austria, September 2004.
[3] K. Hofbauer and H. Hering, "Digital signatures for the analogue radio," in *Proceedings of the 5th NASA Integrated Communications Navigation and Surveillance Conference (ICNS)*, Fairfax, USA, May 2005.
[4] BAE Systems Operations Ltd, "Literature review on terrestrial broadband VHF radio channel models," B-VHF, Report D-15, 2005. [Online]. Available: http:///www.B-VHF.org
[5] M. Pätzold, *Mobile Fading Channels. Modelling, Analysis and Simulation.* John Wiley and Sons Ltd., 2002.
[6] J. D. Parsons, *The Mobile Radio Propagation Channel.* John Wiley and Sons Ltd., 2000.
[7] J. G. Proakis and M. Salehi, *Communication Systems Engineering*, 2nd ed. Prentice Hall, 2001.
[8] B. Sklar, *Digital Communications*, 2nd ed. Prentice-Hall, 2001.
[9] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, 3rd ed. Springer-Verlag, 2004.
[10] E. Haas. Communications systems. [Online]. Available: http://www.kn-s.dlr.de/People/Haas/
[11] ——, "Aeronautical channel modeling," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 2, 2002.
[12] ARINC, "Airborne VHF communications transceiver," Aeronautical Radio Inc., June 2003.
[13] The MathWorks, *MATLAB Communications Toolbox*, 3rd ed., 2004.
[14] K. Metzger. The generic channel simulator. [Online]. Available: http://www.eecs.umich.edu/genchansim/
[15] K. Hofbauer. The generic channel simulator. [Online]. Available: http://www.spsc.tugraz.at/people/hofbauer/gcs/
[16] K. Hofbauer, H. Hering, and G. Kubin, "Speech watermarking for the VHF radio channel," in *Proceedings of the 4th Eurocontrol Innovative Research Workshop*, Dec. 2005.
[17] ——, "A measurement system and measurements database for the aeronautical VHF voice channel," Feb. 2006, submitted for publication.
[18] M. Dickreiter, *Handbuch der Tonstudiotechnik*. KG Saur, 1997, vol. 1.

About the Low Drag of Dolphin Profiles at Supersonic Speeds

Iosif Taposu *

***Abstract*—The paper shows that the wave drag of symmetrical dolphin profiles at supersonic speeds is smaller as compared with that corresponding to any classical airfoil having a rounded leading edge. A simple formula was established for the drag coefficient of dolphin profiles, also it was proved that the bi-directional symmetrical dolphin profiles have the smallest drag. It was performed a comparison with the wave drag of symmetrical classical lenses, also the rhombic airfoils.**

***Index Terms*—Aerodynamics, Drag, New concepts** (innovative ideas)**, Supersonic regime.**

I. NOMENCLATURE

Oxyz – the reference system fixed with the airfoil that lies in the plane Oxz, the origin O being in the leading edge (Fig.1);

2ε - the maximum thickness of airfoil;

x_ε - the position of maximum thickness along the unit chord;

x_1, x_2 - the position of first, respectively the second inflexion point of thickness distribution for dolphin profiles;

C_D - the drag coefficient;

M_∞ - the Mach number ;

g(x) - the one half thickness distribution along the chord;

$g'(x)$ - the first order derivative of function g with the variable x.

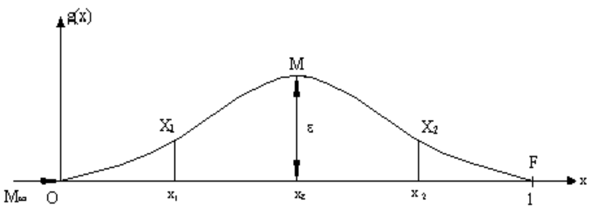


Fig. 1

Manuscript received February 14, 2006.

* Ph. D, Senior Scientist at Aviation Institute INAV-S.A. , Bucharest, Romania. Member GAMM, AIAA. Phone/Fax: +40212326887, E-mail: taposu@inav.ro

II. INTRODUCTION

The dolphin profile family is very rich as compared with the reunion of all classical airfoils having rounded leading edges[1]. The new family generalizes the Joukowski's airfoil concept[2] along two ways: extends the Joukowski's condition at the leading edge, also removes the geometrical restrictions as imposed by the Joukowski's mappings. For a given design point, the ideal dolphin profile (without an imposed technological minimum) is defined using the tangency condition of one half thickness distribution to the skeleton line in the leading and trailing edges, these both being cusped points.

Theoretical and experimental studies proved high performances as compared with the classical airfoils. The most important performance at low speeds as proved in the wind tunnel[3] is a very high critical incidence, over 23 degrees for a thin airfoil and this value is practically double as compared with that corresponding for classical airfoils. Also, water tunnel tests[4] proved for a naval rudder critical deflection angles over 60 degrees, again very high as compared with the reference rudder. Other secondary gains can be put into evidence at low speeds[1]. Numerical simulations proved that along the transonic regime the dolphin profiles have a very low drag, continuous polars as function of Mach numbers, smooth pressure distribution and low aerodynamic noise, also other favorable performances[5].

The actual paper aims to prove that the dolphin profiles have a low wave drag at supersonic speeds as compared with the classical airfoils. To simplify the discussion we have chosen symmetrical airfoils at zero incidence angle only because the thickness distribution is the most important geometrical parameter of which is depending the drag coefficient.

III. THE SYMMETRICAL DOLPHIN PROFILES

According to a simple method for geometrical building the thickness distribution of ideal dolphin profiles along the unit chord is[1]

$$g(x)=\varepsilon.\begin{cases} F_1(x), & 0 \le x \le x_1 \\ F_2(x), & x_1 < x \le x_\varepsilon \\ F_3(x), & x_\varepsilon < x \le x_2 \\ F_4(x), & x_2 < x \le 1 \end{cases} \qquad (1)$$

where

$F_1(x) = \dfrac{x^2}{x_1 x_\varepsilon}$,

$F_2(x) = 1 - \dfrac{(x-x_\varepsilon)^2}{x_\varepsilon(x_\varepsilon - x_1)}$, (2)

$F_3(x) = 1 - \dfrac{(x-x_\varepsilon)^2}{(x_\varepsilon - x_2)(x_\varepsilon - 1)}$,

$F_4(x) = \dfrac{(x-1)^2}{(x_2-1)(x_\varepsilon - 1)}$.

Observe the function g(x) as being derivable along the whole interval [0,1], also its first order and continuous derivative is

$$g'(x)=2\varepsilon.\begin{cases} a_1 x,\ 0\le x\le x_1 \\ a_2(x_\varepsilon - x),\ x_1<x\le x_\varepsilon \\ a_3(x-x_\varepsilon),\ x_\varepsilon < x \le x_2 \\ a_4(x-1),\ x_2<x\le 1 \end{cases} \quad (3)$$

and

$a_1 = \dfrac{1}{x_1 x_\varepsilon}$, $a_2 = \dfrac{1}{x_\varepsilon(x_1 - x_\varepsilon)}$,

$a_3 = \dfrac{1}{(x_2 - x_\varepsilon)(x_\varepsilon - 1)}$ (4)

$a_4 = \dfrac{1}{(x_2-1)(x_\varepsilon - 1)}$.

The wave drag coefficient of any thin symmetrical airfoil at supersonic speeds and zero incidence may be estimated by[3]

$$C_D = \frac{4}{\sqrt{M_\infty^2 - 1}} \int_0^1 g'^2(x)dx, \quad (5)$$

and to compute this integral we need the function $g'^2(x)$ only. From (3) obtain simple $g'^2(x)$, also its primitive $G(x)=\int g'^2(x)dx$ as being

$$G(x) = 4\frac{\varepsilon^2}{3}.\begin{cases} a_1^2 x^3, 0\le x\le x_1 \\ a_2^2(x_\varepsilon - x)^3,\ x_1<x\le x_\varepsilon \\ a_3^2(x-x_\varepsilon)^3,\ x_\varepsilon < x\le x_2 \\ a_4^2(1-x)^3,\ x_2 < x \le 1 \end{cases} \quad (6)$$

Using the primitive (6) then the integral (5) can be computed immediately. If decompose it on the appropriate four intervals where every integral is a positive quantity, then obtain

$$C_D = 16\frac{\varepsilon^2}{3\sqrt{M_\infty^2-1}} *[a_1^2 x_1^3 + a_2^2(x_\varepsilon - x_1)^3 + + a_3^2(x_2 - x_\varepsilon)^3 + a_4^2(1-x_2)^3]. \quad (7)$$

If use (4) in (7) it follows finally the simple formula

$$C_D = \frac{16\varepsilon^2}{3\sqrt{M_\infty^2-1}} \frac{1}{x_\varepsilon(1-x_\varepsilon)}. \quad (8)$$

For given ε and M∞ observe C_D as being minimum if $x_\varepsilon=0.5$, i.e. the bi-directional dolphin profiles have the lowest drag at supersonic speeds. It is useful to remark that the drag is not depending on the position of thickness inflexion points x_1 and x_2 along the chord.

IV. A CLASSICAL LENS

Consider a classical symmetrical airfoil having sharped leading and trailing edges like as a lens.
The one half thickness distribution along the unit chord is the simple parabolic arc (Fig.2)

$$g(x) = x(1-x)tg\frac{\alpha}{2}, \quad (9)$$

where $\alpha=2\mathrm{arctg}\, g'(0) \neq 0$ is the angle between the tangents at contour in the leading edge. The maximum thickness of this airfoil is

$$2\varepsilon = \frac{1}{2}tg\frac{\alpha}{2}, \quad (10)$$

being placed at the middle of chord, i.e. $x_\varepsilon = 0.5$. From (9) obtain the derivative $g'(x)$ as

$$g'(x) = (1-2x)tg\frac{\alpha}{2}. \quad (11)$$

Now using (5) we can compute the drag coefficient and obtain

$$C_D = \frac{4}{\sqrt{M_\infty^2-1}}(tg\frac{\alpha}{2})^2 \int_0^1 (1-2x)^2 dx, \quad (12)$$

i.e. have

$$C_D = \frac{4}{3\sqrt{M_\infty^2-1}}(tg\frac{\alpha}{2})^2 = 64\frac{\varepsilon^2}{3\sqrt{M_\infty^2-1}}. \quad (13)$$

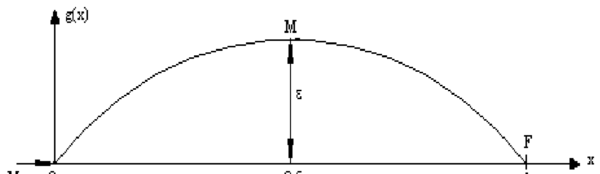


Fig.2

If compare the drag coefficients given in (8) and (13) obtain the ratio

$$k = \frac{1}{4x_\varepsilon(1-x_\varepsilon)}, \quad (14)$$

and for $x_\varepsilon = 0.5$ obtain k = 1. Therefore, if the small ε and M_∞ are given then this biconvex airfoil has the same wave drag as any thin bi-directional symmetrical dolphin profile. For example, if take M_∞ =1.5 and ε=0.03, i.e. the airfoil has the maximum thickness six per cent of the chord, then obtain C_D=0.017. If ε is higher then C_D increases, e.g. for 2ε =0.1 it follows C_D=0.047, the same value being obtained if use (8) for a bi-directional dolphin profile having the maximum thickness 10 per cent of chord.

We have to remark that if apply the same technique for a rhombic airfoil having g(x) like as in Fig.3 obtain the well known formula

$$C_D = 16\frac{\varepsilon^2}{\sqrt{M_\infty^2 - 1}} \quad . \qquad (15)$$

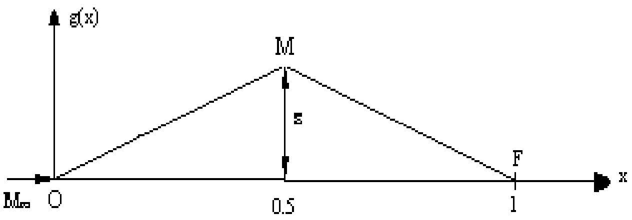


Fig. 3

As compared with the lenses or the conic airfoils it seems that the dolphin profiles are more advantageous because at least in the ideal case the leading and trailing edges may be not stagnation points[1] and so have lower pressure gradients along the whole chord on both the upper and lower sides that means important aerodynamic improvements, particularly lower external noise. The dolphin profiles may be a valid research way to diminish strongly the sonic boom.

V. A JOUKOWSKI'S AIRFOIL

In our analysis when refer to the classical airfoils we have selected a symmetrical Joukowski's airfoil using mainly two reasons: it has a rounded leading edge, also fulfils the condition of being a dolphin profile at the trailing edge. If consider the well known J015 having 2ε =0.15 then the one half thickness distribution in our reference system is[1]

$$g(x) = m(1-x)\sqrt{x(1-x)} \ , \qquad (16)$$

where m = 0.231. From (16) obtain the first order derivative as being

$$g'(x) = \frac{m}{2}\left[\frac{(1-x)\sqrt{1-x}}{\sqrt{x}} - 3\sqrt{x(1-x)}\right], \qquad (17)$$

and

$$g'^2(x) = \frac{m^2}{4}\left(\frac{1}{x} - 9 + 24x - 16x^2\right). \qquad (18)$$

Using (18) in (5) we simple compute C_D and obtain

$$C_D = \frac{m^2}{\sqrt{M_\infty^2 - 1}}\left(\lim_{\delta \to 0} \ln\frac{1}{\delta} - \frac{7}{3}\right). \qquad (19)$$

Also observe as $C_D \to \infty$ when $\delta \to 0$. If consider M_∞ = 1.5 and δ = 10^{-3} obtain C_D = 0.218 that is a big value. Similar conclusions follow for every classical airfoil having a rounded leading edge, particularly those from the NACA family. Therefore, all these airfoil are not appropriate at supersonic speeds, the dolphin profile concept remaining a valid research way to improve the aerodynamic performances for this flight regime.

VI. CONCLUSIONS

Sometimes the old classical results are opportune when we intend to test quickly new concepts in the science. In this case using a simple classical formula for the wave drag of thin symmetrical airfoils at supersonic speeds the paper proved that the very rich family of dolphin profiles may be an appropriate research way to improve the aerodynamic performances at this flight regime. As compared with two classical airfoils having sharp leading/trailing edges, but not dolphin profiles, the drag coefficients have neighbouring values. However at least an important advantage hangs up to the dolphin profiles, namely the riches of this family that allows numerous optimization studies according to the known specific criteria. This initial study must be continued using numerical simulations in real viscous fluid and finally the decisive wind tunnel tests for validation of the theoretical results. All airfoils having rounded leading edges cannot be opportune at supersonic speeds because they have a very high drag, also other negative associated effects like as the strong sonic boom.

VII. REFERENCES

[1] I. Taposu, " Dolphin Profiles, a New Concept in Aerodynamics (in Romanian)". Ed. Tehnica, 243 pages, Bucharest, 2002.

[2] C. Iacob, "Introduction mathematique a la mecanique des fluids". Ed. Academiei /Gauthier Vilars, Bucharest/Paris , 1959.

[3] I. Taposu; P. Spataru, "About the experimental results of a dolphin profile at low speeds". AIAA applied aerodynamics conference, TP2000-4506, Denver, Colorado, August 2000.

[4] D. Obreja, Gh. Iorga, "Experimental results in the water tunnel for a naval rudder (in Romanian)". Report Ac4469, ICEPRONAV Galati, 2001.

[5] O. Manole,."Numerical simulations for dolphin profiles at transonic speeds (in Romanian)". Raport AC-2976, INCAS Bucuresti, 1998.

[6] H. Schlichting, E. Truckenbrodt, "Aerodynamik des Flugzeuges", Bd. I, Springer Verlag, Berlin, 1959.

Assessing Image Difficulty in X-Ray Screening Using Image Processing Algorithms

Anton Bolfing, Stefan Michel, Adrian Schwaninger

***Abstract*—The relevance of aviation security has increased dramatically in the last years. One of the most important tasks is the visual inspection of passenger bags using x-ray machines. In this study we investigated the role of the three image-based factors view difficulty, superposition and bag complexity on human detection of familiar prohibited items (knives) in x-ray images. In Experiment 1 we replicated earlier findings in order to provide converging evidence for the validity of these factors. In Experiment 2 we assessed the subjective perception of the same image-based factors. Twelve participants rated the x-ray images used in Experiment 1. Threat images were rated for view difficulty, superposition, clutter, transparency and general difficulty. Except for clutter ratings obtained in Experiment 2 were significantly correlated with detection performance in Experiment 1. We then developed statistical and image-processing algorithms to calculate the image-based factors automatically from x-ray images. In Experiment 3 it was revealed that most of our computer-generated estimates were well correlated with human ratings of image-based effects obtained in Experiment 2. This shows that our computer-based estimates of view difficulty, superposition, clutter and transparency are perceptually plausible. Using multiple regression analysis we could show in Experiment 4 that our computer estimates were able to predict human performance in Experiment 1 as well as the human ratings obtained in Experiment 2. Applications of such a computational model are discussed for threat image projection systems and for adaptive computer-based training.**

***Index Terms*—Airport security, image difficulty estimation, image processing, statistical model, x-ray screening.**

I. INTRODUCTION

The aim of this study is to evaluate a computational model for image difficulty assessment. Such a model could be important for applications such as Threat Image Projection (TIP) or computer based training. TIP is a software function available on modern x-ray equipment that allows inserting fictional threat items (FTIs) into x-ray images of real passenger bags. TIP is a source of motivation to screeners, provides a means of improving screeners' threat knowledge, and can be used to assess screeners' threat detection performance. Schwaninger, Hardmeier and Hofer (2004) have identified three major image-based factors influencing detection performance: View difficulty of the FTI depending on its rotation, superposition by other objects in the bag, and bag complexity. The latter comprises clutter, the texture's unsteadiness, and transparency, the relative size of dark areas. Current TIP systems project FTIs into real passenger bags based on a random ratio into a random position of the bag. As a consequence, TIP images vary substantially in image difficulty depending on effects of view difficulty, superposition and bag complexity. When TIP data is used to assess screener performance such effects make it difficult to obtain reliable measurements. The main aim of the current work is to develop a computational model using image processing in order to determine x-ray image difficulty while taking effects of view difficulty, superposition, and bag complexity into account.

The study comprises four experiments. The first experiment is a replication of earlier findings to confirm the relevance of three image-based factors in predicting human detection performance. In the second experiment we estimated the relevance of these image-based factors by correlating subjective ratings of them with the hit rate (p(hit)) in detection performance. We expect high correlations to reflect high influence of the subjectively perceived image-based factors on the measured item difficulty p(hit). In Experiment 3 we correlated the computer-based estimates with human ratings to estimate the perceptual plausibility of our computer algorithms for image-based factors estimation. Additionally, this allows us to check for possible intercorrelations among the predictors, which allows us to detect statistical dependencies among them. Finally, in experiment 4 we used multiple linear regression analyses to compare our computational model and human perception with regard to how well they can predict human detection performance.

II. EXPERIMENT I: ORT FINDINGS REPLICATION

A. *Method and Procedure*

1) *Participants*

The sample size of participants was twelve undergraduate students in psychology (5 females). None of the participants has had experience with x-ray images before.

Manuscript received February 24, 2006. This work was supported in part by the Zurich State Police, Airport Division.

A. Bolfing is with Department of Psychology, University of Zurich, Switzerland (e-mail: a.bolfing@psychologie.unizh.ch).

S. Michel is with the Department of Psychology, University of Zurich, Switzerland (e-mail: s.michel@psychologie.unizh.ch).

A. Schwaninger is with the Department of Psychology, University of Zurich, Switzerland (e-mail: a.schwaninger@psychologie.unizh.ch).

2) ORT Test Design

Stimuli were displayed on 17'' TFT screens at a distance of about 100 cm so that x-ray images subtended approximately 10-12 deg of visual angle. The computer program measured outcome (hit, miss, false alarm (FA), correct rejection (CR)) and the time from image onset to final decision key press.

In this study we used the X-Ray Object Recognition Test [4], which contains 256 x-ray images, half of them with an FTI (threat images), the other half without an FTI, i.e. non-threat images. Viewpoint difficulty, superposition and bag complexity are counterbalanced using the following design: 16 (threat items, i.e. 8 guns and 8 knives) x 2 (easy vs. difficult viewpoint) x 2 (easy vs. difficult superposition) x 2 (easy vs. difficult bag complexity) x 2 (threat vs. non-threat images).

The construction of the items in all image-based factor combinations as explained above was lead by visual plausibility criteria. After choosing two sets of x-ray images of harmless bags differing in bag complexity, the sixteen FTI's were projected into the bags in two different view difficulties at two locations with different superpositions each (for details see [4]).

3) Procedure

The X-Ray ORT is fully computer-based. Before starting the test, several practice trials are presented to make sure that the task is understood properly. Immediately prior to the actual test, all threat objects are presented on the screen to reduce any effects of visual knowledge about the shape of threat objects [5]. The participant's task is to decide whether a bag is OK (no threat item present) or NOT OK (threat item present). Each x-ray image disappears after 4 seconds. In addition, participants have to judge the confidence in their answer using a slider control (from "not sure at all" to "very sure"). No feedback is given to their answers. In this study, only trials containing knives have been used for analysis because of their high familiarity. This is important because we want to measure image-based factors related to visual abilities and not the visual knowledge of screeners about threat objects (for details see [5]). In 2005 the same study has been carried out with guns [7], where similar results have been obtained. A study comparing and discussing the differences is currently being conducted.

4) Statistical Analysis

A three-way analysis of variance (ANOVA) with view difficulty, superposition, and bag complexity as within-participant factors was used on percentage of detected threats (hit rate) per participant and factor combination.

B. Results

The main effects of view difficulty, superposition, and bag complexity are shown in Figure 1.

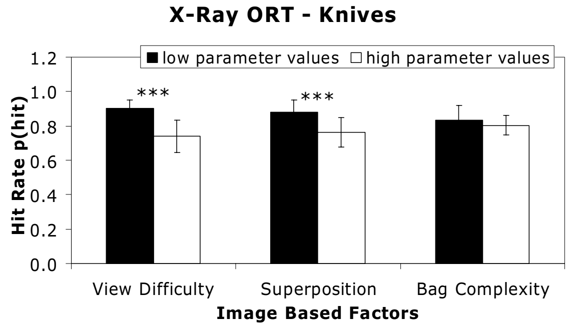


Fig. 1 Main effects of the image-based factors view difficulty, superposition, and bag complexity on hit rate in the X-Ray ORT.

The following results were obtained in the ANOVA. There were clear main effects of view difficulty, η^2=.84, $F(1,11)$=59.06, p<.001, and superposition: η^2=.65, $F(1,11)$=20.48, p<.001. The effect of bag complexity was only marginally significant, η^2=.23, $F(1,11)$=3.30, p=.10. Interaction effects: View difficulty * superposition, η^2=.19, $F(1,11)$=2.54, p=.14, view difficulty * bag complexity, η^2=.52, $F(1,11)$=11.93, p<.01, superposition * bag complexity, η^2=.01, $F(1,11)$=0.06, p= .82, and view difficulty * superposition * bag complexity, η^2=.34, $F(1,11)$=5.60, p<.05.

C. Discussion

We found clear effects of view difficulty and superposition in hit rates, which replicates the results of an earlier study using the X-Ray ORT with novices and screeners by calculating A' scores [4, 5]. Thus, the results of Experiment 1 provide further evidence for the validity of image-based factors as important determinants of x-ray image difficulty. However, the effect of bag complexity was only marginally significant whereas in earlier studies large main effects of bag complexity were found [4, 5]. This is due to the fact that in this study the dependent variable was the hit rate in contrast to our earlier studies in which A' was used. A' is a "non-parametric" measure of detection performance or sensitivity that takes the hit rate and the false alarm rate into account (for details on this and other detection measures in x-ray screening see [6]). Note that bag complexity is the only factor, which is solely dependent on the bags themselves, see section IV.B.1). Therefore, the false alarm rate is the sensitive measure for bag complexity which explains why large effects of bag complexity were found in earlier studies in which hit and false alarm rates were used to calculate A' scores [4, 5].

III. Experiment II: Image-based Factors Rating

A. Method and Procedure

1) Participants

The participants of Experiment 1 took part one week later in Experiment 2.

2) Rating

The same experimental setup was used as in Experiment 1. The participant's task in Experiment 2 was to rate the X-Ray

ORT images in terms of general difficulty and the following image-based factors: View difficulty, superposition, and bag complexity (clutter and transparency). Each of these dimensions could be rated on a 50 point scale from “very low” to “very high” using a slider control.

3) Statistical Analysis

In order to validate the influence of the subjectively perceived image-based factors on the hit rate we correlated the image-based factors ratings with p(hit) derived from Experiment 1 using Pearson’s product-moment correlation.

B. Results

The correlation between objective hit rate and subjectively rated view difficulty was $r(64)=-.53$, $p<.001$, between p(hit) and superposition $r(64)=-.67$, $p<.001$, between p(hit) and clutter $r(64)=-.24$, $p=.06$, and between p(hit) and transparency $r(64)=.31$, $p<.05$.

C. Discussion

The results show clearly that there is a covariation between objective hit rates and the subjective perception of our image-based factors view difficulty, superposition, and bag complexity. This indicates a high perceptual plausibility of these image based factors. However, bag complexity, comprising clutter and transparency, shows the lowest correlation, whereas clutter does not correlate significantly with the hit rate. Again, these findings are consistent with the above mentioned observation that bag complexity, which refers to the bag content only, should be related more to false alarm rates than to hit rates.

Another explanation could be that bag complexity is difficult to rate because of low perceptual plausibility. Indeed, the two components of bag complexity (clutter and transparency) were highly correlated $r(64)=-.89$, $p<.001$. This could imply that novices have a hard time in distinguishing the two components of bag complexity and just give similar ratings for both.

IV. Experiment III: Image-based Factors Correlations between Computer Estimates and Ratings

A. Introduction

Experiment 3 was designed to develop image-processing algorithms for estimating view difficulty, superposition, and bag complexity automatically in x-ray images. These algorithms were then validated by correlating the computer-based estimates with the human ratings from Experiment 2.

B. Method and Procedure

1) Image Processing Algorithms

All image-processing algorithms developed for this purpose are based on theoretical considerations. For each image-based factor the consequences of high and low parameter values of each single image-based factor on the pixel and frequency space have been determined. Different algorithm parameters were optimized by maximizing the correlations between the image-based factor estimates and human detection performance of earlier studies [4, 5, 7].

In the following subsections the image-processing algorithms are described in turn.

a) View Difficulty

Because it is not possible to determine the degree of 3-D rotation (view difficulty) of a physical threat item solely from the 2-D x-ray image, this image-based factor is not being calculated using computational recognition algorithms, but statistically from X-Ray ORT detection performance data.

$$ViewDifficulty \quad VD_j = \left(\left(\sum_{i=1}^{n} p(hit)_i \right) - p(hit)_j \right) \Big/ n-1$$

Eq. 1 shows the equation for estimating view difficulty, whereas *j* denotes the index of the x-ray image in question and n is the number of bags each FTI has been projected to.

The image-based factor view difficulty is basically calculated by averaging the hit rates of a certain threat item in one of the two views presented in the ORT. In order to avoid a circular argument in the statistical model (multiple linear regression, section V.B.1) by partial inclusion of a predictor into the criterion variable, the detection performance of the one item in question is being excluded from this average detection performance estimate.

In this study, each threat item (knives only) is being displayed four times in the same view, 2 (bag complexity low vs. high) x 2 (superposition low vs. high). Therefore, the n in the view difficulty formula equals 4, but the average is calculated over the remaining three items.

b) Superposition

The image-processing algorithm for superposition simply calculates the Euclidian distance between the grayscale pixel intensities of the signal-plus-noise (SN) image and the harmless bag (N, noise) image.

$$Superposition \quad SP = \sqrt{\sum (I_{SN}(x,y) - I_N(x,y))^2}$$

Eq. 2 Image-processing formula of the image-based factor superposition, whereas $I_{SN}(x,y)$ denotes the pixel intensities of a threat image and $I_N(x,y)$ denotes the pixel intensities of the corresponding harmless bag.

For each pixel of a threat image, the pixel intensity difference between the bag with the threat item and the bag without the threat item is calculated and squared. All squared pixel intensity differences are summed up. The final image-based factor is then derived from calculating the square root of this sum of squared pixel intensity differences.

c) Clutter

This image-based factor is designed to express bag item properties like their texture unsteadiness, disarrangement, chaos or just clutter. In terms of the depicted bags themselves, this factor is closely related to the amount of items in the bag as well as their structures in terms of complexity and fineness. The method used in this study is based on the assumption, that such texture unsteadiness can be described mathematically in terms of the amount of edges, i.e. the amount of transitions in luminosity within a certain space frequency range surpassing a certain threshold.

$$Clutter \quad CL_C = \sum_{y}^{height} \sum_{x}^{width} (I_{hp})$$

$$where \quad I_{hp} = I_N \otimes F^{-1}(HP(f_x, f_y))$$

Eq. 3 Image-processing formula of the image-based factor clutter, whereas I_N denotes the pixel intensities of the harmless bag image, F^{-1} denotes the inverse Fourier transformation and $HP(f_x,f_y)$ represents a highpass-filter in Fourier space.

We implemented this mathematical formulation by first applying on the intensity image of the empty bag a convolution kernel, which is derived from a highpass-filter in the Fourier space by inverse Fourier transformation (see Appendix). In a second step, the amount of the resulting pixels, representing edges as described above are being counted.

d) Transparency

The image-based factor transparency reflects the amount to which x-rays are able to penetrate various objects in the bag. This depends on the specific material density of these objects. Heavy metallic materials such as lead are known to be very hard to be penetrated by x-rays. For a screener, the consequence is that he cannot see any objects in front or on the back of such material.

The implementation of the image-processing algorithm for

$$Transparency \quad TR = \frac{\sum_{x,y} (I_N(x,y) < thresh)}{\sum_{x,y} (I_N(x,y) \neq 255)}$$

Eq. 4 Image-processing formula of the image-based factor transparency, whereas $I_N(x,y)$ denotes the pixel intensities of the harmless bag and *thresh* is the pixel intensity threshold beneath which the pixels are counted.

the image-based factor transparency consists in the calculation of the amount of pixels being darker than a certain threshold (<thresh) of the pixel intensity range going from 0 to 255, relative to the bags overall size (< 255, white pixels).

2) Correlations

To evaluate the perceptual plausibility of these image-processing algorithms for estimating view difficulty, superposition, and bag complexity we correlated them with the human ratings obtained in Experiment 2.

C. Results

The correlations were $r(64)=-.47$, $p<.001$ for view difficulty, $r(64)=-.44$, $p<.001$ for superposition, $r(64)=.18$, $p=.16$ for clutter and $r(64)=-.63$, $p<.001$ for transparency.

D. Discussion

Except for clutter all correlations between calculations and ratings are highly significant. Remember the high intercorrelation between the human ratings of the image-based factors clutter and transparency ($r(64)=-.89$, $p<.001$) obtained in Experiment 2. Here, in Experiment 3 we found a quite high intercorrelation between the corresponding calculated image-based factors clutter and transparency ($r(64)=-.55$, $p<.001$). Therefore, we must keep in mind that these two factors are not fully independent. This is not very surprising as we subsume them within the factor bag complexity. Nevertheless we can conclude that our calculations are compatible with the perceptual plausibility of human observers.

V. EXPERIMENT IV: STATISTICAL MODEL: USING MULTIPLE LINEAR REGRESSION ANALYSIS

A. Introduction

Experiment 4 was designed to evaluate the predictive power of our computational model and to compare it to human perception as a tough baseline.

B. Method and Procedure

One multiple regression analysis was carried out using the computationally estimated image-based factors as predictors. A second multiple regression analysis was conducted using the subjectively rated image-based factors as predictors.

1) Multiple Regression Analysis

Weight estimation for a linear model can be achieved using multiple linear regression analysis whereas our image-based factors are the predictors, and the hit rate of human observers obtained in Experiment 1 is the dependent variable. Table 1 shows the abbreviations used later in this section for all dependent and independent variables.

Dependent Variable	statistical Measure	Rating
Item Difficulty	DP: Hit Rate Detection Performance	GD General Difficulty

Independent Variables	Calculations	Ratings
View Difficulty	VD_c	VD_r
Superposition	SP_c	SP_r
Clutter	CL_c	CL_r
Transparency	TR_c	TR_r

In the following, we describe the computational model and the model based on human ratings of perceived view difficulty, superposition and bag complexity.

Linear model using computationally calculated image-based factors as predictors:

$$DP = b_0 + b_1 VD_c + b_2 SP_c + b_3 CL_c + b_4 TR_c + R$$

Linear model using mean values of the subjectively rated image-based factors as predictors:

$$DP = b_0 + b_1VD_r + b_2SP_r + b_3CL_r + b_4TR_r + R$$

Criteria for comparing the different statistical models are:

1. Goodness-of-fit measures
2. Regression coefficient's significances and
3. Percentage of variance in the dependent variable (hit rate) the model is able to explain by its predictors.

C. Results

1) Computational Model

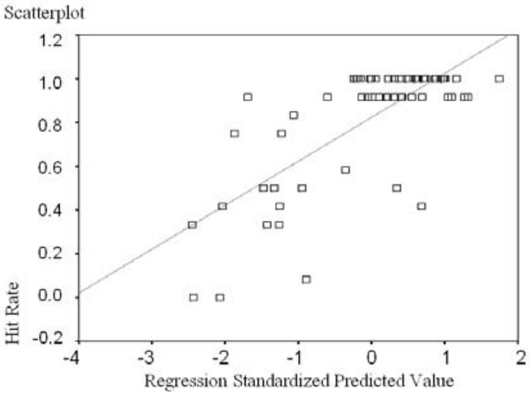


Fig. 2 Scatterplot of the multiple linear regression analysis using the calculated image-based factors as predictors with the standardized predicted values on the abscissa and the measured human performance (hit rates) on the ordinate.

In the table below the basic statistical criteria of the multiple linear regression using calculated image-based factors are listed:

Variable	B	SE B	β
VD_c	.352	.144	.288*
SP_c	.056	.013	.497***
CL_c	.000	.000	-.149
TR_c	-.172	.637	-.029

R^2=.543, $R^2(adj)$=.512, $F(4,59)$=17.49, p<0.001
* p<.05, ** p<.01, ***p<.001

2) Ratings Model

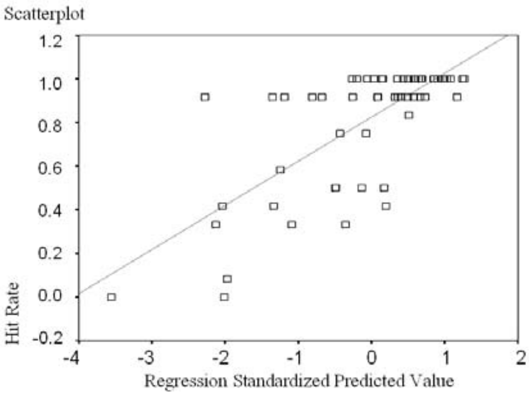


Fig. 3 Scatterplot of the multiple linear regression analysis using the subjectively rated image-based factors as predictors with the standardized predicted values on the abscissa and the measured ORT hit rates on the ordinate.

In the table below the basic statistical criteria of the multiple linear regression using subjectively rated image-based factors are listed:

Variable	B	SE B	β
VD_r	-.010	.003	-.329**
SP_r	-.021	.004	-.675**
CL_r	-.005	.007	-.131
TR_r	-.013	.008	-.338

R^2=.548, $R^2(adj)$=.518, $F(4,59)$=17.91, p<0.001
* p<.05, ** p<.01.

D. Discussion

In Experiment 4 we have developed a computational model to calculate view difficulty, superposition, and bag complexity automatically in x-ray images using image processing algorithms. These algorithms were significantly correlated with human ratings of view difficulty, superposition, and bag complexity. This result shows that our image processing algorithms are perceptually plausible. In order to benchmark our model we compared its predictive power with a model based on human perception. Our computational model using image processing algorithms for automatic estimation of view difficulty, superposition, and bag complexity could predict human detection performance in the X-Ray ORT with a correlation between model prediction and measured performance of r = .74. In order to benchmark our model we have compared it to a linear model using human ratings of view difficulty, superposition, and bag complexity. As we can see from the R^2 values of the regression analyses our computational model performed equally well as a linear model using human ratings of perceived view difficulty, superposition, and bag complexity. It is important to focus on the strength of the impacts of the single image-based factors used as predictors, the beta-weights. As we expected based on Experiment 2 and Experiment 3, the image-based factors solely depending on the empty bags contribute little to the prediction of the hit rate. In both, the computational model as well as the ratings model, only view difficulty and superposition result in significant beta-weights. This problem was discussed already in paragraph IV.D with respect to the different measures hit rate and A'. One is tempted to conclude that bag complexity with its subfactors clutter and transparency could be excluded when hit rates have to be predicted. However, it should be pointed out that in this study only knives were used and different results might be obtained for other types of prohibited items. Moreover, using another formula for estimating bag complexity it might be possible to achieve better results. In any case, there is evidence that image-based factors related to bag complexity are important to predict false alarm rates. This is even more important when TIP is activated because a high false alarm rate of a screener results in long waiting lines.

VI. General Discussion

The results show clearly, that it is possible to develop an automatic system that calculates x-ray image difficulty using image processing algorithms. Further research needs to done regarding the division of image-based factors predicting hit and false alarm rates. Furthermore, there are great chances to enhance the existing predictors and possibly add further predictors of x-ray image difficulty. Especially in the field of bag complexity measures, probably the most challenging ones, there is some more work to be done in the future. As mentioned above, this study was conducted with knives as threat items only. Other object categories like guns, improvised explosive devices (IEDs) and other prohibited items are expected to result in different priorities of view difficulty, superposition, and bag complexity. Such studies can be highly valuable to find out how different properties of threats in terms of their shape and material can influence human detection performance and visual search. As mentioned in section II.A.3), an earlier study was carried out regarding guns as threat items [7] and further studies are planned to investigate IEDs and other threat objects in the same way. Once these data are available we will be able to discuss differences in the effects of the image-based factors in terms of different categories.

APPENDIX

Clutter formula high-pass filter:

$$HP(f_x, f_y) = 1 - \frac{1}{1 + \left(\frac{\sqrt{(f_x^2 + f_y^2)}}{f} \right)^d}$$

This formula denotes the high-pass filter as part of the clutter formula in Experiment 3, whereas f_x and f_y are its frequency components, f is its cut-off frequency and where d is its fall-off. This high-pass filter represents a 2-D matrix in the Fourier frequency-space. Therefore an inverse Fourier transform is applied to transform it into a convolution kernel in the spatial domain.

ACKNOWLEDGMENT

We are thankful to Zurich State Police, Airport Division for their help in creating the stimuli and the good collaboration for conducting the study.

REFERENCES

[1] Schwaninger, A., & Hofer, F., “Evaluation of CBT for increasing threat detection performance in X-ray screening,” in *The Internet Society 2004, Advances in Learning, Commerce and Security*, K. Morgan and M. J. Spector, Eds., Wessex: WIT Press, 2004, pp. 147-156.

[2] Graf, M., Schwaninger, A., Wallraven, C., & Bülthoff, H.H., “Psychophysical results from experiments on recognition & categorization,” Information Society Technologies (IST) programme, Cognitive Vision Systems – CogVis; IST-2000-29375, 2002.

[3] Schwaninger, A., “Object recognition and signal detection,” in Praxisfelder der Wahrnehmungspsychologie, B. Kersten and M.T. Groner, Eds., Bern: Huber, in press.

[4] Hardmeier, D., Hofer, F., & Schwaninger, A. (2005). The object recognition test (ORT) – a reliable tool for measuring visual abilities needed in x-ray screening. *IEEE ICCST Proceedings, 39,* 189-192.

[5] Schwaninger, A., Hardmeier, D., & Hofer, F. (2005). Aviation security screeners visual abilities & visual knowledge measurement. IEEE Aerospace and Electronic Systems, 20(6), 29-35.

[6] Hofer, F. & Schwaninger, A. (2004). Reliable and valid measures of threat detection performance in X-ray screening. *IEEE ICCST Proceedings, 38,* 303-308.

[7] Schwaninger, A., Michel, S., & Bolfing A. (2005). Towards a model for estimating image difficulty in x-ray screening. *IEEE ICCST Proceedings, 39,* 185-188.

The Tromped, an in-flight exercise device to prevent Flight Related Deep Vein Thrombosis

CAROLYN COLLINS, DAVID M. KENNEDY, STEPHEN JERRAMS AND THOMAS CORRIGAN.

***Abstract*—Flight Related Deep Vein Thrombosis (FRDVT) has become a health concern of the general flying public in the last few years, as more and more people take to traveling long haul flights. To dates, airlines have not provided a form of prophylaxis to passengers for the prevention of FRDVT. This paper discusses the reasons why thrombi manifest during flight and the development of a prototype, the Tromped, for use on board commercial aircraft to prevent venous stasis and the possible onset of FRDVT.**

***Index Terms*—Air Plethysmography, Deep Vein Thrombosis, Flight Related Deep vein thrombosis, The Tromped.**

I. INTRODUCTION

DEEP vein thrombosis (DVT), occurs in situations where people are seated for long periods of time during travel on land or in air. Sitting in one position can reduce the venous blood flow velocity by up to two thirds [1].

A thrombus (Figure 1.) grows by layers of platelets, fibrin and red blood cells (RBC) gathering together in the direction of the blood flow [2]. Thrombi form as a result of inherited and/or acquired risk factors. Inherited risk factors are deficiencies in natural inhibitors of coagulation or gene mutations. Acquired risk factors include changes in coagulation factors and fibrinolytic functions. These changes may be as a direct result of a changing health status or directly due to the environment that a person may be exposed to, such as hospitalisation or flying long haul flights [3], [4].

The main determinants thought to influence lower limb thrombi formation, otherwise known as Virchow's Triad include three elements. The first one being changes in blood flow or onset of venous stasis due to cramped seating on long haul flights. Secondly, changes in the vessel wall may be caused by trauma or injury of the popliteal vein as it is compressed against the seat cushion and finally alterations in the blood constituents due to a reduction in the partial pressure of oxygen in the aircraft cabin.

Manuscript received February 2006. This work was supported in part by the DIT. Corresponding author, C. Collins, BSc, is a postgraduate student with the Dublin Institute of Technology, Department of Transport Engineering, Bolton St., Dublin 1. Ireland. (Phone: 00353-1-4023950; fax: 00353-1-4023991; e-mail: Carolyn.collins@dit.ie .

Dr. David Kennedy, BSc, MAI, PhD, Ceng, FIEI, Miee, is the Head of the Department of Mechanical Engineering, Dublin Institute of Technology, Bolton Street, Dublin, Ireland (e-mail: david.kennedy@dit.ie).

Dr. Stephen Jerams, Cert Ed, MSc, PhD, Dublin Institute of Technology, Pembroke St, Dublin, Ireland (email: Stephen.jerams@dit.ie).

Thomas Corrigan, MSc, MRAeS, is the Head of the Transport Engineering Department, Faculty of Engineering, Dublin Institute of Technology, Bolton St. Dublin Ireland. (email: Thomas.corrigan@dit.ie)

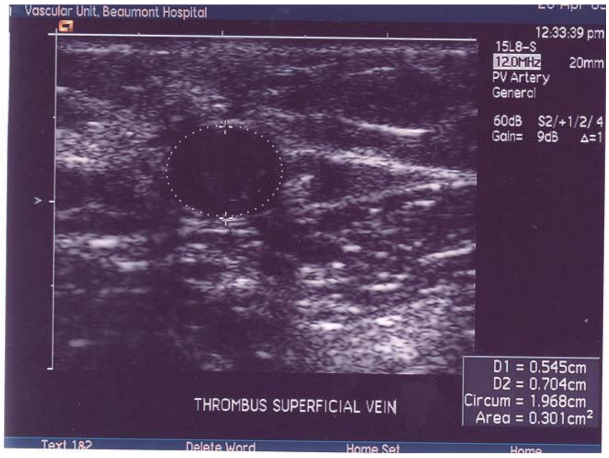


Figure 1. Ultrasound scan of a Thrombus in a superficial vein.

DVT may be symptomatic or asymptomatic (presenting without any signs or symptoms). This is a major area of concern if the thrombus becomes dislodged from its site of formation (it is called an embolus) and moves through the venous system until it becomes lodged again. Areas, which have been known to house these emboli include the brain, the heart or the second, third or fourth order of a pulmonary vessel, hence the term Pulmonary Embolus (PE). If the embolus is large enough to occlude the vessel, it can kill. *"95% of pulmonary emboli arise from thrombi in the deep venous system of the lower extremities"* p.319 [5]. Symptoms of PE may include chest pain, dyspnea or hemoptysis. [6]. *PE "remains a common cause of mortality and its diagnosis is missed in up to 71% of instances"* p. 457 [7].

Venous thromboembolism (VTE) is the term used to describe DVT and or PE. The VTE may remain asymptomatic from anything between 96 hours after the beginning of the flight up to as long as 5 weeks later [8]. *"two thirds of travel-related VTE presents after patients leave the airport"* p. 260 [9]. Many people may be dying directly from flight related DVT and yet this reason may never be attributed to their cause of death.

The Tromped is being developed and medically tested to determine its efficiency in preventing stasis and thus the onset of DVT so that it may be integrated into commercial aircraft design to prevent the occurrence of FRDVT.

II. The Calf Muscle Pump Function

The process of blood flow in the leg, against gravity, towards the heart seems to be instigated by compression of the plantar plexus of the veins, which are situated between the deep and superficial intrinsic muscles of the foot [10]. With contraction of the calf muscles, an ejection volume of venous flow towards the heart occurs, completing the process. The component parts of the muscle pump are the skeletal musculature, intramuscular venous sinusoids and superficial and deep vein compartments [11]. The superficial compartment is made up of the long saphenous vein (LSV), short saphenous vein (SSV) and tributaries of the two veins. The Deep compartment consists of the anterior tibeal, (AT), peroneal tibial (PT) peroneal veins, gastrocnemius and soleal veins [12].

When a person is walking the foot is in contact with the ground 60 per cent of the time and remains off the ground 40% of the time. At the mid segment of the foot is a transverse arch. Therefore the ball of the toes, the heel and the plantar surface of the foot bear the majority of the weight of the body.

During the walking process the three venous pumps, which are the foot pump, the proximal pump and the distal calf pump operate in the following manner. Before the foot bears any weight (i.e. placed on the ground), the ankle is dorsiflexed, emptying the distal calf pump. Weight bearing empties the foot pump into the deep veins of the calf which are relaxed, until required to facilitate heel strike off. The flexing of the plantar area of the foot empties the proximal calf pump into the popliteal (knee area) and femoral veins [10], [13]. All of these movements are recreated during use of the Tromped in situ, thus mimicking the walking process.

Venous flow occurs due to gravity, venous pumping and arterial flow. The peripheral pumping mechanism (calf muscles) alone will cause full venous flow against gravity, when the veins are compressed by the gastrocnemius and the soleus muscles, [14] as seen in Fig. 2.

Exercise results in muscle blood flow increasing to “30±14 ml/ dl/ min reaching 70 ml/ dl/ min”, compared to when the lower limbs are not conducting any exercise, blood flow in the leg averages approximately 300 to 400 ml/min. *“Moderate exercise normally increases total leg blood flow from 5 to 10 times”* p.81 [10]. In contrast stasis will occur due to immobility, during travel by air, as the blood pools in the veins, which become dilated due to prolonged rest. It has been reported that passengers travelling on flights of four hours or more have an increased risk of developing FRDVT, due to immobility “*by between three and five times*” p.19 [15].

So, as activated calf muscles alone are capable of producing venous return, the concept of the Tromped for use during flight was born.

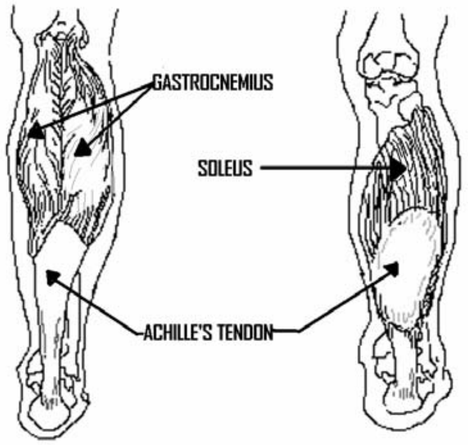


Figure 2. Gastrocnemius and Soleus Muscles

III. Flight Related DVT

“The true incidence of in-flight medical emergencies is unknown as many airlines either do not keep records or do not make them public” p. 1020 [16]. The exact number of medical emergencies on board commercial aircraft is not well known because airlines are not required to report them, even if it requires diversion of the flight [17].

Passengers travelling by air experience unnatural conditions compared to the majority of every day life. They are expected to remain seated in cramped conditions possibly for the duration of the flight. The normal seat pitch on most commercial aircraft is approximately 28 - 31 inches [18]. This in itself is not sufficient leg room for a medium to tall person. Moreover, as most passengers use space below the seat for storage, leg room is subsequently reduced. Therefore due to the fact that there is little or no space to move one’s legs, very little muscle pump action of the lower limbs occurs. Thus, resulting in inadequate return of blood to the heart against gravity and stasis of blood in the lower limbs which could induce hypertension and swelling of the ankles and or calf.

Jet lag may also be a contributing factor to the occurrence of FRDVT. Some passengers suffering from jet lag, upon their arrival at their destination, may try to sleep of its effects and thus become immobile once again after experiencing the cramped seating conditions experienced during long haul flights. These two periods of immobility may encourage a thrombus to grow in size or aid in the onset of a PE.

Passengers who are inclined to sleep through a long haul flight to reduce its length should be persuaded that this could be a dangerous course of action, as it limits movement of the muscles to practically zero, enhancing the possibility of thrombus development due to venous stasis. This stance also should prohibit or encourage passengers to avoid alcohol who may consume it only to aid sleep for the duration or for the most part of the duration of the flight.

The (NZATT) study was designed so as to investigate the frequency and role of risk factors for venous thromboembolism related to travel by air. 878 volunteers completed the study and each person traveled at least 10 hours by air. Ultrasound and CT Pulmonary angiography post travel detected four cases of PE, five cases of DVT and three superficial vein thrombi. *"Our findings suggest that venous thromboembolism is a potentially important health problem to many long-distance air travelers, including those without recognised risk factors"* p.2042 [19]. The authors of this study also recommend further research on prophylactic measures be carried out.

Recently a study into FRDVT confirmed that thrombi formation is as a result of both immobility and the reduced barometric pressure passengers experience on board aircraft in flight [20].

The British Aviation Health Institute built up a database of 544 people, by May 2004, who died or experienced FRDVT. Approximately one third of the recorded victims were under 40 years old, with 74 per cent of mortalities under the age of 60. A total of 84 fatalities occurred, 63% of which were female [21].

PE may develop in passengers several weeks post travel. This would be due to the fact that some DVT may be asymptomatic showing no symptoms. In some cases symptoms of FRDVT may not appear for as long as four weeks after the flight [22].

IV. THE TROMPED PROTOTYPE VERSION I

The first Tromped test jig prototypes, Fig. 3, were produced using aluminium sheets of strength 2024 T4, using a guillotine, a box pan folder and aircraft skin pins. All four test jigs produced had the same dimensions. The base of the test jigs were approximately 177mm wide and 375mm long. The depth of the test jig base is 25mm. The pedals which were attached to the base at one end by a hinge and the front end by the compression spring were 310 mm long, 106mm wide and 18mm deep.

Four springs of different tension, selected randomly were inserted into each of the test jigs by using a stencil with an angle of 25º.

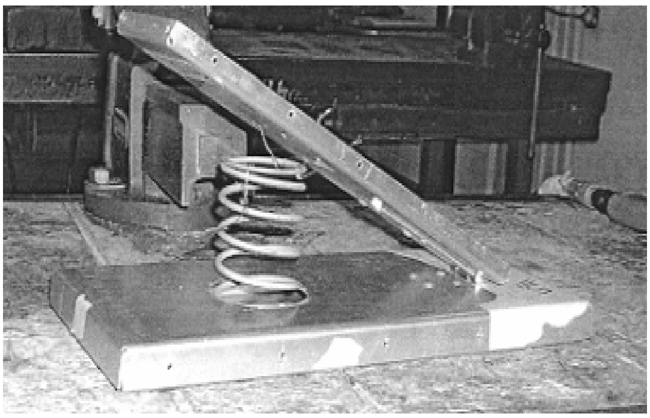


Figure 3. Tromped Prototype version I.

V. TROMPED MEDICAL ASSESSMENT

A medical experiment using an Air Plethysmograph (APG) machine and ten volunteers was conducted in the Non-Invasive Vascular Lab, Beaumont Hospital, Dublin. The aim of this experiment was to determine if any of the springs when compressed, could produce both a sufficient ejection volume fraction (EVF) and a residual volume fraction (RVF) of blood, The normal ejection volume fraction is greater than or equal to 60% [23], [24]. The residual volume (RV) is a measurement of the amount of blood remaining in the veins after exercise. The normal RVF is less than or equal to 35% and is an ideal indicator for the prevention of venous stasis. The APG consists of a Pressure Transducer, amplifier and a pen recorder all connected to a calibrated air chamber which is placed around the calf of the volunteer's right leg. The APG is used to determine relative volume changes in the lower limbs in response to postural alterations and muscular exercise [23]. Each volunteer was asked to conduct a series of calf movements both standing and in the seated position as shown in the diagrammatic representation, Fig. 4.

To begin with each volunteer was asked to lie in the supine position with the right leg elevated at 45 degrees. When a steady base line was recorded on the graph the volunteer was instructed to stand with all the body weight positioned on the left leg. Once the venous filling time (VFT) and venous volume (VV) were recorded, the EVF resulting from one tiptoe movement was recorded. The RVF, the amount of blood remaining in the veins was recorded by the volunteer conducting ten tiptoe movements.

After the ten tiptoe movements the volunteer returned to the initial supine position with the leg elevated. Once a steady baseline was again recorded the volunteer was directed to sit on the edge of the examination couch and place his/her right leg on Tromped Test Jig 1. The sitting VFI and VV were then recorded. The EVF was achieved by compression of the Tromped pedal for five seconds. The RVF was recorded by the volunteer compressing the pedal ten consecutive times before returning to the supine position with the leg elevated. These procedures were repeated for all four Tromped test jigs.

Results of Experiment One are shown in Table I.

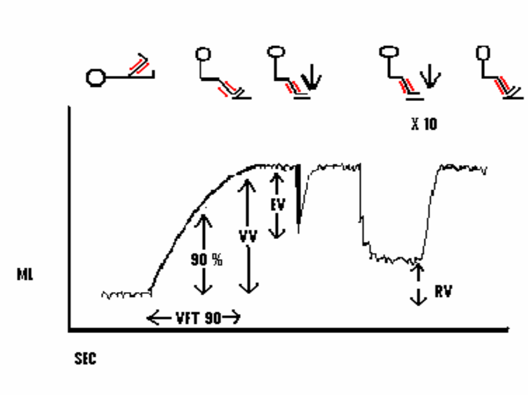


Figure 4. Diagrammatic representation of APG tracing during Tromped Assessment

Only Test Jigs 2 and 3 achieved a greater than 60% EVF, which is comparable to normal standing values, demonstrating that all ten volunteers had normal venous function. Both Test Jig 2 and Test Jig 3 had RVF's less than 35%. Test Jig 3 was reported as most user friendly by all ten volunteers and so this spring's compression value was chosen to be used in the second prototype designs.

Table 1. EVF and RVF obtained from each of the four springs.

	EVFs % (sd)	RVFs % (sd)
Tip-toe	61.016 +/- 16.085	21.745 +/- 11.25
Test Jig 1	54.1 +/- 12.83	31.66 +/- 22.62
Test Jig 2	65.99 +/- 13.83	21.54 +/- 18.87
Test Jig 3	62.173 +/- 11.63	33.71 +/- 25.02
Test Jig 4	46.052 +/- 9.84	40.75 +/- 18.37

VI. TROMPED PROTOTYPE VERSION II

These Test Jigs also consist of one pedal and base attached by a hinge at one end and a compression spring at the front end. As these designs are smaller in length to the first prototype test jig design, the spring is positioned in a completely different position in relation to the foot. An experiment is currently underway comparing the two test jig prototypes of differing lengths, to determine if this difference in spring position in relation to the foot may have some affect if any on the EVF and RVF produced. Figure 5 shows this Tromped version being assessed using the APG.

VII. TROMPED FOOTREST

Upon acquiring the footrest, it was evident that the compression spring used in the first two Tromped prototypes would not be feasible for the third design. The main reason being this spring has a free length of 101.6mm. Therefore four compression springs of a free length ranging from 35mm to 55mm which require the same force to compress it as the spring used in Test Jig 3 to a higher compression value will be selected. These smaller springs in addition to two small plastic pedals, per footrest, will not have a major weight impact on the aircraft.

Figure 5. Volunteer with right leg resting on Test Jig A.

There are four reasons for integrating the Tromped into the aircraft footrest. The first one being that the Tromped would not hamper foot space, which could hinder evacuation time and safety of passengers. The second reason being, that the Tromped could be integrated into a standard piece of equipment, which has already passed international large aircraft aviation requirements such as JAR-125 and FAR-125 approval. Thirdly as it will be integrated into the footrests which are provided by the airlines, two springs and two foot pedals of a light weight material, added to each footrest, will not greatly affect the aircraft gross take-off weight (AGTW). Finally, the author is of the belief (one which will be medically tested using the APG) that this footrest version of the Tromped will be more efficient than any of the previous two prototypes. The user will be required to place the ball of the foot on the pedals, with the arch of the foot resting on the bottom edge of the footrest. The movement then produced by compressing the footrest pedal should increase the plantar flexing of the foot, which empties the proximal calf pump into the popliteal and femoral veins [13] in comparison to when the whole foot is placed flatly on either of the first two prototype designs.

VII. CONSEQUENCES OF DVT AND FRDVT.

"A history of DVT increases the risk of chronic venous insufficiency 25.7 –fold, compared with those without a history of DVT." p. 445 [25] This increase in the risk of chronic venous insufficiency may be as a result of any of the following three factors; a remaining thrombus obstructing venous flow, an increase in venous valve incompetence as a

result of previous thrombus and dilation of veins or calf muscle pump dysfunction [23], [25]. *"limbs with chronic deep venous insufficiency can pump blood from the calf at only about one-fifth the normal rate. After exercise, the volume reduction is approximately half that achieved in normal limbs"* p. 109 [10].

As a duty of care to passengers, airlines may one day provide a form of DVT prophylaxis on board long haul flights.

An episode of DVT may lead to death by PE, and that every survivor of a DVT is at risk of life long health problems. So when a passenger on an airline suffers from a first event of DVT they then must suffer the consequences of this for the rest of their lives in the form of chronic venous disease. This will not only inhibit the distance that they can walk, but will lead to complications such as ulceration possibly, and the increased chance of developing recurrent thrombi, which could kill them a second time round.

VIII. CONCLUSION

At prototype stage, the Tromped has been medically proven to produce a sufficient EVF of greater than 60% thereby preventing venous stasis and the possible onset of FRDVT.

A medical experiment to test the efficiency of the Tromped footrest is currently undergoing ethical evaluation. It is proposed that 18 volunteers, 6 healthy, 6 with a body mass index >30 and six women taking the contraceptive pill will be recruited. Four prototypes will be fitted with springs of free lengths ranging from 35mm to 55mm, each with a different compression value. Efficiency of the Tromped footrests will be tested with each volunteer seated in the aircraft seat, which is compatible to the aircraft footrest.

Once this experiment has been completed, a test rig will be used to determine the duration of continuous use of the Tromped pedals and springs. Thereby permitting a cost benefit analysis to be conducted, this could make a form of FRDVT prophylaxis, such as the Tromped, more feasible to airlines. Especially now, that it has been established immobility due to cramped seating and the reduced cabin barometric pressure experienced, which is essential to fly at high altitudes, both precipitate towards clot formation in passengers.

REFERENCES

[1] Wright, H. P., Osborne, S. B. Effects of Posture on Venous Velocity, Measured with 24 NaCl. *Br. Heart J.* 1952;14:325-30.

[2] Kumar, P., Clark, M. (5th ed), (2002). *Clinical Medicine.* UK. WB Saunders.

[3] Bergqvist, D., Comerde, A.J., Nicolaides, A. N., Scurr, J. H., (1994). Prevention of Venous Thromboembolism. London.Med-Orion.

[4] Heit, J. A., Melton III, L. J., Lohse, C. M., Petterson, T. M., Silverstein, M. D., Mohr, D. N. O'Fallon, W. M. (2001). "Incidence of Venous Thromboembolism in Hospitalized Patients vs Community Residents". *Mayo Clin Proc*. 76, 1102-10

[5] Clement D. I., Gheeraert, P., De Buyzere, M., Duprez, D. (1994). "Medical Patients" in Bergqvist, D., Comerde, A.J., Nicolaides, A. N., Scurr, J. H. (eds) *Prevention of Venous Thromboembolism*. London: Med-Orion. Publishing Company.

[6] Hull, R. D. Raskob, G.E. and Hirsh, J. (1986). "Prophylaxis of venous thromboembolism. An overview". *Chest* 89, 3745-835.

[7] Lingamanaicker, J., Mukherjee, J. J., Fock, K. M., Khoo, T. K. (2001). The Role of Spiral Computed Tomogram in the Diagnosis of Acute Pulmonary Embolism. Singapore Med J. 42(10), 455-59.

[8] Kesteven, P. L. J. (2000). "Traveller's thrombosis". *Thorax.* Vol 55, Suppl 1, S32-S36.

[9] Murphy, J. F. A. (2001). "The Folded Deckchair Position: The problem of long-haul flights". *Irish Medical Journal* 94 (9), 260.

[10] Sumner, D. S. (2000). "Essential Hemodynamic Principles" in Cronenwett, Gbviczki, Johnston, Kempczinski, and Krupski. (eds) *Vascular Surgery, Volume 1*. USA. W. B. Saunders Company

[11] Strandness, D.E., Sumner, D. S. (1975). *Hemodynamics for Surgeons*. New York Grune and Stratton, Inc.

[12] Cole, S. E.A (ed). (2001). *Vascular Laboratory Practice, IPEM, Part IV.* England, Institute of Physics and Engineering in Medicine.

[13] Gardner, A. M.N., Fox, R.H. (1989). *The Return of Blood to the Heart: venous pumps in health and disease.* London: John Libbey & Company Ltd.

[14] Tibbs, D. J. (1992). *Varicose Veins and Related Disorders*. Oxford: Butterworth-Heinemann Ltd.

[15] Templeton, S. (2005) Flights triple blood clot risk. *The Sunday Times*. Sunday January 10 2005. p. 19.

[16] Galili, Y., Bass, A. (2002). "Long Distance Flights and the Risk of Venous Thromboembolism – A Real Threat or Just Another Flight Hysteria?" *IMAJ*. Vol. 4, 1020-22.

[17] Drummond, R., Drummond, A. 2002 On a wing and a prayer: medical emergencies on board commercial aircraft. *Canadian Journal of Emergency Medicine.* Vol. 4.

[18] Murphy, J. F. A. (2001). "The Folded Deckchair Position: The problem of long-haul flights". *Irish Medical Journal* 94 (9), 260.

[19] Beasley, R., Hughes, R.J., Hopkins, R.J., Hill, S., Weatheral, M., Vande, Water, N., Nowitz, M., Milne, D., Ayling, J., Wilsher, M. (2003). Frequency of venous thromboembolism in low to moderate risk long distance air travellers: the New Zealand Air Traveller's Thrombosis (NZATT) study. *The Lancet* 362, 2039-2043.

[20] Schreijer, A. J. M., Cannegieter, S. C., Meijers, J. C. M., Middeldorp, S., Büller, H. R., Rosendaal, F. R. (2006). "Activation of coagulation system during air travel: a crossover study". *The Lancet*, 367 (9513), 832-8.

[21] Lennon, E. (2004) Why fear of flying has just got even worse, DVT is no longer the scourge of just the economy

class. *The Irish Independent.* Tuesday May 25 2004. p. 29.

[22] Gorshtein, A., Levy, Y., Shoenfeld, Y. (2002). “Air Flights and Venous Thromboembolism - A Preventable Condition”. *IMAJ.* Vol. 4, 1080-1.

[23] Nicolaides, A. N., Summer, D. S. (1991). Investigations of Patients with Deep Vein Thrombosis and Chronic Venous Insufficiency. London, UK Med-Orion Publishing Company,.

[24] Ruckley, C.V., Fowkes, F. G. R. and Bradbury, A. W. (1999) *Venous Disease, Epidemiology, Management and Delivery of Care*, Great Britain, Springer-Verlag London Limited.

[25] Blombery, B. and McGrath, B. (2000). “Chronic venous insufficiency in post-thrombotic patients”*Clinical Science* 98, 445-47.

Integrity Enhancement of Terrain Data for Safety Critical Aviation Applications

G. Heidelmeyer, U. Klingauf
Institute of Flight Systems and Automatic Control (FSR),
Technische Universität Darmstadt, D-64287 Darmstadt
Email: {heidelmeyer,klingauf}@fsr.tu-darmstadt.de

Abstract—**Due to the appearance of data driven technologies in the aviation the impact of digital terrain data is growing more and more. Especially for ground near operations (e.g. helicopter missions) the demand for reliable terrain information is increasing rapidly.**

Some years ago digital terrain data only indicated a rudimentary representation of the earth's surface. But due to the development of new earth observation technologies a new generation of elevation models is available. However, it shall be analysed, how far such data derived by remote sensing could be used for aviation purposes.

To satisfy those special requirements the Institute of Flight Systems and Automatic Control (FSR) at the Technische Universität Darmstadt is dealing with the determination of the influencing factors which affect the quality of elevation data during the generation process. Intention of this study is to identify the error bounds based on those factors. By knowing the cause and the impact of the deviations a "Safety Buffer" should be defined in order to increase the integrity of the terrain data (figure 1).

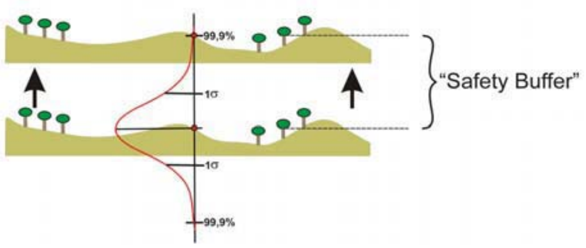


Fig. 1: Schematic depiction of a „Safety Buffer"

I. INTRODUCTION

Because of the extensive cruising radius of an aircraft reliable high quality terrain data with preferably global coverage is demanded. Formerly such terrain data was derived from military sources and a distribution to a civil customer was not usual. Also the retracement of the generation process was nearly impossible and a systematic assessment delicate. So terrain data with global coverage as well as a unique resolution and quality was nearly impossible to obtain.

With the launch of the SRTM-Shuttle mission in February 2000 the earth observation made a big step forward. The NASA as well as the DLR surveyed the surface of our planet with radar sensors out of space in order to provide a unique dataset of digital elevation information. The area under investigation covered the globe from 60° latitude north to 60° latitude south. While the NASA recorded the surface with a C-band radar sensor area wide, the DLR used an X-band sensor to provide height information for a net over the globe's surface with a mesh size of 30 kilometres.

The results gained during the SRTM-mission regarding elevation data are listed below:

1.) NASA C-Band:

-worldwide coverage, public use: *DTED1*
(post spacing about 90m)

-USA coverage, public use: *DTED2*
(post spacing about 30m)

-worldwide coverage, non public: *DTED2*
(post spacing about 30m)

2.) DLR X-Band

-worldwide coverage (net), civil use: *DTED2*
(post spacing about 30m)

At the end of the SRTM-mission a nearly worldwide elevation dataset with a unique quality can be provided. But if this kind of data can satisfy the requirements needed for aviation purposes should still be examined.

For civil aviation the RTCA and EUROCAE composed the document "*User Requirements for Terrain and Obstacle Data*" (RTCA/EUROCAE DO-276/ED-98) [*9*] in which the requirements for terrain data in all phases of flight are exactly defined. Based on this demands it must be considered if the SRTM-data even can be used during safety critical manoeuvres like approach and departure or if an enhancement of the data integrity is needed.

Usually data with a high degree of integrity for critical flight operations is provided for limited areas only. Generation methods like LIDAR (laser-scanning) or airborne radar are normally used. Photogrammetric techniques are losing importance even though they can provide terrain data of very high integrity. All these generation methods are very time consuming and costly.

Therefore a space borne system like SRTM might be a step forward for a worldwide solution.

To determine the quality and integrity of a digital terrain model, two important aspects must be considered. At first the resolution of the data affects the quality. Secondly the surface characteristic of the surveyed area influences the behaviour of the sensors and consequently the quality.
Knowing that SRTM-data is either derived from C-band or X-band radar sensors, it should be possible to figure out how radar is influenced during the surveying process by the surface reflection capacity. Figure 2 shows how the wavelength of radar affects the penetration of the microwaves into a specific surface.

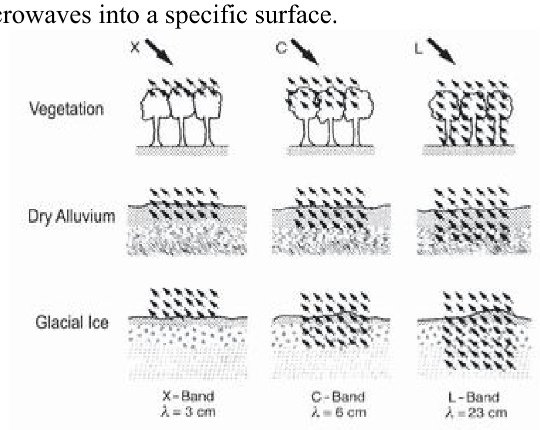


Fig. 2: Penetration into the surface of microwaves depending on the wavelength of radar [1]

There are two kinds of terrain representations: The DTM (Digital Terrain Model) without any cover as the bare earth and the DSM (Digital Surface Model). Ideally the DSM represents the earth's surface like a sheet thrown over the landscape. A DTM traditionally is obtained out of terrestrial survey. For the generation of a DSM remote sensing techniques must be applied to allow a view from above. Those methods are working with active sensors (except photogrammetry) which are scanning the surface of the globe. Because of the penetration into the surface a realistic DSM representation is hardly to create. Most of the surface models comprise the first reflected surface as an in-between of a DTM and the ideal DSM. But for aviation purposes this could be safety critical. Figure 3 shows a depiction of the different terrain representations.

Figure 2 points out that short wave X-band radar only penetrates the surface slightly and therefore should capture the real surface of the earth quite good. On the other hand C-band radar with a longer wave length penetrates deeper into the surface. Thus such kind of radar type seems to be more appropriate for a bare earth model

Concerning the applicability of the SRTM-data as a DTM multiple studies (see [8], [10], [11]) has been realized, but an assessment of the data appropriate as a DSM for aviation use is still missing.

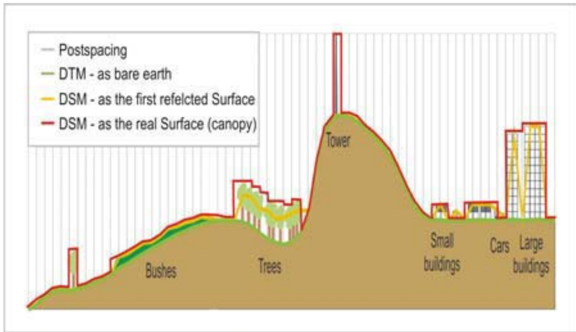


Fig. 3: Comparison of different terrain representations

II. CONCEPTUAL FORMULATION

It should be considered, if elevation information surveyed with active sensors is suitable to represent a DSM. Therefore terrain data derived through radar and laser sensors will be checked against control points and a reference terrain model. The reference model is generated by photogrammetric methods and passed a detailed verification process to guarantee the needed integrity.
Area under investigation is a region around Michelstadt in the Odenwald (Hesse). The Odenwald is a low mountain range in the south-west of Germany. The examined terrain models are described in as followed:

(1) Photogrammetric reference elevation model

Using stereo-photogrammetric methods elevation data with nearly every resolution and quality can be generated. But due to the huge amount of labour this procedure is only appropriate for bounded areas. Therefore this data will never be available for larger areas respectively worldwide.
In the scope of this study a terrain model with a post spacing of 2m (stripped 1m model) was chosen to be the reference. This height model was generated at the FSR using aerial imagery provided by the Hessian surveying office.
The horizontal reference system is the geographic coordinate system WGS84. The vertical reference of this elevation model is the WGS84 Ellipsoid.

(2)Laser Scanning elevation model

The elevation data of the laser scanning model (mission March 2000) also has been provided by the Hessian surveying office. The delivered data was already pre-processed and not in the initial LIDAR-format. Due to this processing it was not possible to associate the degree of the reflection to the height values.
The hand over coordinate system has been Gauss-Krueger and the vertical reference was similar to EGM96. To be conformant to the reference model the data have been transformed to WGS84 as the horizontal and vertical reference system.

(3) SRTM X-Band elevation model

The SRTM X-SAR data have been recorded during the SRTM-mission in February 2000 and is provided by the DLR. The X-band sensor operated with wave length of 3,1 cm at a frequency of 9,6 Gigahertz. The resolution of the elevation model is conformant to DTED2 representing a post spacing of about 30m (1 arc sec.). The horizontal and vertical coordinate system of the SRTM X-band data matches with the reference model. No coordinate transformation was needed.

(4) SRTM C-Band elevation model

The C-band data also was generated during the SRTM-mission. The SIR-C radar sensor developed by the NASA worked with a wave length of 6,0 cm at a frequency of 5,3 Gigahertz. The uncorrected elevation data for research issues is downloadable on the USGS website. For Europe elevation data in DTED1 format is available with a cell size of 90 m (3 arc seconds).

The horizontal reference is also the WGS84 coordinate system. The vertical information refers to the EGM96 geoid. So the height information had to be transformed to the WGS84 ellipsoid.

Summarising all elevation models showing the originator and the data characteristic are listed in table 1.

TABLE 1
Investigated terrain models

Generation Method	Originator	Resolution	Format
Photogrammetry	TUD/HLVA	2 m	-
Laser Scanning	HLVA	5 m	-
SRTM X-Band	DLR	30 m	DTED2
SRTM C-Band	NASA/USGS	90 m	DTED1

III. VERIFICATION OF THE TERRAIN MODELS

To execute the verification three procedures have been implemented. At first the terrain models were verified with the aid of control points. Secondly complete terrain models have been considered. Those both methods can be performed independently so that a proof of the results is possible. Finally an assessment regarding the specific surface characteristics was conducted.

A. Verification with control points

148 control points have been extracted manually out of a geo-referenced aerial photography block. This aerial photography block has been orientated with reference points surveyed with a Differential GPS (DGPS) system on site.

To prove the accuracy of a terrain model to be used as a DSM, it is important to select points on the surface of the ground coverage. At first multiple points directly on the ground like grassland, farmland and urban flanks have been surveyed, but also points on the surface of vertical spread objects like buildings or forests have been extracted. Finally these points were projected onto the surface of the elevation model to determine the height difference between the control points and the DSM under investigation.

The distribution of the control points over the investigated area at Michelstadt/Odenwald is depicted in figure 4.

Fig. 4 Distribution of the control points over the area under investigation

B. Verification using a reference model

Contrary to the verification with control points complete terrain models are compared. By subtracting the height model of the dataset under investigation from the one of the reference model a difference height model can be determined. This difference model comprises the deviation between the two datasets (see figure 5).

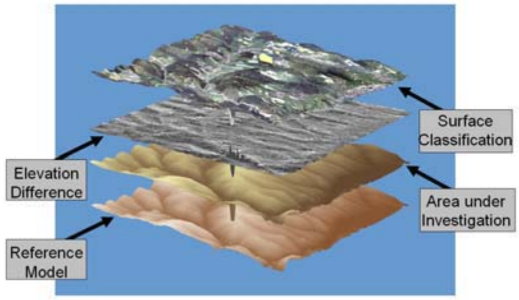


Fig. 5: Verification using a reference model

Since the surface characteristic has a significant impact on the accuracy of the height information in a next step the deviation between the two models divided by the surface characteristic should be determined.

C. Assessment of terrain models regarding specific surface characteristics

To determine the deviation based on the reflection capacity a classification of the ground cover must be performed. The classification is done with help of the topographic-cartographic information system provided by the Hessian surveying office called ATKIS. This dataset provides a nation wide unique stock of digital topographic information appropriate for a classification. Out of this data set 16 classes have been extracted which should impact the generation process. Figure 6 shows the distribution of the different topographic classes over the area under investigation. A listing of the codes with the associate feature class is listed in table 4.

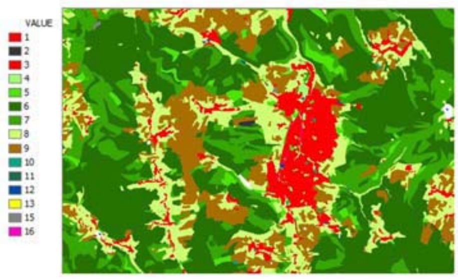


Fig. 6: Depiction of the topographic elements in the area under investigation

Using each topographic class as a template the associated part of the terrain model can been detached from the difference model. As an example the process is shown in figure 7 for the detachment of high-density areas. Finally the deviation between the reference model and the height dataset for each class can be determined. Based on these results for each class a safety buffer for the enhancement of the height model can be defined.

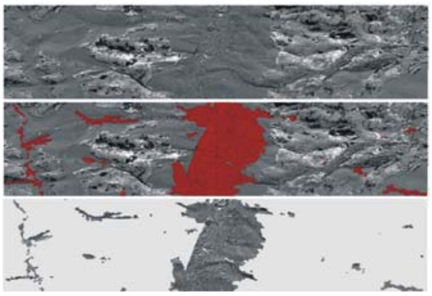


Fig. 7: Detachment of „High-density areas"

IV. INTERPRETATION OF THE RESULTS

A: Results of the control point verification

By projecting 148 control points onto the terrain models the deviation was calculated. The distribution of the difference values between the control points and the photogrammetric model referred to the surface type is shown in diagram 1. Negative values are representing points of the terrain model higher than the control points. The distribution of the positive and negative values shows that no offset in one specific direction is detectable. But it must be remarked, that the automatic generation process works more reliable on terrain sections without any cover than in sections covered with e.g. forest.

The same method was used to verify the three other terrain models (laser scanning, X-band and C-band). Finally some quality parameters like the maximum difference in non canopied areas (2D) and canopied areas (3D), the standard deviation, the mean value as well as "Confidence Level 90%" has been determined.

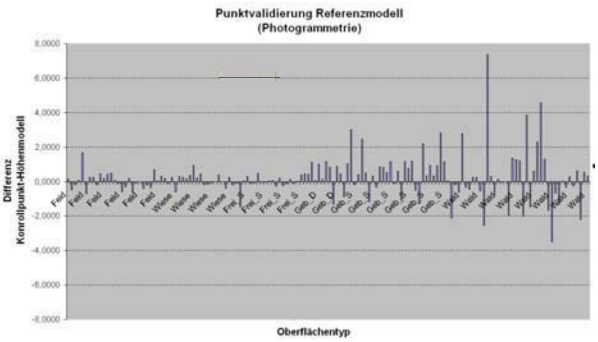


Diagram1: Point verification of the photogrammetric model

The "Level of Confidence 90%" describes the boundary which can be kept by 90% of the considered data. Hereby it must be remarked, that for aviation puposes terrain information higher than reality is not critical. Therefore only areas represented too low are of interest for the calculation of the "Level of Confidence 90%" value. A listing of the quality parameters is given in table 2.

As we can see after the terrain verification based on control points the photogrammetric generated model provides the best results. With 1,15 m as the "Level of Confidence 90%" the reference model shows a degree of accuracy that can not be provided by any of the other models (X-band 10,8m; C-band 9,3m; laser scanning 7,8m). The results of the radar models comply with the information given in the product specification (see [2]). The result for the Laser scanning data is not representative, because no distinction between the degrees of reflection could be made. Especially in forested areas this is problematically.

As we can see after the point verification the photogrammatic elevation model is suitable to be a reference model for the next step of the verification.

B. Results based on reference model verification

Following a verification of entire terrain datasets was performed. The laser scanning, the X-band as well as the C-band model has been subtracted from the reference model and finally the generated difference between the datasets was examined. The quality parameters like mean value, standard deviation and "Level of Confidence 90%" also has been determined. The results as well as the maximal difference are shown in table 3.

By comparing the results of the point verification with those with a reference model we see that there is a difference for the standard deviation and the mean value. A reason for this could be that point verifications always have a random character and a quantitative analysis is only limited possible. But for the "Level of Confidence 90%" the results are matching quite well (laser scanning 7,8m / 8,0m; X-band radar 10,8 m / 11,0 m; C-Band Radar 9,3 m / 11,0 m). Due to this outcome we can ascertain that both methods are useable for the assessment of elevation data.

C. Determination of a „Safety Buffer"

Finally the creation of a "Safety Buffer" was performed to enhance the data able to fulfil the requirements of a certain scenario.

The chosen scenario is to validate if the SRTM X-band data satisfies the requirements of the RTCA/DO-276 Document for the approach and take off area of civil aircrafts (area 2). The area 2 describes a perimeter of 45 km around the airport. In this area the recommended "Level of Confidence 90%" is about 3 meters. This implies that only 10% of the height values of a dataset shall be lower than 3 meters compared to the real height. If the needed integrity of the data can not be provided for this critical phase of a flight in a next step a "Safety Buffer" will be defined. This buffer will be added to the initial dataset and then the elevation model will be verified again.

At first we take a look at the already determined quality parameters. As we can see in table 2 and 3 the "Level of Confidence 90%" is 10,8m (point verification) and 11m (reference model verification). Hence the requirements given in the document RTCA/DO-276 could not be fulfilled and a data enhancement shall be performed to reach the desired integrity. The distribution of the difference values is given in diagram 2. The balance point of the curve lies at 5,0 m. Out of this fact we can see that the SRTM X-band data is altogether 5,0 m lower than the terrain in reality. This error seems to accrue by the penetration capacity of the radar. Anyway it should be remarked, that these facts are critical for ground near operations of an aircraft.

The simplest way to increase the integrity of the dataset is to add a global buffer of 8,0 m. A disadvantage of this method would be that this buffer impacts the entire dataset even though the error is not distributed equally. After adding a global buffer of 8,0 m the verification process has been restarted. As we can see in diagram 3 the balance point moved from 5,0 m to -3,0 m.

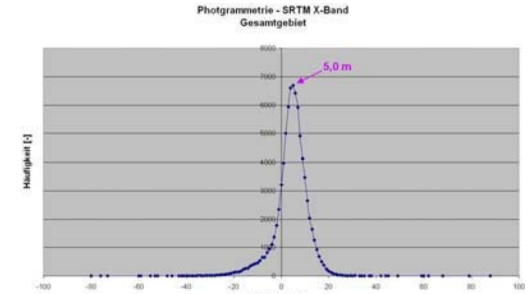


Diagram 2: Distribution of the difference between the X-band data and reference terrain model

As it could be expected this procedure lifts the entire dataset about the required value, but a real enhancement of the data can not be proved. A real enhancement of the dataset would be reached, if the balance point of the curve lies exactly on the zero point of the difference axis. Then the surface profile of the reference is nearly reached.

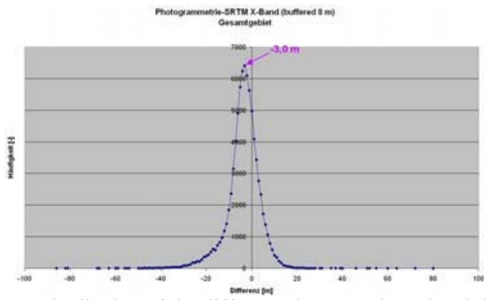


Diagram 3: Distribution of the difference between the X-band data with a global buffer of 8m and the reference terrain model

As we can see after the point verification it must be expected that ground coverage and appearing errors are correlating.

To take this factor into account the already determined difference model has been split regarding the surface characteristics with help of the topographic information of ATKIS. Then for each class the "Level of Confidence 90%" was determined. The results are shown in table 4. Because of the distribution of the error we can see that a global buffer laid on the terrain does not eliminate the effective source of errors, the 3D objects.

The verification results show that, altogether the X-band radar model is lower than the reality (mean value-point verification: 4,22 m; mean value-reference model verification 6,39m). Therefore all terrain sections must be lifted by the minimum "Safety Buffer" of each surface class (see table 4).

Now the buffered data was reassembled and the verification process was performed again. The results of the re-verification are listed in table 5. Comparing the results of the verification with the reference model before and after the adding of a "Safety Buffer" we will recognise that for the reference model verification the mean value decreased about 7,57 m. This matches approximately the value the terrain was to low to comply with the RTCA requirements. The standard deviation before the verification (6,85 m) and after the verification (7,28 m) is nearly the same. The aspired "Level of Confidence 90%" of 3,0 m could not be reached completely. But in comparison to the initial data the value could be decreased about 7,0 m to 3,5 m. Taking a look on diagram 3 we can see that the balance point of the curve nearly matches the zero point of the difference axis. Only an offset of 1,0 m is left. This fact points out, that the addition of a particular buffer lifts the terrain model exactly there where it is needed. Also we achieve the goal to modify the data in such a way that it matches nearly with the reference. With these results we can prove that it is possible to increase the "Level of Confidence" by eliminating the errors where they accrue.

To reach completely the aspired "Level of Confidence 90%" an addition of the minimum "Safety Buffer" is not sufficient and we must assume that the buffer must be defined more spaciously.

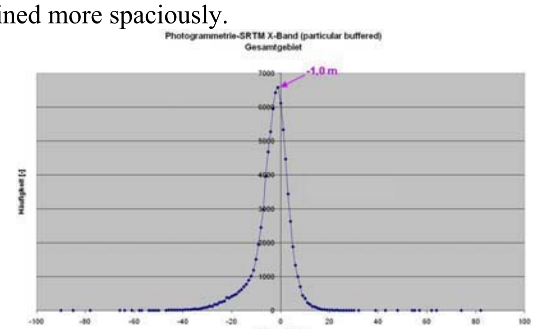


Diagram 4: Distribution of the difference between the X-band data with a particular buffer and the reference terrain model

TABLE. 2
Analysis of the control point – terrain model comparison

Control Point - Terrain Model Comparison							
Terrain Model	*max. Diff.$_{2D}$* [m]	*max. Diff.$_{3D}$* [m]	σ_{2D} [m]	σ_{3D} [m]	σ_{ges} [m]	*Mean Value* [m]	*Conf. Level 90%* [m}
Photogrammetry	1,67	7,35	0,42	1,58	1,21	0,22	1,15
Laser Scanning	-6,00	42,35	1,75	7,32	6,13	3,02	7,8
SRTM Radar X-Band	11,83	38,73	3,41	4,62	4,89	6,39	10,8
SRTM Radar C-Band	-12,14	38,97	3,29	5,19	5,79	3,62	9,3

TABLE 3
Analysis of the reference model – terrain model comparison

Reference Model – Terrain Model Comparison					
Terrain Model	*max. Difference (over Ref.)* [m]	*max. Difference (under Ref.)* [m]	*Mean Value* [m]	σ_{ges} [m]	*Conf. Level 90%* [m}
Laser Scanning – Reference	66	65	0,75	8,32	8
SRTM X-Band – Reference	80	88	4,22	6,85	11
SRTM C-Band – Reference	34	32	2,07	7,40	11

TABLE 4
Quality parameters after the subdivision into characteristic areas

Surface Type	*Coding*	*Level of Confidence 90% [m]*	*Minimum Safety Buffer [m]*
High Density Areas	1	9,5	6,5
Circulation Area	2	7,5	4,5
Significant Buildings	3	16,0	13,0
Urban Flanks	4	12,0	9,0
Deciduous Wood	5	14,5	11,5
Coniferous Wood	6	12,0	9,0
Mixed Wood	7	14,0	11,0
Greenland	8	8,5	5,5
Farmland	9	8,0	5,0
Garden Plot	10	8,0	5,0
Water	12	10,0	7,0
Quarry	15	9,0	6,0
Unclassified Areas	16	7,0	4,0

TABLE 5
Quality parameters of the corrected SRTM X-Band model

Verifcation Method	*Standard Deviation*	*Mean Value*	*Confidence Level 90%*
Control points	5,69	-0,82	5,00
Reference model	7,28	-3,35	3,5

V. SUMMARY AND FORECAST

In this study the usability of four different terrain models based on radar sensors (SRTM X- and C-band), laser scanning and stereo photogrammetry have been assessed about their usability for aviation purposes. The quality analysis has been conducted by comparing the areas under investigation with control points as well as with a complete reference terrain model. Additionally the error in correlation to the surface characteristic of the considered area has been analysed. For flat regions like grass- or farmland all generation techniques can provide satisfactory results. For areas with vertical spread cover like wood or areas of high density the accuracy of the terrain data is decreasing significantly. Therefore the conclusion is obvious, that the penetration of the sensor into the surface is conspicuous. So the deviation of the model under investigation and the reference has been investigated. Under the consideration of the surface characteristics the terrain model has been split in 16 different classes. In a next step the determined difference for each class has been added as a "Safety Buffer" to the terrain model to eliminate the error where it accrues. Then the verification process was restarted to asses the impact of the buffer.

Resulting we can say that this method is useable for an enhancement of terrain data for aviation purposes by reaching an improvement of the terrain profile regarding the reference and therefore the reality.

However it should be remarked that this examinations have been performed for a limited area only. To prove these results it would be necessary to realise a study in a wide range area. Also some other surface characteristics like desert, snow and ice should be considered. Also the influence of the terrain scope should be taken into account. Finally the information about the reflection intensity recorded during the generation should be regarded in order to examine if there is a context between the accuracy of a terrain model and of the intensity of the returned beam.

REFERENCES

[1] ALBERTZ, Jörg 2001: Einführung in die Fernerkundung: Grundlagen der Interpretation von Luft- und Satellitenbildern / Jörg Albertz. - 2., überarb. und erw. Aufl. - Darmstadt: Wiss. Buchges.

[2] DLR: Beschreibung der Mission X-SAR/SRTM: http://www.op.dlr.de/~hounam/srtm-website_de/beschreibung.html (3.6.2004)

[3] DLR: 2003: Product Description SRTM/PD-03/11/03. SRTM DTED Format, Version 1.1.

[4] DLR: 1999: X-SAR/SRTM. Shuttle Radar Topography Mission. Mapping the Earth from Space. (Pr-200-27e-6/99).

[5] DLR:http://www.dlr.de/ft/institut/abteilungen/flugzeuge/assessment (12.2.2005)

[6] DLR http://www.dlr.de/ft/institut/abteilungen/flugzeuge/uav) (12.2.2005)

[7] DLR 2004: TerraSAR-X. The German Remote Sensing Satellite

[8] KOCH, A. 2002: Qualitätsuntersuchung und Validierung von SRTM ITED-2 Daten in Niedersachsen. Schlussbericht des Instituts für Photogrammetrie und Geoinformation der Universität Hannover.

[9] RTCA 2002: User Requirements for Terrain and Obstacle Data. RTCA/EUROCAE DO-276/ED-98.

[10] SCHIECHE, Bernd, T. BUSCHE, P. HURLEMANN u. I. Janssen 2002: Lower Saxony Validation with digital elevation Models of SRTM - Los Vademos. GEG – Gesellschaft für digitale Erdbeobachtung und Geoinformation mbH Göttingen.

[11] THIEL, Karl-Heinz u. Manfred REICH 2002: Validierung von X-SAR SRTM Höhendaten mit Laserhöhen- und Laserintensitätsdaten. Institut für Navigation der Universität Stuttgart. Stuttgart

Session 7

Decision Support and Human Factors

Approaches to Operative Decision Support in the Management of Aircraft Turnaround Disruptions

Jürgen Kuster and Dietmar Jannach

Abstract— **With the availability of comprehensively integrated Airport Operational Databases, as motivated by the emergence of Collaborative Decision Making for example, new forms of automated information processing become available: Particularly, proactive Decision Support Systems can be developed based on the concepts of Artificial Intelligence. From various potential fields of application, this paper focuses on the problem of managing disruptions within the aircraft turnaround, the most typical airport ground process. With a view to supporting respective system development intentions, it describes the results of the evaluation of several approaches towards the operative provision of decision support: Particularly, strengths and limitations of simulation, planning and scheduling are discussed in the regarded context. Based on the respective findings, it is finally shown how the Resource-Constrained Project Scheduling Problem (RCPSP) – the most promising approach in our view – can be extended for its application to aircraft turnaround disruption management.**

Index Terms— **Decision Support System, Aircraft Turnaround Process, Disruption Management, Collaborative Decision Making**

I. Introduction

Even though the process that aircrafts go through between arrival and departure is well-defined in the so-called *aircraft turnaround*, a high level of time-criticality as well as various forms of dependencies make its execution prone to disruptions, delays and inefficiencies. Since each of these irregularities causes costs by provoking deviations from a predetermined and optimized schedule, continuous intervention is necessary.

In the current situation, respective Disruption Management (DM) is typically the task of human operators: They evaluate alternatives with respect to changes of priorities and resource assignments in order to minimize delays and associated penalties, aiming at the maximal efficiency of aircraft turnaround execution. In the process of identifying optimal interventions and schedule modifications, they base their decisions on (1) the available information as well as (2) the qualitative assessment of the current situation (according to individual experience and domain knowledge): Both of these aspects have to be considered if performance improvements are intended.

As regards the former, enhancements in the quality and availability of relevant information are addressed by the concept of Collaborative Decision Making (CDM): Targeting at a higher level of decision quality and overall system performance in the domain of air traffic, operative collaboration between air traffic management, airlines and airports is improved through intensified information sharing and the development of a common situational awareness (see [1], [2] for example): Relevant data is made accessible for all involved stakeholders, at the time and in the quality required. These new forms of shared knowledge implicitly open new possibilities to coordinate decisions and dispositions early and easily.

As regards the latter aspect – the assessment of situations and respective interventions according to individual experience and domain knowledge – particularly the application of intelligent systems provides promising opportunities: Based on concepts of Artificial Intelligence (AI), the available information can be analyzed automatically and therefore form the basis for the proactive provision of decision support in the operative management of aircraft turnaround disruptions. However, only few such applications have yet been developed, even though the increasing application of CDM concepts has led to the emergence of comprehensively integrated databases, which are an important prerequisite of efficient decision support.

With a view to supporting the development of near real-time Decision Support Systems (DSS), we thoroughly analyze and evaluate potential techniques for solving respective problems in this paper: Their applicability, potentials and limitations are assessed before a novel conceptual framework for the realization of operative DSS is introduced. The work presented herein is motivated by the insights and findings of a study conducted in collaboration with Deutsche Lufthansa AG, aiming at the identification of elementary requirements and potentials of turnaround-related decision support.

This paper is organized as follows: Section II gives an overview on typical problems in the application domain and illustrates the concept of disruption management by means of a typical example. Section III discusses different candidate techniques for the basis of future decision support systems: In particular, simulation, planning and scheduling approaches to the development of DSS are presented, respective strengths and weaknesses are described. Based on the findings of this analysis, Section IV introduces a novel approach for using the Resource-Constrained Project Scheduling Problem (RCPSP) as the basis of turnaround-related DSS. Section V summarizes the gained results and gives an outlook on future work.

II. Problem Statement

The process considered as *aircraft turnaround* combines all activities carried out at the airport while an aircraft is on ground. For presentation purposes we will limit ourselves to the following version of the process, which basically corresponds to the combination of core processes as mentioned by Carr [3]: The turnaround starts when the plane reaches its

J. Kuster and D. Jannach are with the Department of Business Informatics and Application Systems of the University of Klagenfurt, Universitätsstraße 65-67, 9020 Klagenfurt, Austria (e-mail: {jkuster, dietmar}@ifit.uni-klu.ac.at)

final gate or stand position. All incoming passengers leave the aircraft (*deboarding*) before it is *fueled*, *cleaned* and *catered* simultaneously. After the end of the last of these activities, the outgoing passengers enter the aircraft (*boarding*). The process finishes when the airplane finally leaves its position, heading for the runway.

Disruptions and unintended modifications of the predetermined schedule typically result from arrival delays, resource breakdowns or unavailabilities, flight cancelations, emergency landings, time slot modifications, bad weather conditions or strikes, for example. Thereby caused activity delays often ripple through the entire process to finally result in departure delays: They thus may even affect processes at remote airports. Since a link between processes conceptually corresponds to a link between process steps, we can consider a simple disruption of a single turnaround for illustrative purposes: We assume an instance of the process, in which a delay occurs already prior to deboarding. It propagates the entire turnaround due to the lack of slack times and finally results in a departure delay. Figure 1 illustrates this process: Grey bars indicate the execution times predicted *before*, white bars the ones predicted *after* the occurrence of the disruption.

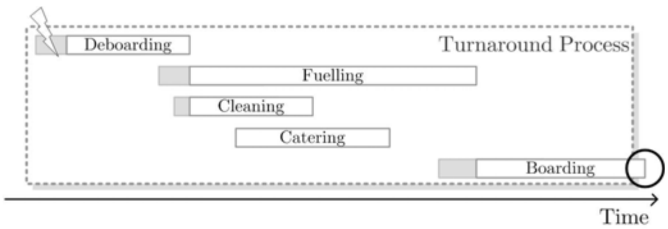


Fig. 1. Simplified Version of the Turnaround Process

In case of disruptions, deciding operators typically have several possibilities to intervene and to re-optimize the given schedule: They might accelerate processes through the assignment of additional resources, quality reductions or the parallelization of activities. They also might change priorities and reschedule affected operations, cancel flights or simply tell the apron workers to hurry up. All in all, their decisions aim at the minimization of delays and inefficiencies. For the regarded instance of a disrupted aircraft turnaround, let us assume to have a set of three potential options available: First, deboarding can be accelerated through the assignment of additional buses. Second, cleaning can be shortened if – in exchange – the cabin is additionally inspected by the cabin crew prior to boarding. And third, fueling and boarding can be parallelized if the fire brigade is available for supervision.

The operator decides on an intervention by selecting either one or a combination of several of these $n = 3$ options. If the schedule contains m processes, which are linked through resource requirements and order constraints, the number of interventions that have to be considered theoretically goes up to 2^{n*m} (including the option of not intervening at all). In realistic scenarios where the number of both possibilities and processes is high, it is therefore impossible for the human operator to analyze all of the given alternatives: Instead, the number of actually considered options is reduced to a reasonable and manageable amount based on intuition and individual experience.

Since premature restrictions of the search space may result in suboptimal decisions, the kind of DSS considered herein is intended to support the operator through the automated filtering and the proposal of relevant forms of interventions: The intelligent system shall therefore assess problematic situations and available options. This represents a complex task due to the following reasons:

- *Process Complexity.* Scheduling the aircraft turnaround can be considered complex due to a high number of dependencies: Processes are interrelated through shared resource requirements and resource flows (consider transfer passengers and cargo). Moreover, their execution depends on the current state of the environment, consisting of the given weather situation, air space restrictions as well as relevant guidelines, business rules and laws, for example.
- *Data Complexity.* Various forms of deficiencies can be distinguished for the information, which is typically available at an airport: *Inaccurateness* means that the available data does not describe reality correctly. *Inconsistency* means that various sources of information describe reality differently. *Incompleteness* means that reality is only partly described by the available data. *Instability* means that changes in the real world can implicate changes of data contents at any time. The minimization of inaccurateness, inconsistency and incompleteness is addressed by the concept of CDM and thereby motivated applications such as Lufthansa's ALLEGRO, which is concerned with the automated collection and aggregation of turnaround-related time stamp information [3]. Although these efforts are intended to finally make more precise and stable predictions possible, handling the inevitable instability remains a principle challenge in the domain of air traffic.
- *Operations Complexity.* The process of decision making is complex by itself. On the one hand, there is a high number of aircraft turnarounds that have to be managed: About 800 aircrafts per day are handled by the Hub Control Center in Frankfurt, Germany, for example. An intelligent system has to regard all of them in context: And ideally, even temporally and locally distant processes are considered if they are potentially affected. On the other hand, only little time is available for the identification of interventions: Minimum ground times are usually only about 35 to 45 minutes in Europe and even lower at US hubs [4]. It is obvious that almost no time remains for the management of disruptions and the disposition of interventions within this short period.

As regards the development of decision support systems for the operative management of turnaround disruptions, the following requirements can therefore be summarized:

- *The underlying modeling concept has to be easy and intuitive.* Since the domain of air traffic is characterized by frequent changes of operative conditions, it is of particular interest to use a model which is intelligible to and therefore maintainable by the final user. It has to support the flexible formulation and modification of process structures, time and resource constraints as well as potential interventions, costs and associated goals.

- *Applied algorithms have to provide results within reasonable response times.* The operative use of a DSS in turnaround disruption management is only possible if interventions are assessed and proposed in near real-time. The system has to find a tradeoff between optimality and computational effort: Identifying the optimal solution is desirable but not imperative.
- *Decision support has to be provided in a proactive way.* The given situation has to be analyzed and evaluated continuously, existing and pending disruptions have to be detected, a proposal on appropriate forms of interventions has to be generated and presented to the user, finally.

It is particularly the former two aspects that have to be considered in the following evaluation of fundamental approaches for the development of automated DSS.

III. Potential Approaches

In this section, simulation, planning and scheduling are presented and evaluated as candidate techniques for the provision of decision support in the operative management of turnaround process disruptions.

A. Simulation

Simulation is the process of designing a simplified model of the real world and conducting experiments on this model, aiming at an improved understanding of the regarded system's behavior or the optimization of its operation. It is typically applied whenever considered systems are too complex to be efficiently modeled through analytical methods: Particularly stochastic and dynamic real-world systems fall into this category. Depending on the points in time at which system state changes may occur, computer-based simulation can be classified into discrete-event, continuous or hybrid forms.

The power of simulation lies in the provided possibility to analyze even highly complex real-world systems in experimental execution: Performance numbers can be estimated for new policies, parameters or operating conditions; alternative designs or scenarios can be compared and evaluated; answers to ,,what-if" questions can be found. Computer-based simulation makes it possible to predict the behavior of the real-world system under various settings without disturbing actual operations or causing additional costs [5].

However, simulation itself is not a design tool, but is intended to provide the evaluation of a design which is input to the system [6]. For this reason, *simulation optimization* combines the methodologies of simulation and optimization in the search for an optimal system configuration: Given a relationship between input and output parameters which is so complex that simulation is used to estimate the output, it is concerned with the identification of a feasible setting of input variables, which minimizes or maximizes the expected value. The optimization procedure can be based on random search algorithms, sample path optimization or metaheuristics, for example (see [7] for an introductory overview). For a decision support system, which is typically concerned with the identification of an optimum, simulation therefore always has to be combined with some kind of search method.

Computer-based simulation has been applied in a variety of applications which are intended to support decision processes in the domain of airport operations: Gatersleben et al. [8] investigated passenger flows and analyzed the causes of existing bottlenecks at Schiphol Airport in the Netherlands. Similarly, Kiran et al. [9] regarded passenger and aircraft flow for the international terminal at Istanbul Ataturk Airport. Cao et al. [10] evaluated queue structures and working schedules for the check-in system at Ottawa International Airport. Lee et al. [11] analyzed the time requirements for the deicing procedure at Detroit Metro Airport and Schumacher [12] evaluated the effects of runway capacity changes at Atlanta Hartsfield International Airport by means of simulation.

Since one of the main purposes of computer-based simulation is the evaluation of various input parameters and operating policies, it has traditionally primarily been applied during system design and configuration, where typical goals are the identification of bottlenecks, requirements and optimal distributions of resources. The time required for a simulation run is not actually crucial during these early phases of system analysis. As regards the application of respective techniques in operative and near real-time decision support systems, however, the computational overhead which is typically associated with simulation [13] becomes a problem.

Even though the underlying idea of evaluating and comparing different scenarios fits pretty well with the regarded problem of turnaround process disruption management, basing a respective decision support system on simulation optimization may have its limitations:

- What is considered simulation is typically rather concerned with the analysis of artificial cases, where events and system state changes are generated randomly according to some predetermined probability distribution: It rather aims at the analysis of potential future situations than at the evaluation of an actual present situation.
- Simulation is all about abstracting the relationship between input and output variables: Each output value is considered the result of a specific combination of input values, no system-internal dependencies are analyzed. It is rather the experimental execution of a scenario based on a potential setup that is of interest.
- Due to the fact that the system transforming input into output is considered a black box in simulation, optimization can be characterized as trial-and-error procedure [14]: Such approaches are inefficient whenever the internal structure of the problem can reveal information on the reasons of disruptions and appropriate modification options: This is the case for the regarded domain.
- The inefficiency of optimization causes high levels of computational overhead: More combinations than actually required have to be evaluated in separate simulation runs. This clearly contradicts with the posted requirements of maximum efficiency and near real-time performance.

For these reasons, we propose to focus on more analytical approaches in the identification of an optimal form of intervention: Respective methods will be discussed in the following.

B. Planning

Planning is the process of selecting a sequence of actions suitable for the achievement of a previously defined goal: AI Planning is concerned with the development of automated methods for the identification of such sequences (see [15] for an overview). In this context, a planning problem is composed of the descriptions of an initial state, a goal and various actions, used to transform a given system state into another.

The aircraft turnaround is closely related to planning by nature, since already the process of originally assigning starting times to the set of relevant activities falls into the category of (temporal) planning. And also the regarded problem of managing disruptions is exactly what can be considered a planning problem: It is about *planning* interventions to return to an original schedule or about *replanning* all future activities with respect to some predefined optimization criterion. Accordingly, the exemplary case described in Section II can be regarded as the following replanning problem:

- *Relevant Actions.* The set of plannable actions consists of all remaining (i.e. future) process steps as well as all predetermined possibilities of intervention. Therefore, `deboarding`, `fueling`, `cleaning`, `catering` and `boarding` have to be considered as well as the options of assigning additional buses (`additional-buses`), reducing the scale of cleaning (`shortened-cleaning`) or parallelizing fueling and boarding (`parallelize`).
- *Initial State.* The initial state describes the current real-world situation. Since in the regarded example the disruption occurs prior to the start of the first process step, the respective initial state description merely defines that nothing has started yet.
- *Goal.* The goal of the planning process defines that all remaining activities have to be considered in the plan and that some additional optimization criterion has to be met: This auxiliary condition makes sure that optional process modifications are also taken into account. In our example, the goal consists in the identification of a plan, in which the aircraft turnaround ends on time (or at least with a minimum of delay). Moreover, it is of utmost importance in this kind of replanning problem that deviations from the original schedule are avoided as far as possible.

Note that the necessity of considering all future activities for plan generation is due to the fact that the consequences associated with the application of interventions can only be evaluated efficiently if regarded within the context of all actually and potentially affected processes.

Another reason for the consideration of planning as a basic technique is the fact, that during recent years remarkable algorithmic advances have been made in the development of automated planning systems, which can be used to solve the formulated problem: These make especially domain-independent planning applicable to realistic problems in realistic sizes (see [16] for example). Automated planning and respective systems are usually classified into the following categories [17]:

- *Domain-Independent Planning.* Concepts and systems of the most abstract class can be applied to any problem in any domain.
- *Domain-Configurable Planning.* Domain-specific knowledge is encoded directly in the provided model. This approach to planning therefore considers the fact that often a certain idea of how problems can be solved is given. Typically, knowledge concerning the order of activities falls into this category: Regarding the aircraft turnaround domain, it can – for example – be defined that deboarding always has to be executed before catering, which itself has to be finished before boarding starts.
- *Domain-Specific Planning.* Representations and algorithms are optimized for and usually limited to a certain application within a predefined domain.

Since we are particularly interested in more general and reusable approaches, we will focus on the evaluation of the two former categories in the following.

1) Domain-Independent Planning: In this section, potentials and limitations of domain-independent planning will be analyzed. After an introductory overview on formal modeling concepts, it will be discussed if and how certain parts of the regarded problem can be encoded. Respective solvers will be presented briefly before the adequacy of this form of planning for aircraft turnaround DM is summarized.

a) Formalisms and Languages: The Stanford Research Institute Problem Solver (STRIPS) was developed by Fikes and Nilsson [18]: It provides a simple and compact way of expressing planning problems, based on objects (i.e. variables or constants) and simple positive facts (i.e. ground literals without variable symbols), which in conjunction describe world states and goals. STRIPS is particularly famous for its way of action representation: A list of propositions which have to be true forms the set of preconditions, a delete-list describes the propositions which will become false, an add-list the ones which will become true upon action execution.

The Action Description Language (ADL) as proposed by Pednault [19] intends to combine the advantages of the semantically powerful Situation Calculus [20] and the rather syntactically oriented STRIPS notation. For this purpose, STRIPS is extended by negations, quantifiers, disjunctions as well as quantified and conditional effects.

During the last few years, the Planning Domain Definition Language (PDDL), which derived from ADL and several other formalisms, became the de-facto standard language for the formal description of planning problems. In domain definitions, types, constants, axioms, predicates and actions are described whereas problem definitions are used to formulate the initial state and the goal of the planning process. PDDL was introduced in 1998 in its initial version [21] and has been discussed, modified and extended ever since. With version 2.1 [22], numeric fluents, respective plan metrics for quality evaluation and durative actions have been introduced. PDDL 2.2 [23] brought derived predicates as well as timed initial literals, which can be used to express time-triggered state modifications. Finally, version 3.0 of the Planning Domain Definition Language [24] introduced state trajectory constraints as well as preferences (i.e. soft constraints). We will focus on this last and current version, which has been developed for the International Planning Competition (IPC) held in 2006, in the following.

b) Domain and Problem Encoding: The Planning Domain Definition Languages provides many of the required modeling concepts for the description of the aircraft turnaround. Numeric values are available for the dynamic modification of resource availabilities and requirements. Concepts of modeling time make it possible to schedule processes and to regard them within their temporal context: Even conditions and effects can be timed in an elementary way. However, when encoding the turnaround DM problem for the subsequent resolution within an existing domain-independent planning system, several particularities have to be considered. These are discussed in more detail in the following.

The problem that *hierarchical actions* are not supported by PDDL and that it therefore is not possible to explicitly group individual activities in one process `turnaround` can be resolved by the use of auxiliary predicates in the following structure: A process is modeled as a separate activity, which triggers the execution of its contained elements in a start effect and has the (post)condition assigned that the last of its subactivities has to be finished at its end. The resulting abstract action is described in PDDL notation in the following listing: `?p` corresponds to an instance of `process`, for which the `time` function describes the remaining duration. The `ready` and `stepready` literals are true whenever the associated process or process step is ready for execution. Similarly, `finished` and `stepfinished` are true as soon as execution has finished. In the `deboarding` activity, the preconditions must contain a reference to the associated `stepready` literal, and one end effect of the `boarding` activity has to be the activation of the respective `stepfinished` literal.

```
(:durative-action turnaround
 :parameters (?p process)
 :duration (= ?duration (time ?p))
 :condition (and (at start (ready ?p))
                 (at end (stepfinished boarding ?p)))
 :effect    (and (at start (stepready deboarding ?p))
                 (at end (finished ?p))))
```

Disjunctive conditions are required for the support of different activity execution states and varying process structures. As regards the former, consider the distinction between pending and running operations: The action `catering` has to be scheduled immediately if all requirements concerning predecessors and temporal constraints are fulfilled *or* if activity execution has already started. As regards the latter, consider the simple case of process step parallelization: The action `boarding` shall start if `fueling` has been finished *or* if the specific requirements for parallel execution are fulfilled. Although PDDL does not support disjunctions of temporally annotated expressions, disjunctive conditions can be used as long as only one temporal reference is regarded. The following listing illustrates a potential definition of the `catering` action.

```
(:durative-action catering
 :parameters (?p - process)
 :duration (= ?duration (steptime catering ?p))
 :condition (and
    (at start (not(stepfinished catering ?p))
    (at start (or (stepstarted catering ?p)
                  (stepfinished deboarding ?p))))
    (at start (stepready catering ?p)))
 :effect (and (at start(stepstarted catering ?p))
              (at end(stepfinished catering ?p))))
```

Again, the literals `stepready`, `stepstarted` and `stepfinished` are used to describe the current state of an activity. `steptime` quantifies the time required by a certain process step. The first condition in the above example defines that the action `catering` can only be started if it has not already been executed. The second condition states that the activity must either be running or that its predecessor `deboarding` has to be finished. And finally, the last condition defines that the `stepready` literal has to be true for the regarded instance.

Another problem is associated with the question of how the *effects* of an intervention can be handled: If we assume *discrete* costs (which means that all costs are charged for the first step of intervention realization), expenses can simply be associated with the execution of the initial activity of the reparation process. Whereas the assumption of discreteness is mostly acceptable for the costs of interventions, it is typically not for positive effects. The mere assignment of additional buses, for example, is not sufficient for the acceleration of `boarding` and `deboarding` by a predetermined amount of time: Instead, the buses help to accelerate the process *as long as* they are available for the turnaround. The problem of such *continuous* effects can be addressed in various ways:

- *Internal clock.* PDDL itself provides possibilities to describe continuous effects: An internal clock value `#t` can be used within the effect definition of a durative action to dynamically modify fluent values. However, since it is not possible to change an activity's duration after its initialization, this approach is inappropriate for the implementation of process acceleration.
- *Unit-wise Planning.* The logical consequence of this restriction is a reduction in the level of abstraction through the consideration of singular units: Instead of regarding `deboarding` as a rather abstract activity, the exact way of each passenger is therefore planned. However, the high number of relevant entities makes plan generation complex and the difficulty of estimating the duration of individual movements makes plans sensitive and fragile.
- *Reservation.* Another possibility to handle continuous effects is to make them discrete: This can be done by grouping all intervention-related activities into one abstract (and non-preemptive) process, which lasts as long as respective resources are required. Basically, this approach corresponds to the disposition of durative resource reservations. Since, however, it not possible to access start and end times of related (affected) activities, the required duration can not be determined exactly and has thus to be estimated in a generous way, which is (of course) highly inefficient.

The problem of providing decision support in the operative management of turnaround disruptions is a *replanning* problem: It requires the consideration of an existing schedule. Unfortunately, however, PDDL does not support respective plan modifications explicitly. The only thing possible is the definition of temporal constraints by the use of timed initial literals: Release times can be considered if the respective activity depends on the truth value of a specific literal, which is only activated at a certain point in time. To make sure, for example, that `catering` for turnaround `TA1` does not start before a given time t, a timed initial literal `(at` t `(stepready catering TA1))` shall be added to the problem description. Note, however, that this is not sufficient for comprehensive replanning, since only minimum or maximum times can be considered this way. A more flexible solution would be the use of `preference` descriptions in the goal definition: But unfortunately, even these PDDL 3.0 specific constructs do not allow full rescheduling, since it is not possible to access activity-related time points for the assessment of plans.

Traditionally, the only criterion of *plan evaluation* was goal compliance. It was only with PDDL 2.1 and the introduction of plan metrics that more flexible quality assessments became possible. In PDDL 3.0 one can finally also define *weak* goals, which shall but do not need to be fulfilled for a plan to be accepted: The `preference` element is used for the description of respective constraints, their violation can be considered by using associated `is-violated` expressions within the `:metric` section. Accordingly, the goal section in a problem definition for aircraft turnaround disruption management should look like the PDDL segment described in the following listing:

```
(:goal (and
   (all-turnarounds-finished)
   (forall (?p - process)
           (preference in-time (punctual ?p)))))
(:metric minimize (+
   (* (is-violated in-time) (cost-per-delay)))
   (* (number-delayminutes) (cost-per-delayminute)))
```

The goal definition consists of two conjunctive parts: Based on an `all-turnarounds-finished` literal, the first one defines the (hard) goal that all processes must be finished within the generated plan. The second one describes the preference `in-time`: For all instances of `process` the literal `punctual` shall be true. And also the plan metric is composed of two parts: The former considers the number of delayed turnarounds (the number of processes for which `punctual` is not true), the latter considers overall delay minutes. For both positions, associated cost values are taken from the `cost-per-delay` or `cost-per-delayminute` literal, respectively.

Whereas timed initial literals can be used as discussed above to identify if a turnaround finishes with delay and to update the truth value of `punctual` accordingly, the number of delay minutes – represented by `number-delayminutes` – can not be determined at all: It would be necessary to have access either to the current time within the action's effect declaration (in order to update the respective fluent) or to the ending time of an activity within the goal and metrics definition. However, neither is supported by PDDL, unfortunately.

c) Problem Solution: Many domain-independent planning systems are available: All of them differ in available search procedures (optimal or suboptimal planning), supported version of PDDL, the extent of implementation of respective concepts or applied algorithms. Since PDDL 3.0 has been released for the upcoming IPC only, no respective planners are yet available. Instead, some of the candidates of the last competition shall be introduced briefly.

- *CPT*[1]. The Constraint Programming Temporal (CPT) planner is an optimal planner for temporal STRIPS: However, neither durative actions with pre-, invariant and post-conditions nor numerics or metrics are supported.
- *HSP*[2]. The planning systems grouped under the name Heuristic Search Planners (HSP*) are optimal planners based on heuristic search methods. They process a PDDL2.1-like input language and can cope with durative actions, numerics and metrics: Only the fact that activity durations must not be defined through dynamically changeable values can make modeling difficult.
- *LPG-td*[3]. The Local Search for Planning Graphs planner with timed initial literals and derived predicates (LPG-td) is one of the few planning systems supporting PDDL 2.2 in its full extent. LPG-td is a suboptimal planner based on stochastic algorithms: Generated plans are therefore not necessarily identical.

d) Summary: PDDL and associated planning systems have been considered particularly due to the fact, that aircraft turnaround disruption management represents a replanning problem by nature: An existing schedule has to be updated and modified according to some real-world disruption. The algorithmic advances that have been made recently represent another argument for the consideration of domain-independent planning. Although the conceptual framework provided by PDDL allows the description of even complex structures and dependencies, there are various peculiarities of the regarded domain which are not currently supported: A major problem is the lack of possibilities to access the current time or start and end time values of contained activities: Especially the planning goal can not be described correctly within these boundaries. Another fundamental problem is the fact that no explicit support for replanning is provided. As regards associated planners, most of the potentially relevant systems are not yet able to cope with PDDL 2.2 in its full complexity. The performance of some of them has been evaluated for the airport ground-traffic control problem in a feasibility study by Trüg et al. [25], who – despite all recently made advances – declared generic planning systems as still too weak for their application to realistic problems in realistic sizes.

2) Domain-Configurable Planning: After the introduction of the concept of Hierarchical Task Networks, several planning systems are presented: The analysis of their strengths and weaknesses forms the basis for the finally provided assessment of domain-configurable planning in the regarded context.

a) Hierarchical Task Networks: In domain-configurable planning the representation of domain-specific knowledge is the core task. The most common approach is to use Hierarchical Task Networks (HTN, see [26], [27] for example), which provide possibilities to describe actions in groups and ordered structures. The following elements shall be introduced:

- A *task network* corresponds to the problem description, defining the set of tasks that have to be executed, associated ordering constraints as well as variable instantiations.

[1]Available at http://www.cril.univ-artois.fr/ vidal/cpt.en.html
[2]Available at http://www.ida.liu.se/∼pahas/hsps
[3]Available at http://zeus.ing.unibs.it/lpg

- Two categories of *tasks* can be distinguished: A *primitive* task corresponds to an action in STRIPS-style planning: It describes a state transition and can be solved by an *operator*. A *non-primitive* task, however, can not be executed directly: Since it may contain several other tasks, the planner has to figure out how to accomplish it best. A *method* is used to describe respective ways of realization.
- A *plan* corresponds to a sequence of ground primitive tasks: It describes how a given task network can be accomplished.

In contrast to STRIPS-style planning, HTN are rather task- than goal oriented: A task network is not restricted to attainment goals. It describes the tasks to execute rather than the states to achieve. HTN planning is therefore based on task decomposition and conflict resolution: Non-primitive tasks are expanded and reduced until a feasible sequence of executable primitive tasks is identified. The fact, that there exists no common language for the definition of HTN represents another difference between domain-independent and domain-configurable planning.

As far as the regarded problem of aircraft turnaround management is concerned, we assume that the finalization of the process represents a non-primitive task (`turnaround`). As such, it can be accomplished in various ways: All contained activities correspond to primitive tasks grouped in the respective method, where interventions may or may not be applied. In SHOP2's (see below) LISP-like notation, the default version of the `turnaround` task might look like described by the following listing, where `ready ?p` represents a precondition:

```
(:method (turnaround ?p)
         (ready ?p)
         ((!deboarding ?p)
          (!fueling ?p)
          (!cleaning ?p)
          (!catering ?p)
          (!boarding ?p)))
```

b) HTN Planning Systems: Domain-configurable planning can be positioned between domain-independent and domain-specific approaches: Domain-independent planning techiques are used to compile a domain-specific planner based on a certain domain description. The following systems are potential candidates for the resolution of HTNs:

- *SHOP*[4]. Due to its notable performance at the IPC 2002 and the continuous emergence of new derivatives (SHOP2 as well as the Java-based JSHOP and JSHOP2), the Simple Hierarchical Order Planner (SHOP) can be considered as one of the most vivid HTN-based planning systems right now. The current version 1.2 of SHOP2 supports many of the features provided by PDDL: Derived predicates, quantifiers, conditional paths and numerics are available. Even late-bound values, external function calls and debugging are possible.
  However, SHOP2 provides no explicit support for time and durations. The Multi-Timeline Preprocessing (MTP) approach [28], which proposes a way of describing durative activities arithmetically, is not sufficient for the regarded context since especially the modeling of resource requirements reveals problems: Whereas it is possible to consider periods of unavailability in the planning procedure, temporal shifts of requiring processes to the earliest period of availability are not.
  While the lack of support for disjunctive preconditions can be resolved by the use of separate `:method` definitions, another elementary drawback is the fact that the Simple Hierarchical Order Planner provides no explicit possibilities of considering existing plans. Moreover, the fact that SHOP2 represents a fully task-oriented planning system has to be kept in mind: It is not possible to resolve operator preconditions automatically.
  As regards its performance, an instance of the problem presented in Section II has been implemented as far as possible: By assigning durations and cost values to all activities and delays, an example in which the `parallelize` option represents the optimal form of intervention has been constructed. On a standard desktop PC, the identification of this optimum took SHOP2 about 2 seconds, which is obviously insufficient for the application of the HTN planner to realistic problems in realistic sizes.
- *HyHTN*[5]. The Hybrid HTN Planner (HTN) combines HTN resolution with fast-forward search: It therefore supports task decomposition as well as precondition achievement. HyHTN processes Object Centered Language (OCLh) descriptions, which provide no support for time and numerics.
- *SIPE-2*[6]. The System for Interactive Planning and Execution (SIPE-2) was developed by the AI center at SRI International. It supports resources and temporal constraints and can be used for execution monitoring, user interventions and replanning. However, the system is a commercial product with no evaluation version available. Moreover, it seems orphaned when the copyright is declared only from 1989 to 1995.
- *O-Plan*[7]. The Open Planning Architecture (O-Plan) provides explicit support for resources and temporal constraints: Precedence relationships, durations and allowed time windows can be assigned to activities. However, also this planning system seems to lie dormant right now.

c) Summary: All in all, the field of domain-configurable planning seems a lot less active than the one of domain-independent planning: Hardly any HTN planning systems have been developed during recent years, the number of respective IPC participants is small. Recent developments focus on planners which basically suffer from the same problems as described for PDDL-based systems: It is mainly a lack of comprehensive support for durative elements, temporal requirements and replanning, which therefore makes the use of domain-configurable planning systems impossible in the regarded context. Moreover, the observed performance indicates that respective planners are currently too weak for the application to realistic problem sizes.

[4]Available at http://www.cs.umd.edu/projects/shop

[5]Available at http://scom.hud.ac.uk/scomdl2/hyhtn/index.html
[6]Available at http://www.ai.sri.com/∼sipe
[7]Available at http://www.aiai.ed.ac.uk/∼oplan

C. Scheduling

Scheduling is the process of assigning times and resources to activities, respecting precedence and availability constraints. Compared to planning, which rather focuses on *what* to do, scheduling is therefore mainly concerned with the question of *when* to take action. Respective concepts have been considered as a basic technique for the DSS of interest since turnaround DM can be regarded as a rescheduling problem: The starting times of future activities have to be updated according to a given real-world disruption. Moreover, the limitations that have been identified for automated planning systems implicitly suggest to focus on rather time-oriented approaches.

1) Conceptual Framework: From the set of available classes of scheduling problems (see [29] for an overview), it is particularly the Resource-Constrained Project Scheduling Problem (RCPSP) which shall be considered for modeling the aircraft turnaround domain: As a generalization of the job-shop scheduling problem it provides possibilities to define arbitrarily linked activities which are processed on arbitrary resource types. It can be formalized as follows: A project consists of a set of activities $\mathcal{A} = \{0, 1, ..., a, a+1\}$: The first and the last element correspond to fictitious start and end activities, which require no resources and have a duration of 0. Each remaining $i \in \mathcal{A}$ has a non-negative duration d_i assigned. For activity execution, a set of renewable resource types $\mathcal{R} = \{1, ..., r\}$ is available, with a constant amount of u_k units disposable for a type k. Activities are ordered by a set of precedence constraints $\mathcal{P}$: The existence of $p_{i,j} \in \mathcal{P}$ states that activity j must not start before the end of activity i. The relationship between activities and resource types is defined by a set of resource requirements $\mathcal{Q}$: Activity i requires $q_{i,k} \in \mathcal{Q}$ units of resource type k throughout its execution.

Describing an RCPSP is easy and intuitive: Activities, resources and constraints represent abstract constructs which allow the definition of problems on a conceptual level. In this context, the turnaround process can be described as a resource-constrained project, with process steps modeled as ordered activities. Unfortunately, however, the description of available interventions is not possible within the Resource-Constrained Project Scheduling Problem: Neither optional activities nor dynamic structural modifications are supported.

2) Solution Methodologies: For the resolution of scheduling problems, basically the following methodologies can be distinguished (for comprehensive overviews see [30], [31]):

- *Mathematical Programming (MP).* Particularly early research focused on the identification of optimal solutions based on MP: Integer Programming [32], Dynamic Programming [33] and Branch-and-Bound procedures [34], [35] have been applied, for example. Integer linear formulations of the RCPSP are mostly based on time-indexed variables, where one decision variable is necessary for any potential combination of activity and starting time [36]: This makes problems large and implies high computational requirements. Another difficulty results from the underlying low-level formalism: Modeling usually requires deep mathematical knowledge and resulting models are quite hard to understand.
- *Constraint Programming (CP).* In Constraint Programming, problems are modeled as Constraint Satisfaction Problems (CSP): Respective instances contain a set of variables, associated domains, and a set of constraints describing the relationships between the variables. In the process of searching an optimal solution, constraint propagation and consistency enforcing techniques can be applied: The domains of the contained variables are reduced as far as possible in a preprocessing step. Afterwards, particularly techniques of Artificial Intelligence and (increasingly) of Operations Research are used for the identification of an optimum (see [37], for example). Scheduling problems can be expressed as CSP based on precedence and resource constraints: Compared to mathematic formulations, Constraint-Based Scheduling (CBS, see [38] for example) thus provides possibilities of higher level modeling, which means that the complexity of defining and maintaining domain models is significantly reduced. However, the problem of rapidly increasing problem sizes remains, due to the common approach of applying time-indexed variables: Computational efforts remain high for the identification of optimal solutions.
- *Heuristic Procedures.* The combinatorial nature of the RCPSP makes the determination of exact optimal solutions difficult: The successful application of such optimization approaches and the resolution of problems within reasonable time have only been reported for relatively small problem sizes. By contrast, Heuristic Procedures rather focus on the generation of *good* solutions in shorter time. They make it possible to dynamically trade off between solution quality and the time spent searching. In the context of the RCPSP, it is particularly an abstraction of temporal concepts, which makes remarkable reductions of the search space possible: Instead of working directly with time values, a common approach is to use abstract representations of the schedule (activity lists, random keys, etc.), which can be optimized more easily [39]. As regards the optimization procedure itself, various forms of metaheuristics have been implemented: Tabu search, simulated annealing and genetic algorithms represent only examples, comprehensive overviews and evaluations are available (as in [40] for example).

For the problem of providing decision support in turnaround DM, the requirement of near real-time performance clearly suggests the use of heuristic procedures.

IV. A Framework for Near Real-Time Decision Support Systems

According to the evaluation results, we propose to base near real-time decision support systems in the operative management of turnaround disruptions on the concept of the Resource-Constrained Project Scheduling Problem. It is considered the most promising approach due to the following reasons:

- The RCPSP provides explicit support for many of the required concepts: It is possible to model time, resources and respective constraints in an easy and intuitive way.

- Heuristic and metaheuristic procedures can be used for the identification and optimization of solutions. Incremental search makes it possible to provide good results even in minimal time. This corresponds to the requirements of operative disruption management, where short performance times are rather of interest than global optima.
- The RCPSP has been and still is subject to extensive research: Respective work focuses on conceptual extensions as well as the further improvement of algorithms.

In this section, we propose an approach to overcome the previously mentioned limitations of the RCPSP. It is shown, how the existing modeling concept can be extended to support optional activities and structural process modifications before some remarks on resolution methodologies are made.

A. Extending the RCPSP

It has already been mentioned that the RCPSP is not sufficient for the description of optional process modifications. Therefore, we propose its generalization in the Extended Resource-Constrained Project Scheduling Problem (*x-RCPSP*), where available interventions are described by means of alternative activities: The application of an intervention corresponds to the switch from a default version of an activity to one of its alternatives. Accordingly, only the subset of currently chosen (*active*) activities is considered and contained in the final schedule.

The *x-RCPSP* is based on the idea of placing an abstract layer on top of the elementary constructs of the classical RCPSP: $\mathcal{A}^+$, $\mathcal{P}^+$ and $\mathcal{Q}^+$ contain all potentially relevant activities, precedence constraints and resource requirements. From these supersets, concrete instances of the RCPSP can be generated: $\mathcal{A} \subseteq \mathcal{A}^+$, $\mathcal{P} \subseteq \mathcal{P}^+$ and $\mathcal{Q} \subseteq \mathcal{Q}^+$ group actually relevant elements. In their instantiation process, a choice on a set of active activities is made before all associated precedence constraints and resource requirements are activated. The modification of the activation state is driven by three sets:

- $\mathcal{A}^0$ describes the reference process: Contained activities are activated by default and are preferred over others.
- $\mathcal{X}^+$ describes valid activity substitutions: The existence of $x_{i,j} \in \mathcal{X}^+$ states that activity i can be replaced by activity j. Note, that this relationship is not commutative.
- $\mathcal{M}^+$ describes mutual dependencies: The existence of an element $m^{\ominus}_{i,j} \in \mathcal{M}^+$ states that activity j has to be deactivated upon the activation of i whereas the existence of $m^{\oplus}_{i,j} \in \mathcal{M}^+$ states that j has always to be activated along with activity i.

Based on these constructs, the aircraft turnaround can be formalized as is summarized in Table I: $Start$ and End correspond to fictious start and end activities, all other process steps are represented by the first three letters of their name. Deb^{Bus} is the accelerated alternative to deboarding, as is Cle^{Red} to cleaning. Ins is the activity of cabin inspection by the crew and Fue^{Par} is the parallelized version of fueling. As regards the notation of relations, $i \rightarrow j$ defines that $p_{i,j} \in \mathcal{P}^+$, $i \triangleright n \times k$ defines that $q_{i,k} = n \in \mathcal{Q}^+$, $i \Rightarrow j$ defines that $x_{i,j} \in \mathcal{X}^+$, $i \Leftrightarrow j$ defines that $x_{i,j}, x_{j,i} \in \mathcal{X}^+$, $i \oplus j$ defines that $m^{\oplus}_{i,j} \in \mathcal{M}^+$ and $i \ominus j$ defines that $m^{\ominus}_{i,j} \in \mathcal{M}^+$.

TABLE I
FORMAL DESCRIPTION OF THE EXEMPLARY TURNAROUND PROCESS

Set	Content
$\mathcal{R}$	Bus, $Firebrigade$
$\mathcal{A}^0$	$Start$, Deb, Fue, Cat, Cle, Boa, End
$\mathcal{A}^+$	$Start$, Deb, Deb^{Bus}, Fue, Fue^{Par}, Cat, Cle, Cle^{Red}, Ins, Boa, End
$\mathcal{P}^+$	$Start \rightarrow Deb$, $Start \rightarrow Deb^{Bus}$, $Deb \rightarrow Fue$, $Deb \rightarrow Fue^{Par}$, $Deb \rightarrow Cat$, $Deb \rightarrow Cle$, $Deb \rightarrow Cle^{Red}$, $Deb^{Bus} \rightarrow Fue$, $Deb^{Bus} \rightarrow Fue^{Par}$, $Deb^{Bus} \rightarrow Cat$, $Deb^{Bus} \rightarrow Cle$, $Deb^{Bus} \rightarrow Cle^{Red}$, $Fue \rightarrow Boa$, $Fue^{Par} \rightarrow End$, $Cat \rightarrow Boa$, $Cle \rightarrow Boa$, $Cle^{Red} \rightarrow Ins$, $Ins \rightarrow Boa$, $Boa \rightarrow End$
$\mathcal{Q}^+$	$Deb \triangleright 1 \times Bus$, $Deb^{Bus} \triangleright 2 \times Bus$, $Fue^{Par} \triangleright 1 \times Firebrigade$
$\mathcal{X}^+$	$Deb \Leftrightarrow Deb^{Bus}$, $Fue \Leftrightarrow Fue^{Par}$, $Cle \Leftrightarrow Cle^{Red}$
$\mathcal{M}^+$	$Cle^{Red} \oplus Ins$, $Cle \ominus Ins$

The knowledge on available interventions is therefore encoded directly within the domain model. The option of assigning additional buses can be applied by switching from Deb to Deb^{Bus}, as made possible by the existence of $Deb \Leftrightarrow Deb^{Bus}$ in $\mathcal{X}^+$. For the switch from Cle to Cle^{Red}, the set of mutual dependencies $\mathcal{M}^+$ defines that the *optional activity* of cabin inspection has to be activated. And the *structural modification* of parallelizing fueling and boarding corresponds to a simple switch from a default version Fue to an alternative Fue^{Par}, that is not linked to boarding through a precedence relation.

B. Problem Resolution

The aim of DM is the identification of an optimal set of interventions: Potential options include both mere temporal shifts of activities as well as structural modifications. While the former kind is covered by the existing RCPSP-specific optimization procedures, it is particularly its combination with the latter type of intervention that has to be addressed in the *x-RCPSP*. Apart from responding to the question of *when* to start activities, it has also to be considered *what* steps to execute at all: Planning has to be combined with scheduling.

For this purpose and a first evaluation of the proposed extensions, we have adopted an evolutionary algorithm for the heuristic resolution of the RCPSP [41] to our specific requirements: As a metaheuristic approach, the respective procedure makes it possible to identify good solutions even in short time. Optimization is accomplished through the continuous evolution of a population (consisting of different solutions): Based on the idea that combinations of good solutions might result in even better ones, the best solutions are selected, recombined and slightly modified in each iteration. Recombination (so-called *crossover*) is based on two parents, with one of them prescribing the activation state of the child and the other one defining the order of contained activities. Modification (so-called *mutation*) is based on the random shift or exchange of the child's active process steps.

On a first prototype implementation, 20 instances of the exemplary process have been regarded in context: The best solution found within 4 seconds on a standard PC was only about 25% below the theoretical optimum. These results are promising, particularly since the current implementation is still unoptimized and does not consider domain-specific knowledge. For further details see [42].

V. Conclusion

In this paper, potential approaches to the problem of providing decision support in the operative management of aircraft turnaround disruptions have been evaluated: Potentials and limitations of simulation, planning and scheduling approaches have been discussed and assessed in this context. Finally, it has been shown how the RCPSP – as the most promising approach in our view – can be extended by support for optional activities and alternative process execution paths for its application to the regarded problem: The respective approach to modeling interventions can easily be adopted to other problem classes.

Future work will be directed at the conduction of additional measurements, the further enhancement of the discussed optimization procedures and the development of a first prototype version of a respective decision support system: Results from practical testing and a final evaluation based on real-world test cases from industrial partners will be provided.

Acknowledgment

This work is part of the *cdm@airports* project, which is carried out in cooperation with FREQUENTIS GmbH (Austria) and is partly funded by grants from FFF (Austria).

References

[1] R. Hoffman, M. Ball, A. Odoni, W. Hall, and M. Wambsganss, “Collaborative decision making in air traffic flow management,” UC Berkeley, Tech. Rep., 1999.

[2] M. Wambsganss, “Collaborative decision making through dynamic information transfer,” *Air Traffic Control Quarterly*, vol. 4, pp. 107–123, 1997.

[3] F. R. Carr, “Robust decision support tools for airport surface traffic,” Ph.D. dissertation, 2004.

[4] C.-L. Wu and R. E. Caves, “Modelling and optimization of aircraft turnaround time at an airport,” *Transportation Planning & Technology*, vol. 27, no. 1, pp. 47–66, 2004.

[5] A. M. Law and W. D. Kelton, *Simulation Modelling and Analysis*. McGraw-Hill, 1982.

[6] M. Mollaghasemi and G. Evans, “Multicriteria design of manufacturing systems through simulation optimization,” *IEEE Transactions on System, Man, and Cybernetics*, vol. 24, no. 9, 1994.

[7] M. C. Fu, F. W. Glover, and J. April, “Simulation optimization: A review, new developments, and applications,” in *Proceedings of the 2005 Winter Simulation Conference*, 2005, pp. 83–95.

[8] M. R. Gatersleben and S. W. van der Weij, “Analysis and simulation of passenger flows in an airport terminal,” in *Proceedings of the 1999 Winter Simulation Conference*, 1999, pp. 1226–1231.

[9] A. S. Kiran, T. Cetinkaya, and S. Og, “Simulation modeling and analysis of a new international terminal,” in *Proceedings of the 2000 Winter Simulation Conference*, 2000, pp. 1168–1172.

[10] Y. Cao, A. L. Nsakanda, and I. Pressman, “A simulation study of the passenger check in system at the ottawa international airport,” in *Proceedings of the 2003 Summer Computer Simulation Conference*, 2003, pp. 573–579.

[11] J. T. Lee, S. Chen, and A. Daskalakis, “Deicing decision support tool,” in *Digital Avionics Systems Conference*, 2004, pp. 2.E.6.1–2.E.6.9.

[12] B. Schumacher, “Proactive flight schedule evaluation at delta air lines,” in *Proceedings of the 1999 Winter Simulation Conference*, 1999, pp. 1232–1237.

[13] C. M. Harmonosky and S. F. Robohn, “Investigating the application potential of simulation for real time control decisions,” *International Journal of Computer Integrated Manufacturing*, vol. 8, p. 126132, 1995.

[14] J. Vankateswaran and Y. J. Son, “Impact of modelling approximations in supply chain analysis - an experimental study,” *International Journal of Production Research*, vol. 42, no. 15, pp. 2971–2992, 2004.

[15] J. Rintanen and J. Hoffmann, “An overview of recent algorithms for ai planning,” *Künstliche Intelligenz*, vol. 2, pp. 5–11, 2001.

[16] D. S. Weld, “Recent advances in planning,” *AI Magazine*, vol. 20, no. 2, pp. 93–123, 1999.

[17] M. Ghallab, D. Nau, and P. Traverso, *Automated planning: theory and practice*. Morgan Kaufmann Publishers, 2004.

[18] R. E. Fikes and N. J. Nilsson, “STRIPS: a new approach to the application of theorem proving to problem solving,” *Artificial Intelligence*, vol. 2, pp. 189–208, 1971.

[19] E. P. D. Pednault, “ADL: exploring the middle ground between STRIPS and the situation calculus,” in *Proceedings of the first international conference on Principles of knowledge representation and reasoning*. Morgan Kaufmann Publishers Inc., 1989, pp. 324–332.

[20] J. McCarthy, “Situations, actions and causal laws,” Stanford University, Tech. Rep., 1963, reprinted in: *Semantic Information Processing*, M. Minsky, Ed. MIT Press, 1968, pp. 410-417.

[21] M. Ghallab, A. Howe, C. Knoblock, D. McDermott, A. Ram, M. Veloso, D. Weld, and D. Wilkins, “PDDL—the planning domain definition language,” Yale University, Tech. Rep., 1998.

[22] M. Fox and D. Long, “PDDL 2.1: An extension to PDDL for expressing temporal planning domains,” *Journal of Artificial Intelligence Research*, vol. 20, pp. 61–124, 2003.

[23] S. Edelkamp and J. Hoffmann, “PDDL 2.2: The language for the classical part of the 4th international planning competition,” Tech. Rep., 2004.

[24] A. Gerevini and D. Long, “Plan constraints and preferences in PDDL3 - the language of the fifth international planning competition,” University of Brescia, Italy, Tech. Rep., 2005.

[25] S. Trüg, J. Hoffmann, and B. Nebel, “Applying automatic planning systems to airport ground-traffic control a feasibility study,” in *KI 2004*, ser. Lecture Notes in Computer Science 3238, S. Biundo, T. Frühwirth, and G. Palm, Eds. Springer Verlag, 2004, pp. 183–197.

[26] E. D. Sacerdoti, *A Structure jor Plans and Behavior*. Elsevier Scientific Publishing, 1977.

[27] K. Erol, J. Hendler, and D. S. Nau, “Semantics for hierarchical task-network planning,” Computer Science Department, University of Maryland, Tech. Rep., 1994.

[28] D. Nau, T.-C. Au, O. Ilghami, U. Kuter, W. Murdock, D. Wu, and F. Yaman, “SHOP2: An HTN planning system,” *Journal on Artificial Intelligence Research*, vol. 20, 2003.

[29] P. Brucker, *Complex Scheduling Problems*, ser. Osnabrücker Schriften zur Mathematik, 1999, no. 214.

[30] O. Icmeli, S. S. Erenguc, and C. J. Zappe, “Project scheduling problems: A survey,” *International Journal of Operations & Production Management*, no. 11, pp. 80–91, 1993.

[31] W. Herroelen, B. D. Reyck, and E. Demeulemeester, “Resource-constrained project scheduling: a survey of recent developments,” *Computers and Operations Research*, vol. 25, no. 4, pp. 279–302, 1998.

[32] A. A. B. Pritsker, L. J. Watters, and P. M. Wolfe, “Multiproject scheduling with limited resources: A zero-one programming approach,” *Management Science*, vol. 16, no. 1, pp. 93–108, September 1969.

[33] S. E. Elmaghraby, “Resource allocation via dynamic programming in activity networks,” *European Journal of Operational Research*, vol. 64, no. 2, pp. 199–215, 1993.

[34] J. Stinson, E. Davis, and B. Khumawala, “Multiple resource-constrained scheduling using branch-and-bound,” *AIIE Transactions*, vol. 10, no. 3, pp. 252–259, 1978.

[35] E. Demeulemeester and W. Herroelen, “A branch-and-bound procedure for the multiple resource-constrained project scheduling problem,” *Management Science*, vol. 38, no. 12, pp. 1803–1818, December 1992.

[36] S. Demassey, C. Artigues, and P. Michelon, “Constraint-propagation-based cutting planes: An application to the resource-constrained project scheduling problem,” *Journal on Computing*, vol. 17, pp. 52–65, 2005.

[37] P. Baptiste, C. L. Pape, and W. Nuijten, “Incorporating efficient operations research algorithms in constraint-based scheduling,” in *1^{st} Joint Workshop on Artificial Intelligence and Operational Research*, 1995.

[38] P. Baptiste, C. LePape, and W. Nuijten, *Constraint-Based Scheduling*. Kluwer Academic Publishers, 2001.

[39] R. Kolisch and S. Hartmann, “Heuristic algorithms for solving the resource-constrained project scheduling problem: Classification and computational analysis,” in *Project scheduling: Recent models, algorithms, and applications*, J. Weglarz, Ed., 1999, pp. 147–178.

[40] S. Hartmann and R. Kolisch, “Experimental evaluation of state-of-the-art heuristics for resource constrained project scheduling,” *European Journal for Operational Research*, vol. 127, no. 2, pp. 394–407, 2000.

[41] K. Hindi, H. Yang, and K. Fleszar, “An evolutionary algorithm for resource-constrained project scheduling,” *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 5, pp. 512–518, 2002.

[42] J. Kuster and D. Jannach, “Extending the resource-constrained project scheduling problem for disruption management,” in *Proceedings of the 3rd IEEE Conference on Intelligent Systems*, 2006, to appear.

The feasibility of a new air traffic control concept from a human factors perspective

M. Heese, R. Schuen-Medwed and K. W. Kallus

Abstract— **This document reports the feasibility of the Central European Air Traffic Services (CEATS) Multi Executive Sector (MESEC) concept, tested in a real-time simulation project designed by EUROCONTROL CEATS Research, Development and Simulation Centre (CRDS) Budapest to satisfy the requirements of the CEATS Strategy, Planning and Development Unit (CSPDU) Prague, focusing on the three main human factors 'Co-operation', 'Situation Awareness' (SA) and 'Workload'. Method: Based on the current literature regarding the assessment of those human factors in air traffic control (ATC) and in order to answer the basic question whether the MESEC concept is feasible or not, it was hypothesized that there is a significant difference between the MESEC organisation and the conventional organisation regarding the three main human factors co-operation, SA and workload. Furthermore it was assumed that co-operation and workload will increase in the MESEC organisation in higher traffic load conditions, while SA should decrease collapsing across the two different controller working positions (executive controller (EC) and planning controller (PC)). Results: It was found out that there is a significant difference between the conventional organisation and the MESEC low traffic organisation regarding SA and workload. Moreover co-operation and workload increase significantly in the MESEC organisation in higher traffic load conditions, while SA is significantly decreased. Conclusions: The MESEC concept from a human factors perspective is feasible with restrictions on the amount and the complexity of the traffic, the number of conflicts as well as the working procedures. The recommendation for future MESEC simulations is to include investigations on a bigger sample size, a clear definition of working procedures, task sharing and responsibilities, changes in the Human Machine Interface (HMI) as well as having available system co-ordination and tools such as the medium-term conflict detection.**

***Index Terms*— air traffic control, cooperation, situation awareness, workload.**

Manuscript received February 21, 2006. This work was sponsored by Dr. Philippe Debels at the European Organisation for the Safety of Air Navigation (EUROCONTROL) CEATS Research-, Development- and Simulation Centre in Budapest (Hungary).

M. Heese was with the Department of Psychology (Section Working-, Organisational- and Environmental Psychology) at the University of Graz (Austria) (e-mail: Michaela.heese@feichtinger.biz).

R. Schuen-Medwed is with the CEATS Research-, Development- and Simulation Centre in Budapest (Hungary) (e-mail: renee.schuen-medwed@eurocontrol.int).

K. W. Kallus is with the Department of Psychology (Section Working-, Organisational- and Environmental Psychology) at the University of Graz (Austria) (e-mail: wolfgang.kallus@uni-graz.at).

I. Introduction

Stanley Roscoe [1] gives a complete overview of the development of human factors in aviation as a major source of engineering psychology, defined as the science of human behavior in the operation of systems. Among the earliest experimental studies of the human factors in equipment design were those made during World War II at the Applied Psychology Unit of Cambridge University, England, under the leadership of Sir Frederick Bartlett.

Human factors in aviation are involved with the study of human capabilities, limitations, and behaviours and the integration of that knowledge into the system we design for them with the goals of enhancing safety, performance, and the general well-being of the operators of the system [2].

The latest air traffic forecasts predict that traffic will double by 2020 due to the development of various factors such as globalisation, economic factors, business and environmental costs [3]. The worldwide rapid growth of air travel is putting immense pressure on the ATC system [4]. Due to airspace capacity limitations, several proposals have been put forward for modernizing ATC to meet the demands for enhanced capacity, efficiency and safety. Foremost among these proposals for changes in ATC procedures is for example the effort to give the users more freedom in their operations within the airspace. These new future Air Traffic Management (ATM) concepts may mean a change of the ATCOs' roles and responsibilities. The role of the controllers tends to change from one of active control to passive monitoring.

One of these concepts is the Multi Executive Sector Concept (MESEC) designed by the EUROCONTROL CEATS Research, Development and Simulation Centre (CRDS) Budapest [5] characterized by one planning controller (PC) working together with two executive controllers (ECs), while usually two-person teams handle one sector. In a conventional organisation, one PC is preplanning the traffic for the EC to handle the aircraft, as long as it crosses their sector (area of responsibility). If the traffic is expected to increase for a longer period, a new sector is opened by the supervisor and another pair of controllers starts to work in this new sector.

The focus of this study is on the three main human factors co-operation, situation awareness (SA) and workload as defined in the following.

A. Co-operation

Although the concept of co-operation has been appearing in nearly every research field until today, it seems to be difficult to find a complete definition. In the social sciences co-operation is associated with a common group term, while in economic and organisational sciences as well as computer sciences co-operation is linked to co-ordination. Moreover co-operation is attached to communication in all disciplines. Piepenburg [6] tried to define interdependabilities of co-operation, communication and co-ordination in a work related context. Following his ideas co-operation refers to a conscious, systematic and harmonized activity between two or more individuals, that ensures the achievement of each individual's objectives to the same degree. Hoc and Lemoine [7] investigated the concept of co-operation in the scope of the cognitive evaluation of human-human and human-machine co-operation modes in ATC.

Finally, gathering together past definitions of the co-operation concept and taking into account practical experiences of co-operation in ATC, for this study co-operation is defined as 'a positive interference opposite to competition (negative interference) between two or more controllers working towards a common goal based on task sharing, principal roles and responsibilities written down in the ATC manual, summarizing verbal and nonverbal communication, as well as the term 'co-ordination' in the form of human machine interface (HMI) inputs. Besides, individual factors based on each person's personality, knowledge and experience also influence co-operative behaviour and the quality of the controller's team performance.

B. Situation Awareness

Regarding the concept of SA Endsley [8] defines SA as the 'perception of the elements in the environment within a volume of time and space, comprehension of their meaning and the projection of their status in the near future'. Furthermore Endsley [9] highlights that 'SA in the aviation setting is challenged by the limitations of human attention and working memory. The development of relevant long-term memory stores, goal-directed processing, and automaticity of actions through experience and training are seen as the primary mechanisms used for overcoming these limitations to achieve high levels of SA and successful performance.' To complete the picture it has to be mentioned that while Endsley [9] focuses on a memory-oriented assessment of SA, and Finnie and Taylor [10] deal with the element of perception and behaviour and Kallus et al. [11] work on a process oriented approach based on the model of cognitive aspects of ATC which refers to the ATCOs' 'mental picture' and 'mental model' based on Hoffmann's [12] action-oriented psychological models.

C. Workload

DeWaard [13] tried to characterize mental workload as the specification of the amount of information processing capacity that is used for task performance. Issues regarding how the goal is reached and individual restrictions imposed upon performance are included in the concept. Hilburn and Jorna [14] first tried to distinguish the terms 'task load', as the demand imposed by the ATC task, and 'traffic load', as the 'number-of-aircraft-under-control' from the ATCO's actual 'workload'. System factors such as airspace demands, interface demands and other task demands contribute to task load, while operator factors like skill, strategy and experience determine workload.

II. EMPIRICAL BACKGROUND

Regarding the assessment of co-operation, SA and workload related to traffic load in ATC, this study is oriented on the 'The moderator effects of task load on the interplay between en-route intra-sector team communications, SA and mental workload' conducted by Bailey et al. [15]. Using a high-fidelity ATC simulator Bailey et al. [15] investigated certified ATC specialists in ten three persons teams, who performed routine ATC tasks within a single factor (low vs. high workload) repeated measure design.

Due to the lack of multivariate studies, which interpret co-operation, SA and workload, the precise relationship between these three human factors remains unclear. Nevertheless one of the more common findings across studies is that as the average number of aircraft increases, there is a corresponding increase in controller-pilot communications, mental workload, and a decrease in SA [16 and 17].

Furthermore Brookings & Wilson [18] investigated the 'Psychophysiological responses to changes in workload during simulated ATC' and showed that the eye blink frequency decreases in higher traffic load and higher traffic complexity conditions. Jorna [19] also highlighted that the measurement of psychophysiological factors such as respiration rate, temperature, heart rate, heart rate variability, eye movements and eye blinks, brain activity and temperature can provide valuable and unique information, which is not available through either subjective or behavioural measures.

A. Question 1

Is the Multi Executive Sector (MESEC) concept concerning the three main human factors co-operation, SA and workload feasible? Note that the term feasibility can have many different meanings such as operability or practicability and that it is closely related to other factors such as applicability, affectivity or efficiency. Therefore in this study feasibility is defined as the capability of performing tasks successfully with the available resources related to co-operation, SA and workload.

B. Question 2

How does the traffic load affect the main three human factors within the MESEC organisation?

III. METHOD

A. Experimental Design

This pilot study is based on an incomplete multifactor, multivariate, within subject design and investigates the effect of the two independent variables 'traffic organisation' (conventional vs. MESEC organisation) and 'traffic load' (low= 40ac/h; medium= 60ac/h and high= 75ac/h) on the three dependant variables co-operation, SA and workload. Five different conditions were randomly assigned to 32 (four exercises per day, 8 days long) follow ups assuming that each condition appears at least once in each row and each column. It should be pointed out that the experimental plan was conducted for the needs of the MESEC small scale real time simulation 4 (SSRTS4) at CRDS Budapest [20] including civil and military traffic, as well as group briefings. Thus eight civil and four military controllers (n=12) were participating at 28 exercises and four briefings (N=32). It should be noted that in the conventional exercises two civil and two military controllers were sitting at a measured position. In the MESEC exercises three civil controllers and two military controllers and in the conventional exercises with military involvement two civil controllers and one military controller were sitting on a measured position. The remaining controllers were working on feed and pilot positions, which were not measured.

The dependent variable co-operation is collected from the radio telecommunication (rt) calls and the Human Machine Interface (HMI) recordings stored in the EUROCONTROL CRDS Log Analyser. Specifically four determinants served as dependent variables: (a) the average number of HMI clicks in specific the cleared and rejected flight levels, the assigned speeds and headings, as well as route directings and the assigned vertical rate and (b) HMI co-ordinations meaning counter proposals, rejects and accepts, as well as the (c) average number and the (d) average duration of the rt calls. Besides, the data gained from the systematic observation of the ATCOs' behaviour and interview data served as indicators for the qualitative results supporting the quantitative variables of the concept of co-operation.

The dependent variable SA is based on four closed questions on a 5-point rating scale adapted to the SASHA (= Situational Awareness for Solution for Human-Automation Partnerships in European ATM) [21] Questionnaire presented after each simulation exercise. Specifically, (a) the average value of the first three questions aiming at the prediction of the evolving traffic situation, the awareness of what the team member was doing and the focus on single problems in an area or sector and (b) the average value of the overall SA served as indicators for the statistical analysis. Additionally two open questions were asked in the interview regarding the prediction of the evolving traffic situation and temporarily loosing SA. The outcome of the qualitative content analysis following Mayring [22] supported the quantitative results of the variable SA.

The dependent variable workload is deducted from the average value of the Instantaneous Self Assessment (ISA) [23] recorded from the simulator and the average value of six closed questions on a 5-point rating scale of the NASA Task Load Index (TLX) [24].

The eye blinks were derived from the ElectroOculoGram with a sensitivity of 0.7 mV, a time constant of 0.25 Hz and a storage tact of 50 times per second [25] recorded with the Physiomodulsystem. The average value of the eye blink frequency per minute (60 seconds) served as dependent variable.

B. Participants

In total four military (OAT) and eight civil (GAT) male Air Traffic Control Officers (ATCOs), between the ages of 26 and 50 participated in simulated radar-based ATC tasks using four-person teams (GAT EC/PC and OAT EC/PC). The participants were recruited from the CEATS Simulation Expert Groups from Austria, the Czech Republic, Hungary, Northern Italy, Slovenia and Slovakia and were all licensed ACC (Area Control Centre) controller for their country. The majority had more than ten years of experience, while two ATCOs only had four to six years experience. All participants were paid 68 Euro as a daily allowance plus travel expenses in addition to their regular salaries.

C. Apparatus

The radar-based ATC exercises were simulated on the ESCAPE simulator in the operations room at CRDS Budapest. Additionally, the EUROCONTROL Recording and Graphical display On-line (ERGO) [23] component provided subjective values of the controller's workload (very low-very high) on a five-point rating scale in real time. All voice recordings were handled by the AudioLan Voice Communication System attached to the simulator, while the Human Machine Interface (HMI) recordings were directly collected through the simulator. Finally the recorded data were available in the CRDS Log Analyser Tool Version 2.9. All control positions were provided with a communications touch panel with telephone and intercom ('press and hold') functionality.

The eye blinks were recorded via the 'Physiomodule'-system [26] consisting of a small opto-coupled interface with three slots to charge up to three portable Physiomodules.

D. Test material

The CRDS engineers conducted 28 ATC exercises lasting 80 minutes each. The exercises were set up from the three different traffic organisations and the three different traffic loads according to the sector capacities. The simulation area compromised the CEATS Upper Information Region (UIR: FL 285-660) of the eastern part of the Slovak Republic and Hungary (D4 sector), as well as the middle part of Hungary, defined as sector D7. In order to feed traffic into the simulation, parts of the neighbouring Flight Information Regions (Czech Republic and Poland = Feed North, Romania

= Feed East and Croatia = Feed South) were included in the simulation. The MESEC SSRTS4 was run in the fixed route network environment ARN V4. The meteorological conditions were set to calm surface wind velocity, standard linear wind direction and speed above FL 300. Short term conflict detection and area proximity warning was available to civil and military controllers for a minimum vertical separation of 600 ft below and 1600 ft above FL410. 1600 ft were also set as a minimum vertical separation above FL290 between RVSM flights and non-RVSM state flights. The minimum horizontal separation was 4.9 nautical miles.

E. Analysis

Due to the fact that the data collection was mainly based upon subjective assessment techniques as well as psychophysiological data each controller was first treated as a single case. The single case analysis was based on so-called ‘single case randomization tests’ [27 and 28]. Afterwards the single case data were meta-analytically agglutinated [29]. For p-values Krauth [30] recommends the following formula, proposed by Shirata (1981), because the test statistic is approximately standard normal distributed. The test statistic for Shirata’s method is defined as ‘K’ delivered by the following formula: $K = -\sum_{i=1}^{k} \ln(P_i/(1-P_i))$. K therefore results from summing up the natural logarithm of the actual probabilities divided by the possible probabilities multiplied by (-1). According to Shirata the null hypothesis is rejected, if K’ is smaller than the critical z-value of the standard normal distribution. K’ is calculated with the formula: $K' = (\frac{3}{k\pi^2})1/2\,K$. Next the alpha level of 5% was adjusted following the Bonferoni-Holm method [31].

F. Hypothesis 1

There is a significant difference between the conventional organisation and the MESEC low traffic organisation regarding the three main human factors, co-operation, SA and workload collapsing across the two different controller working positions (EC/PC).

G. Hypothesis 2

There is a significant difference between the conventional organisation and the MESEC low traffic organisation regarding eye blinks to be related to the three main human factors co-operation, SA and workload on the planning controller position.

H. Hypothesis 3

More co-operation within the MESEC organisation will occur under high traffic load conditions, as compared to lower traffic load conditions collapsing across the two different controller working positions (EC/PC).

I. Hypothesis 4

ATCOs will perceive greater difficulty maintaining SA within the MESEC organisation under higher traffic load conditions, as compared with lower traffic load conditions collapsing across the two different controller working positions (EC/PC).

J. Hypothesis 5

ATCOs will experience greater mental workload within the MESEC organisation under higher traffic load conditions, as compared with lower traffic load conditions collapsing across the two different controller working positions (EC/PC).

IV. RESULTS

The statistical analysis was conducted for two hypotheses groups. It has to be highlighted that the available conditions per case were summarized to execute the randomisation tests. The first hypothesis group (later indicated as D versus A) was related to the first basic question whether or not there was a significant difference between the conventional organisation (D) and the MESEC low traffic organisation (A) regarding co-operation, SA and workload. The conventional organisation and the MESEC low organisation were based on the same traffic sample. The second group of hypotheses (later indicated as A</>B</>C) according to the second basic question investigates whether or not there was a significant difference between the MESEC low (A), medium (B) and high traffic condition (C) regarding co-operation, SA and workload.

The probability (‘p’)-values indicate the relevant test statistic. The error risk alpha was controlled at a 5% level. The resulted test statistics in the second last row of each table are compared to the adjusted alpha levels in the last row to decide whether the null hypotheses shall be maintained or have to be rejected. Statistically significant data are printed in bold.

A. Co-operation

As presented in Table I there was no significant difference between the conventional organisation (D) and the MESEC low traffic organisation (A) related to the concept of co-operation according to the average number of HMI clicks, HMI co-ordinations and rt calls as well as the average duration of rt calls ($K_{1(n=7,\ -2.22)}$=-1.62, ns.; $K_{2(n=7,\ -2.46)}$=-1.12, ns.; $K_{3(n=7,\ -1.96)}$=-.88, ns.; $K_{4(n=7,\ -1.85)}$=-.68, ns.).

TABLE I
SINGLE-CASE AGGLUTINATION FOR CO-OPERATION (D≠A)

		Dependant variables			
n	No.	HMI clicks	HMI coord.	rt calls (rtc)	Φ rtc duration
1	2	.31	.71	.02	.03
2	4	.50	.48	.57	.53
3	5	.73	.64	.49	.45
4	7	.29	.28	.69	.77
5	9	.11	.61	.56	.44
6	10	.58	.15	.64	.68
7	12	.70	.43	.76	.76
8	13	missing	missing	missing	missing
Shirata's K		-1.62	-1.12	-0.88	-0.68
Shirata's K'		0.0134	0.0193	0.0247	0.0318
Adjusted α		0.0080	0.0100	0.0125	0.0167

Table II shows that there was significantly more co-operation found within the MESEC organisation under high traffic load conditions, as compared to lower traffic load conditions collapsing across the two different controller working positions (EC/PC) ($K_{1(n=8,\ -2.34)}$=-6.92, p<.05; $K_{2(n=8,\ -2.30)}$=-3.40, p<.05; $K_{3(n=8,\ -2.41)}$=-3.71, p<.05; $K_{4(n=8,\ -2.74)}$=-7.00, p<.05).

TABLE II
SINGLE-CASE AGGLUTINATION FOR CO-OPERATION (A<B<C)

		Dependant variables			
n	No.	HMI clicks	HMI coord.	rt calls (rtc)	Φ rtc duration
1	2	.27	.60	.03	.03
2	4	.41	.47	.58	.05
3	5	.53	.68	.63	.62
4	7	.98	.73	.84	.82
5	9	.10	.59	.42	.33
6	10	.42	.11	.60	.62
7	12	.55	.37	.72	.69
8	13	.09	.18	.08	.09
Shirata's K		**-6,92****	**-3,40****	**-3,71****	**-7,00****
Shirata's K'		**0.0026**	**0.0056**	**0.0051**	**0.0027**
Adjusted α		0.0050	0.0070	0.0063	0.0060

** $K<Z_{CRIT}$

B. Situation Awareness

As demonstrated in Table III there was a significant difference detected related to the mean SA score gained from the first three questions of the SAWL questionnaire comparing the conventional- (D) and the MESEC low traffic (A) organisation ($K_{1(n=7,\ -2.11)}$=-10.21, p<.05). According to the overall SA score no significant differences could be found ($K_{2(n=7,\ -2.11)}$=.04, ns.).

TABLE III
SINGLE-CASE AGGLUTINATION FOR SA SCORES (D≠A)

		Dependant variables	
n	No.	Mean SA (Q1-3)	Overall SA
1	2	.05	1
2	4	.61	.63
3	5	.42	.57
4	7	.31	.79
5	9	.02	.50
6	10	.01	.45
7	12	.26	.13
8	13	missing	missing
Shirata's K		**-10.21****	0.04
Shirata's K'		**0.0021**	0.5811
Adjusted α		0.0045	0.0500

** $K<Z_{CRIT}$

It was true that the ATCO's had greater difficulties maintaining SA in the MESEC high traffic load (C) condition compared to the MESEC medium (B) and the MESEC low traffic (A) organisation collapsing across the two different controller working positions (EC/PC) ($K_{1(n=8,\ -0.15)}$=-10.08, p<.05; $K_{2(n=8,\ -0.15)}$=.04, p<.05;) (view Table IV). According to the overall SA score no significant differences could be found ($K_{2(n=7,\ -0.15)}$=.04, ns.).

TABLE IV
SINGLE-CASE AGGLUTINATION FOR SA SCORES (A>B>C)

		Dependant variables	
n	No.	Mean SA (Q1-3)	Overall SA
1	2	.14	1
2	4	.36	.26
3	5	.42	.77
4	7	.22	.73
5	9	.04	.45
6	10	.03	.31
7	12	.29	.22
8	13	.41	.76
Shirata's K		**-10.08****	0.04
Shirata's K'		**0.0019**	0.4340
Adjusted α		0.0038	0.0250

** $K<Z_{CRIT}$

C. *Workload*

The third variable investigated within the MESEC concept was workload gained once from the ISA recordings and once from the NASA TLX. As depicted in Table V there was a significant difference between the conventional (D) and the MESEC low traffic organisation (A) in both data sources ($K_{1(n=7, -2.98)}$=-14.56, p<.05; $K_{2(n=7, -2.88)}$=-10.37, p<.05).

TABLE V
SINGLE-CASE AGGLUTINATION FOR WORKLOAD RATINGS (D≠A)

		Dependant variables	
n	No.	ISA recordings	NASA TLX
1	2	.11	.68
2	4	.08	.02
3	5	.01	.35
4	7	.32	.34
5	9	.01	.02
6	10	.50	.23
7	12	.48	.30
8	13	missing	missing
Shirata's K		**-14.56****	**-10.37****
Shirata's K'		**0.0015**	**0.0021**
Adjusted α		0.0036	0.0042

** $K<Z_{CRIT}$

Table VI shows that the workload was rated significantly lower in the MESEC low traffic condition (A) compared to the higher traffic load conditions (B and C) for both data sources collapsing across the two different controller working positions (EC/PC) ($K_{1(n=8, -3.00)}$=-25.79, p<.05; $K_{2(n=8, -3.00)}$=-21.81, p<.05).

TABLE VI
SINGLE-CASE AGGLUTINATION FOR WORKLOAD RATINGS (A<B<C)

		Dependant variables	
n	No.	ISA recordings	NASA TLX
1	2	.01	.70
2	4	.14	.03
3	5	.08	.11
4	7	.13	.45
5	9	.02	.04
6	10	.34	.09
7	12	.40	.22
8	13	.73	1
Shirata's K		**-25,79****	**-21,81****
Shirata's K'		**0.0007**	**0.0009**
Adjusted α		0.0031	0.0033

** $K<Z_{CRIT}$

V. DISCUSSION

A. *Statistical Conclusion*

The first hypothesis, whether there is a significant difference between the conventional organisation and the MESEC low traffic organisation related to the concept of co-operation, has to be rejected according to the average number of HMI clicks, HMI co-ordinations and rt calls as well as the average duration of rt calls. There is significantly more co-operation found within the MESEC organisation under high traffic load conditions, as compared to lower traffic load conditions collapsing across the two different controller working positions (EC/PC). Bailey et al. [31] also found out that in higher task load conditions the total number of communication events increases compared to lower task load conditions, but the difference was not statistically significant. In fact the concept of co-operation in ATC has still not been sufficiently investigated.

In conclusion MESEC is possible, provided that the proposed definition of co-operation is sufficient. Looking at the interview data the ATCOs see restrictions on the traffic amount, the traffic complexity, the working methods and the technical support.

There is a significant difference detected related to the mean SA score gained from the first three questions of the SAWL questionnaire comparing the conventional- and the MESEC low traffic organisation. According to the overall SA score no significant differences can be found. It is true that there are greater difficulties maintaining SA in the MESEC high traffic load condition compared to the MESEC medium and the MESEC low traffic organisation collapsing across the two different controller working positions (EC/PC). The results in this study correspond to Bailey et al.'s [31] results to a significant SA in higher task load conditions. The amount of SA is continuously rated high in the questionnaires, which indicates that MESEC is workable even in higher traffic load conditions. But according to the interviews the ATCOs came up with many concerns related to the SA within the MESEC concept.

Thus in conclusion it is no problem to maintain SA within MESEC in low traffic conditions, with low traffic complexity, only a few crossing conflicts, when everything works smoothly so that every person from the three persons team working in two sectors is fully in the loop of what's going on.

The first hypothesis group testing whether there is a significant difference related to workload between the conventional and the MESEC low traffic organisation is true for the ISA records and the NASA TLX score. Moreover the workload is rated significantly lower in the MESEC low traffic condition compared to the higher traffic load conditions collapsing across the two different controller working positions (EC/PC).

The results for workload also orientate on Bailey et al.'s

[31] assumptions that workload in increased in higher traffic load conditions.

Table VII gives a summary of the basic questions, hypotheses and statistical conclusions.

Although the workload gained from the ISA recordings is generally rated lower compared to the NASA Task Load Index, it can be assumed that the two subjective assessment tools are correlated in the MESEC high traffic organisation (compare Fig 1).

TABLE VII
OVERVIEW OF THE BASIC QUESTIONS, HYPOTHESES AND STATISTICAL CONCLUSIONS

Questions	**Data source**	**Decision for H_1**
Q1. Is the MESEC concept regarding the three main human factors co-operation co-operation (CO-OP), situation awareness (SA), and workload (WL) feasible?	**all**	**See Hypotheses Group 1**
Hyp. 1.1. There is a significant difference between the conventional organisation and the MESEC low traffic organisation regarding the three main human factors, CO-OP, SA and WL collapsing across the two different controller working positions (EC/PC).	all (except eye blinks)	CO-OP: REJECTED SA + WL: KEPT
Hyp. 1.2. There is a significant difference between the conventional organisation and the MESEC low traffic organisation regarding eye blinks to be related to the three main human factors, CO-OP, SA and WL on the planning controller position.	Eye blinks + syste-matic ob-servation s of the ATCOs' behaviou r	EX-CLUDED
Q2. How does the traffic load affect the main three human factors within the MESEC organisation?	**all**	**See Hypotheses Group 2**
Hyp.2.1. More CO-OP within the MESEC organisation will occur under high traffic load conditions, as compared to lower traffic load conditions collapsing across the two different controller working positions (EC/PC).	Audio-Lan recordsHMI inputs, observa-tions, inter-views	KEPT
Hyp. 2.2. ATCOs will perceive greater difficulty maintaining SA within the MESEC organisation under higher traffic load conditions, as compared with lower traffic load conditions collapsing across the two different controller working positions (EC/PC).	SA question-naire, inter-views	KEPT
Hyp. 2.3. ATCOs will experience greater mental WL within the MESEC organisation under higher traffic load conditions, as compared	NASA TLX (Part1), ISA	KEPT

It was assumed that there is a significant difference between the conventional organisation and the MESEC low traffic organisation regarding eye blinks to be related to the three main human factors co-operation, SA and workload on the planning controller position. But due to a malfunction of the recording system the eye blink data are only complete for one single case (compare Fig 2).

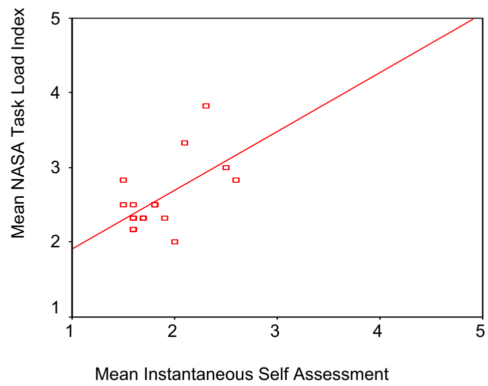


Fig. 1. Mean NASA Task Load Index correlated to the mean Instantaneous Self Assessment Score (MESEC high traffic condition). The mean Instantaneous Self Assessment Score and the mean NASA Task Load Index on the EC and PC position were summarized (n=18) to calculate a Pearson correlation under the assumption that the data are interval scaled and standard normal distributed. (1=very low, 2=low, 3=medium, 4=high, 5=very high)

Nevertheless it is empirically proved that the eye blink rate decreases while workload increases. So far there are no consistent relationships found regarding the eye blink rate and the concept of co-operation or SA in ATC.

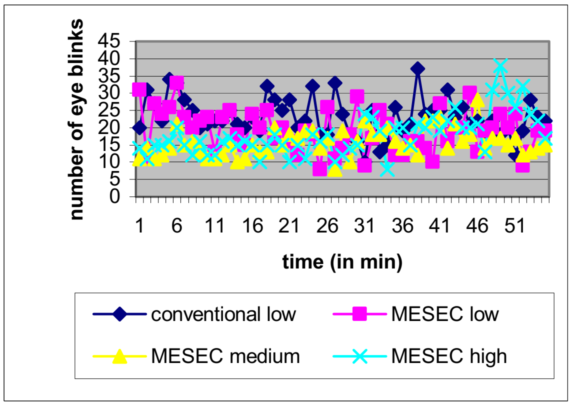


Fig. 2. Number of eye blinks per minute for Case 3. The first and the last ten minutes in the figure were recorded as baseline during the exercise feeding and clearing.

B. Recommendations

One of the general strengths of the MESEC concept is that it can be used for short term peaks in quiet traffic flows and less complex traffic. MESEC can therefore be useful for situations during nightshifts for example, when the traffic amount is too high for two persons and not high enough for four persons. The conventional procedure is to split a sector or rather open a new sector when the traffic amount rises, which means that two persons, completely unfamiliar with the traffic situation start to work. This handover is a quite time-consuming procedure. The new team is first briefed on the actual traffic situation by the existing team. Then they sit down on the controller working position, familiarise with the traffic picture and finally disclose that they are ready to take over. Not until then the supervisor adapts the sector configuration to the actual traffic situation. In MESEC there is always the PC available who is totally in the loop, because he actually stays on his position. Therefore one more benefit of MESEC is that the handover of a sector is a lot easier and less time critical as the procedure is no longer possible in a very high traffic load.

On the other hand one of the major concerns within MESEC is the fact that three persons working instead of four persons are definitely not safer, because eight eyes are supposed to monitor the traffic more careful than six eyes. The main challenge for the MESEC concept will be to gain the ATCOs acceptance regarding further development and implementation. It is therefore absolutely necessary that the benefits are shown clearly from an operational perspective. So far, the ATCOs, who participated in the MESEC SSRTS 4 could envisage applying the MESEC concept subject to the following restrictions. It is only possible to have MESEC (a) in a manageable airspace, (b) in low density traffic, (c) with less complex traffic, (d) with identical information on the screen, (e) if the speakers of both collaborating ECs are on, (f) if the the exit flight level problem is solved in terms of unique procedures, (g) if the task sharing and responsibilities are clearly defined, (h) if system coordination (SYSCO) is available and (i) if tools such as the medium term conflict detection (MTCD) are on-hand.

ACKNOWLEDGMENT

Thanks to my supervising professor Dr. K. Wolfgang Kallus for his professional guidance, stimulating suggestions and encouragement. Special thanks to Fidel Chetcuti for the correction of the English language and his friendly support and to Dr. Sophie Dusire and Mag. Renée Schuen-Medwed for their professional help and critical remarks. A million thanks to the air traffic controllers, who participated at the MESEC SSRTS4, for all their patience and help during the two weeks of investigations!

REFERENCES

[1] Roscoe, St. N., (1997). *The Adolescence of Engineering Psychology*, Human Factors History Monograph Series, Vol.1; Published by the Human Factors and Ergonomics Society. Retrieved 3rd Nov., 2005 from http://www.hfes.org/PublicationMaintenance/FeaturedDocuments/27/adolescencehtml.html

[2] Koonce, J. M. (1979). A Historical Overview of Human Factors in Aviation. In D.J. Garlad, W. Hopkin and V.D. Hopkin (Eds.) *Handbook of Human Factors*. (p.3) Hillsdale, N.J.: Lawrence Erlbaum.

[3] EUROCONTROL (2004). *EUROCONTROL Capacity Enhancement Planning.* EATMP Infocentre, Brussels, Belgium. Traffic Forecast 2004-2025. Retrieved April 19, 2005, from http://www.eurocontrol.int/statfor/public/subsite_ homepage/homepage.html

[4] Metzger, U., Parasuraman, R. (2001). The role of the air traffic controller in future air traffic management: An empirical study of active control versus passive monitoring. *Human-Factors.* 43 (4), 519-528.

[5] EUROCONTROL CRDS (2005). *CEATS-Central European Air Traffic Services Programme.* Retrieved Feb. 9, 2005, from http://www.eurocontrol.int/ceats/public/subsite_homepage/homepage.html

[6] Piepenburg, U. (1991). Ein Konzept von Kooperation und die technische Unterstützung kooperativer Prozesse. [A cooperation concept and the technical support of kooperative processes] In H. Oberquelle (Ed.) *Kooperative Arbeit und Computrunterstützung* p. 79-99. Stuttgard, Germany: Hogrefe.

[7] Hoc, J. M. and Lemoine, M. P. (1998). Cognitive evaluation of human-human and human-machine co-operation modes in air traffic control. *The International Journal of Aviation Psychology, 8(1),* 1-32. Lawrence Erlbaum Associates.

[8] Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In: *Proceedings of the Human Factors Society 32nd Annual Meeting* (Vol. 1), 97-101. Santa Monica: Human Factors Society.

[9] Endsley M. R. (1999). Situation Awareness in Aviation Systems. *SA Technologies* 257-276.

[10] Finnie, S. & Taylor, R. (1998). The Cognitive Cockpit. *Flight Deck International*, UK & International Press.

[11] Kallus, K.W., Barbarino, M. & VanDamme, D. (1997). *Model of the Cognitive Aspects of Air Traffic Control.* EUROCONTROL, Brüssel, Ref.No. HUM.ET1.ST01.1000-Rep-02, (55 S.).

[12] Hoffmann, J. (1992). Probleme der Begriffsbildungsforschung: Von S-R Verbindungen zu S-R-K Einheiten. [Problems in the concept formation research: from S-R combinations to S-R-K Units] In: *Sprache & Kognition*, 11, Heft 4, S. 223-238.

[13] De Waard, D. (1996). *The measurement of driver's mental workload.* Ph.D. Thesis, University of Groningen. Haren, The Netherlands: University of Groningen, Traffic Research Center.

[14] Hilburn, B., Jorna, P. G. A. M. (1999). Workload in Air Traffic Control. In P.A. Hancock and P. Desmond [Eds.] *Stress, Workload, and Fatigue: Theory, Research, and Practice.* Hillsdale, New Jersey, USA: Lawrence Erlbaum Associates.

[15] Bailey, L. L., & Willems, B. F. (2002). The moderator effects of taskload on the interplay between en route intra-sector team communications, situation awareness, and mental workload. US: *Aviation Medicine*. Report Nr. DOT/FAA/AM-02/18.

[16] Morrison, R. & Wright R. H. (1989). ATC control and communication problems: An overview of recent ASRS data. In R.S. Jensen (Ed.) *Proceedings of the Fifth International Symposium on Aviation Psychology*, 901-07.

[17] Morrow, D., Lee, A. & Rodvolt, M. (1993). Analyzing problems in routine controller-pilot communication. *International Journal of Aviation Psychology 3,* 285-302.

[18] Brookings, J. B., Wilson, G. F. & Swain C.R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology 42*, 361-377.

[19] Jorna, P. G. A. M. (1998). Automation and free(er) flight: exploring the unexpected. In *Proceedings of the Third International Conference on Automation Technology and Human Performance*. Norfolk, Virginia, USA. March 24-27,1998.

[20] EUROCONTROL CRDS (2004). *CEATS Prototyping Real/Time Simulation No. 2*. Budapest, Hungary: CRDS/SIM/rtS-2834-PET. Retrieved Feb. 9, 2005, from http://www.eurocontrol.int/crds/ documents.html

[21] EUROCONTROL (2003). *The development of Situation Awareness Measures in ATM Systems.* Brussels, Belgium, by O. Straeter and M. Woldring. Retrieved Feb. 9, 2005, from http://www.eurocontrol.int/ceats/public/subsite_homepage/homepage.html

[22] Mayring, P. H. (2000). *Qualitative Inhaltsanalyse* [Qualitative content analysis]. Forum Qualitative Sozialforschung / Forum: Qualitative Social

Research [Online Journal], 1 (2). Retrieved, May 18th, 2005 from http://qualitative-research .net/fqs/fqs-d/2-00inhalt-d.htm.

[23] EUROCONTROL Experimental Centre (EEC) (1996). *ERGO (Version 2) for Instantanteous Self Assessment of workload in a real-time ATC simulation environment.* Bretígny, France: H. Hering and G. Coatleven. Retrieved Sept. 9, 2005 from http://www.eurocontrol.int/eec/gallery/content/public/documents/EEC_notes/1996/EEC_note_1996_10.pdf

[24] Hart, S. G. & Stavelend, L. E. (1988). Development of the NASA-TLX (Task Load Index): results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.) *Human mental workload*. Holland: Elsevier.

[25] Schandry, R. (1998). Augenbewegungen und Lidschlag. [Eye movements and Eye Blink] In R. Schandry *Lehrbuch Psychophysiologie Körperliche Indikatoren psychischen Geschehens* (pp.274-278, 281-282). Weinheim, Germany: Beltz Psychologie Verlags Union.

[26] Rimkus Medical Engineering. *The Physiomodul-System.* Retrieved March 10th, 2005, from www.rimkusmed.de

[27] Edgington, E. S. (1995). *Randomization tests.* 3. ed., rev. and expanded, New York: Dekker.

[28] Todman, J. B. & Dugard, P. (2001). *Single-Case and small-n experimental designs. A practical guide to randomisation tests*. NY: Lawrence Erlbaum Associates.

[29] Kallus, K.W. (1986). Agglutination of Single-Case Data: Appliction to Individual Data in Experimental Designs. *Pharmacopsychiatry 19*, 333-334. Stuttgart, Germany: Georg Thieme.

[30] Krauth, J. (1985). Meta-Analyse und Effektmessung in der Psychologie. [Metaanalysis and effect measurement in psychology]. In F.J. Hehl, V. Ebel and W. Ruch (Eds.) *Psychologische Diagnostik.* Kinder, Familie, Schule, Sport. (p-151-175). Deutscher Psychologen Verlag. Bonn, Germany.

[31] Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics,* 6, 65-70.

M. Heese was born in Graz (Austria) in 1982. After the grammar school with special emphasis on languages, M. Heese started the studies of Psychology with specialization in the field of Working-, Organisational- and Environmental Psychology at the University of Graz. Since November 2005 M. Heese holds the degree of a diploma psychologist.

She absolved an extended traineeship at the European Organisation for the Safety of Air Navigation (EUROCONTROL) in Brussels from July-December 2004 supervised by Dr. Manfred Barbarino followed by another four months at the EUROCONTROL CEATS Research-, Development- and Simulation Centre (CRDS) Budapest from February to May 2005 supervised by Mag. Renée Schuen-Medwed. From June-October 2005 she was participating in the Project '*The use of Red Bull Sugarfree® for the retention of central attention functions in complex control tasks on pilots*.' supervised by Dr. K.Wolfgang Kallus and Dr. Dagmar Deixelberger-Fritz..

Since April 2006 she is working as FEAST (First European Air Traffic Controller Selection Test) Service Assistant in the Human Factors Unit at EUROCONTROL Brussels.

Mag. Heese is the secretary of the Austrian Aviation Psychology Association (AAPA) and entitled Junior Aviation Psychologist with the European Association for Aviation Psychology (EAAP).

Impact of the Introduction of CPDLC on Controllers' Team Situation Awareness

Vincent Kapp, Nicole Boudes, Céline Classe

***Abstract*— Previous experiments involving Controller Pilot Data Link Communications (CPDLC) highlighted several issues:**

-- The risks of redundant actions between the executive controller and the planning controller,

--Defaults of mutual comprehension,

--Decrease of situation awareness that have a deep impact on the efficiency and the quality of the cooperation.

There are two factors that may contribute to jeopardize the current level of cooperation in a Data-Link environment:

--The fact that the data-link communications are asynchronous and silent and the loss of shared VHF communication associated,

--Modifications in the working methods, in particular in the communication task sharing.

Experimentation involving real time simulations aims at showing that data-link tends to support a "non manifest" assistance of the planning controller to the detriment of an active and explicit assistance.

***Index Terms*—Air Traffic Control, Team Situation Awareness, Cooperation, Controller Pilot Data Link Communication.**

I. Introduction

THE introduction of the CPDLC is often argued by the limits of the VHF (Very High Frequency): insufficient availability of radio channels, congestion, bad quality of the transmission, understanding problems, call sign confusion, etc. Therefore CPDLC is considered as a way to bring a better efficiency by allowing fast transfer of information.

On one hand, the selected addressing of the message and its unambiguous coding is expected to increase the reliability of the communication. But, on the other hand, former studies [1] showed that CPDLC generates side effects that cause qualitative changes, and modifies substantially the activity.

The new paradigm is that CPDLC can not be considered as a mere replacement of a way of communication by another. There is a permanent listening pattern on a controlling working position: actually the planning controller (PC) dedicates part of his/her resources to listening to the VHF. In the same way, the executive controller (EC) pays attention to the telephone calls, for coordination with the adjacent sectors.

Perceived information is used to create and maintain team situation awareness, i.e. the representation of the team mate activity as well as a common representation of the situations of traffic and pilots intentions. As CPDLC induces silent and asynchronous communications for the message conveyed by data-link, a risk of deterioration of the process previously described has been identified. In order to compensate this loss, the Controller Working Position (CWP) should be designed as a "common space of information" [3]. This issue motivated this preliminary study that focuses on the impact of the CPDLC on the shared representations of the controllers. This paper will firstly discuss the notion of shared representation in the ATC context and then focus on the experimentations that were conducted.

II. Shared representations

The combined activity between EC and PC obviously exceeds formal task attribution. Actually, the exchange of information (verbal or non verbal) contributes to build a common cognitive environment. The mutual understanding of the situation depends on the quantity and the quality of the information handled by the controllers, the working methods and the level of expertise of the operators. The party line, i.e. the perception of communications coming from other actors, contributes to the development of the team situation awareness [8]-[6] Loosing team situation awareness modifies consequently the diagrams of perception and the effectiveness of the co-operation between actors. The probability of misunderstandings is increased if this loss of information is not compensated to a certain extent.

The shared information is included in a common operative referential that refers to the information and know-how that are needed to complete a given task in common. In other words, all the processes linked to the cooperation between EC and PC strongly depend on the ability to share contextual information. This contextual information consists in different elements characterizing the situation: flight plans, history of actions, pilots' intentions, inter-sector coordination, tools specificities, etc.

The sharing of contextual information relies on mutual actions, artifacts, and communications:

A. Mutual Actions

The visibility of the actions performed by one of the

Manuscript received February 24, 2006. This work was supported in part by the SDER Department of DSNA.

Vincent Kapp is with the research center of DSNA, Toulouse, France (e-mail: kapp@ cena.fr).

Nicole Boudes is with SOFREAVIA, Toulouse, France (e-mail: boudesn@ tlse.sofreavia.fr).

Céline classe is with SOFREAVIA, Toulouse, France (e-mail: classec@tlse.sofreavia.fr).

operators, on the IHM or the strips bay, helps his team mate to understand the current situation and allows the anticipation of future actions.

In a non CPDLC environment, each instruction and pilot request can be heard on the radio. The monitoring of the radio also produces expectations allowing the coordination of future actions and the detection of potential errors. Shared representations (and consequently the level of cooperation) therefore depend on direct or peripheral auditory and visual perception.

B. *Artifacts*

Artifacts (basically the screen, the paper strips, the mouse, and so on) constitute a precious source of shared information: they are visually accessible and can be used by both controllers.

C. *Communication*

The verbal and non verbal communications (gestures, postures, looks, etc.), between EC and PC contribute to share contextual information. These exchanges, although they do not always directly aim at it, inform about the activity, the knowledge and the intention of the team mate.

In relation to that last item, the information shared among the controllers is also the result of the explicit co-operation between executive and planning controllers.

In accordance with the theoretical background, we consider that shared representations show the following features:

-- They are identical between both controllers on a given controlling working position.

--They are a subset of a controller's knowledge.

--They concern several elements, namely the current position of the traffic, the performed actions, the actions to do, the pilots intentions and the displayed data on the HMI.

--They are updated thanks to verbal and non verbal communication.

--They are also updated thanks to common tools, i.e. paper strips and the radar image.

Shared representations, as they are at the intersection of individual representations, do evolve: the recovery zone of those individual representations can vary according to the "quantity" of knowledge and information held at once by each controller.

The experiments allowed us to emphasize the gaps between individual representations, without judging or analyzing the "quality" of the representations.

The questions were the following:

-- Do the controllers share actions on the traffic if both controllers have access to CPDLC?

--Does CPDLC modify the quality of the individual representations on the traffic?

--Does CPDLC modify the shared representations among the controllers at a given time, compared to a scenario without CPDLC? Does the communication task sharing influence this modification?

III. THE EXPERIMENTAL DEVICE

In order to gather data for analysis, we defined the following experimental conditions:

1) Organizations with CPDLC were compared to a reference situation without CPDLC. The same traffic type was played in the same technical environment (cf. below) and with the same experimental constraints.
2) Two CPDLC organizations differing on the experimental conditions, i.e. the instructions given to the controllers, were compared. This point will be developed later in the paper.

A. *Simulation facilities*

The working environment consisted in a CWP including EC and PC function. The implemented CPDLC applications were: ACL (simple clearances and pilots requests), ACM (transfer of communications) plus dedicated tools for the VHF/CPDLC shifts management and for the display of the messages.

Some DAPs (Downlink Aircraft Parameters, i.e. the Indicated Air Speed, magnetic heading and rate of climb/descent) were displayed in order to limit the VHF communications.

During this experimentation, controllers were provided with a basic minimal CPDLC HMI: all the CPDLC messages are reproduced and available on the HMI of each controller, but neither support nor specific means were given to facilitate information sharing.

In order to control and limit the role of the artifacts in building the shared representations, observability of the actions were also limited: a panel was put between the two radar images, limiting the information taken from mutual actions and artifacts (see picture below).

Fig. 1. Observation of the controller's activity

B. *Experimental instructions*

The first generic instruction given to the controllers was to manage the traffic with maximum security, even though the environment was only conceived for experimental purpose.

A special instruction was not to talk spontaneously about the CPDLC actions. This instruction was inspired by former observations in Data-Link environments in which additional communications tended to compensate the loss of information. Controllers were also asked not to verbalize any item of their traffic analysis.

Two differentiated experimental instructions were proposed, based on the hypothesis that task sharing influences the construction of shared representations.

--Instruction A: EC manages the CPDLC communications. He/she can delegate to PC some tasks (assume, transfer, some clearances…) as a last resort.

--Instruction B: PC transfers autonomously the flights that are at the periphery of the sector (at his/her own discretion). PC assumes autonomously the flights. PC can send CPDLC clearances without informing the executive controller under the experimenter supervision. These "mute" clearances are spontaneous clearances or responses to pilot requests.

PC was given 6 red cards at the start of the simulation, for every mute clearance a card was considered to be played. It was the experimenter responsibility that all cards were played during the simulation session (by reminding the PC about the experimental instruction).

These highly specific experimental instructions had been defined with the cooperation of controllers during pre-experimental runs. The instruction B had been iteratively tested and refined during several simulation sessions.

IV. Methodology of Data Gathering

The experimental setting was the following: 45 minutes of simulation including an interruption break (that is to say a break with a freeze of the radar image), and a debriefing at the end of the simulation session. The complete session lasted 1h45 minimum.

The interruption had to be unpredictable for the controllers. During the break, controllers were separated, so that the information given by one could not be perceived by the other. The data gathering relative to the interruption was of two types:

--The controller gave a description of the situation; a methodical analysis of the state of each flight and the associated dialogues was performed. The analyst noted down the responses and their certainty levels (based on the controller's commentaries) on a grid.

--Once the grid was filled, a small individual debriefing was performed with the controllers in order to collect his/her current and future comprehension of the situation. This debriefing allowed to associate the context to the gathered traffic data.

The final debriefing purpose was to analyze the gaps between the controllers information on the basis of a set of selected situations relevant. In order to facilitate the confrontation of the individual data, the simulation, that had been previously completely recorded, was replayed and comments were directly addressed in reference to the situations displayed.

Observations of the activity were conducted for each simulation session. The objective was to detect any deteriorated situations for recall and comment during the final debriefing.

V. Results

The results can be divided into 3 parts: distribution of the actions, level of exactness of the responses and gaps between EC and PC. The analysis process was applied to 228 flights corresponding to 8137 pieces of data.

TABLE I
GENERAL DISTRIBUTION OF THE ACTIONS BEFORE INTERRUPTION

	VHF Only		CPDLC A		CPDLC B		Total
	EC	PC	EC	PC	EC	PC	
Assume	74	0	66	6	44	32	222
Transfer	25	0	28	1	21	9	84
Request	11	0	21	0	15	3	50
Clearance	62	0	81	1	87	16	247
Total	172	0	196	8	167	60	603

A. *Distribution of the actions*

The table below presents the results given by the recording of the performed actions on the system. For the simulation without CPDLC, there is a correspondence between the data input and the clearances that are given via VHF.

In CPDLC B, the PC is actually implied in the sending of clearances and responses to requests. In CPDLC A, the EC did not delegate CPDLC communications to the PC, although during the debriefing they said that the EC could have needed more support and would have benefited from a task delegation.

Two factors may explain this paradox:

1) The experimental environment (lack of communication, the hiding panel and CPDLC) did not allow the PC to anticipate the EC's needs
2) Explicit task delegation by giving all the necessary contextual elements to the PC would have added too much workload to the EC who preferred to perform all actions by himself/herself.

In CPDLC B, the PC did not go beyond the instruction of carrying out 6 clearances and/or responses to requests (average of 6,5).

Another interesting result is the fact that more clearances were issued in CPDLC B than in the other experimental situations despite the fact that the same exercises were played.

We observed that some flights that had been transferred by one of the controllers were transferred a second time by his/her team mate, more particularly in CPDLC B (situation were the communication tasks, especially the transfer tasks were shared among EC and PC). In those cases, the delays between the two transfer actions were greater than the delay of the transfer feed-back on the flight label. The controller who did the second transfer did probably not perceive that the flight had already been transferred by the other controller. The HMI data sharing shows here some of its limits

B. Exactness of the responses

1) *Response categories*

Each action (assumes, transfers, clearances…) recalled by the controller during the interruption is a set of several pieces of data, (that altogether are at the number of 8137):

--The nature of the action: presence or absence (response "yes", "no" or "I do not know").

--The value of the action: type and value of parameters, contain of the responses (Wilco,.Unable...).

--The one who did the action (PC, EC).

According to the information given by the controllers, we built three groups of responses: exact, incomplete and erroneous:

--The exact responses correspond to data given by the controllers that are perfectly in accordance with the electronic data recording.

--The incomplete responses correspond to data given by the controllers that do not cover all the information related to an action.

--The erroneous responses correspond to data given by the controller that are not in accordance with the electronic data recording.

TABLE II
PERCENTAGE OF EXACT RESPONSES FOR THE PERFORMED ACTIONS

	VHF Only		CPDLC A		CPDLC B	
	EC	CO	EC	PC	EC	PC
Assume	96	97.3	93	91.7	50	51.3
Transfer	84	56	72,4	48,3	40	40
Request's response	50	28,6	28,6	19	27,8	16,7
Clearance	59,7	19,3	44,3	15,2	37,2	29,4

The data given by the controllers, and their level of accuracy, are our indicators of individual representation of the flights at the moment of the interruption.

2) *Exact responses*

The data above show that the tendency is a general degradation of individual representations in CPDLC A and B, in comparison with the situation without CPDLC.

For comparable scenarios, it is therefore possible to assign this degradation to the fact that the controllers used a silent way of communication to exchange messages with the pilots. Three factors contributed to the deterioration:

a. Air/ground communications were performed via a «minimal» HMI, the information shared on both screen were limited to the status of the ongoing CPDLC dialogues, consultation of the dialogue detail was possible on specific action.
b. The experimental device constrained the verbal exchanges between the controllers, for instance they had to avoid to verbalize their actions.
c. The hiding panel restricted the possibility to get data from the team mate's screen.

These measures, as foreseen, prevented controllers from acquiring the air-ground information from other media: verbal communication between EC and PC and observation of the team mate's gestures. The fact that forms of cooperation compensations could not take place, allowed us to highlight the impact of the silent air ground communications.

A first interpretation of these results is on the level of the HMI's usability, which raises the following questions: are the displays intrusive enough and legible? Should some information be more prioritized like some CPDLC message content? Is some information missing like the degree of urgency associated to consultation or reply?

Another way of analyzing those results is to attribute the level of exactness of the controllers' responses to the functioning of the memory. Although the methodology of data gathering aimed at reducing the memory factor by various means (such as maps of the sector given as support, corrections of some minor mistakes) it is not possible to totally separate memorizing bias. Nevertheless, memorization and mental representation are tightly related. The controllers build their mental representation by putting in relation data that are stored in long term memory with data stored in immediate memory and with real time perceived data. Thus, deterioration in the capacity to recall information is revealing of the quality of the mental representations.

The recall accuracy for the clearances varied according to the controller's role:

--The EC meets especially difficulties in the CPDLC organizations, where he acts without talking. This suggests that the verbalization of a clearance improves its recall.

--The PC meets more difficulties in organizations where it is the EC that preferentially manages the communications. A difficulty to follow the course of the actions performed by the EC has been detected in CPDLC scenarios.

CPDLC B scenario, in which both PC and EC could use CPDLC, rose a new problematic: the controllers may not know anymore who did some of the action on the CWP (assume, transfer or clearances sending). This phenomenon shall still be analyzed in terms of operational consequences. This effect is more noticeable for the EC when recalling clearances that where made by the PC.

d. The gaps between EC and PC

3) *Gaps categories*

There are 4 types of gaps between the controllers for a given flight:

--No gap: the individual data collected on a given flight are strictly identical between the EC and PC. These data may be in accordance or not with what actually happened.

--Big gap: the individual data collected on a given flight are very different between the EC and PC and never complete.

TABLE III
DISTRIBUTION OF THE GAPS ACCORDING TO THE ORGANIZATION

Gaps	VHF	CPDLC A	CPDLC B	TOTAL
No Gap	13	12	10	35
Big	16	28	17	61
Little	18	12	30	60
In nature	14	13	7	34
Total	61	65	64	190

--Little gap: the individual data collected on a given flight are not very different between the EC and PC and never complete.

--Gap in nature: the number of individual data collected on a given flight is identical although different in nature.

i. CPDLC A and CPDLC B

The flights corresponding to big gaps between controllers were flights on which clearances were issued that were consulted, even though those clearances could not be correctly recalled.

ii. CPDLC A

This organization more frequently led to the big gaps. The EC is more often inexact than he/she was in the VHF scenario. He/she forgets to recall some of the actions or adds some that did not actually occur.

The number of little gaps is the lower.

When there is no gap, errors consist in the addition of clearances or requests as wrong identification of the originator of the action.

iii. CPDLC B

Among the flights with no gaps between individual representations, in this situation we found the less number of flights with complete and exact data.

The number of big gaps is the lower. The EC, in this organization, is more often incomplete than in the VHF or in CPDLC A organization, and also more than what the PC can be.

The number of little gaps is the most important, and the cases are more varied. In this group, the number of flights with complete data recalls increases clearly for the PC.

The proportion of gaps in nature is also the most important and the biggest difference is on the identification of the originator of the action.

2) DISCUSSION

a. Shared representation

In CPDLC A, the data collected from the debriefing strongly stressed the fact that both controllers were in a highly uncomfortable position, indeed dangerous. No controllers were satisfied with the work accomplished, nor with the traffic control conditions.

--The EC worked alone. At some times, he/she had troubles to remain ahead of the traffic. If the traffic had increased « it should have been necessary to split the sector ».

--The PC was passive and felt out of the loop quit quickly. It was practically impossible for him/her to properly understand and monitor the traffic. The PC did not feel at ease and showed a lack of interest about its activity.

The controllers' feelings were not linked to the CPDLC as a new tool. Indeed, the controllers showed a positive attitude towards this way of communication that has the potential to release resource and to limit the interruptions.

Thus, these data confirm that there are big gaps between controllers' individual mental representation, and show that the EC is less exact than in the VHF organization. In most cases, a controller recalls the maximum number of data on a flight whereas his/her team mate can only recall one assume and/or a transfer.

In CPDLC B, some PCs felt to rather succeed in following the traffic, although it demanded a great degree of resources, whereas others among them stressed their uncertainty about the actions executed by the EC. Even though they doubt to have perceived everything, half of them have the feeling to have followed the traffic more accurately than in the CPDLC A organization.

b. The cooperation

In CPDLC A, the results showed that no cooperation took place. We noticed that controllers held only few data that were common and exact for the same flights and it is relevant to think that they were aware of the gaps of their individual representations.

In CPDLC B, the EC had indeed no specific expectation towards the PC. Only one team agreed prior the simulation on a sort of minimal delegation plan on transfers and assumes for some flights. Thus, the ECs expressed the fact that globally the PC's actions did not bother their own activity. These results are consistent with the fact that there are less big gaps in this organization and consequently more little gaps. Indeed, the more the representations are shared and the more the PC has the capability to integrate his/her actions in the EC's activity, with security and effectiveness.

The PC generally performed actions that were not disturbing for the EC and « that were supposedly convenient ». For this reason, they usually limited themselves to direct clearances for flights entering into the sector, or in initial descent. On the other hand, the PC responded rarely to the pilots' requests and never issued clearances conveying a heading, statement that has been confirmed and described during the debriefing as "non desirable delegation".

Nevertheless, we have examples that show that the PC sometimes initiated actions that did not bother the actions already performed by the EC, but caused trouble the EC's plan of actions. Consequently, in those cases the EC had to change his/her strategy. Furthermore, cases of double actions on a same flight had been noticed.

In any case, it is not possible to consider that the PC could use the CPDLC without a good working method including a form of organized tasks sharing that would have been validated.

c. The Follow-up of the PC's actions

In CPDLC B: the observation of PC's actions, when in contradiction with the EC's plan of actions, show the consequences linked to the fact that the HMI, designed as a « minimal » HMI (used in this experimentation), did not allow to infer on the EC's intentions. These situations may be potentially hazardous.

The EC had no means to anticipate the actions of the PC. The manifestations of those actions appeared rather «suddenly» on the flight labels, forcing the EC to make the corresponding consultations. On the other hand, despite the coding of the CPDLC, the PC had no means to anticipate the future data/activity although the consultation might be considered sometimes less urgent. Indeed it is the lack of data on the urgency of the consultation that brings discomfort, mainly by the fact that the coding of a transfer or a clearance is exactly the same.

The possibility of not seeing the actions of the other controller was mentioned by the controllers as a risk for the security.

3) CONCLUSION

a. Synthesis

The results of the analysis converge towards the idea that the introduction of the CPDLC on a CWP may induce a damage of the individual representations, especially for the PC, and a divergence of the shared representations on a given flight

This deterioration (of individual and shared representations) differs according to the way the CPDLC is used within the CWP and more generally to the communication task sharing that will be foreseen. It is consistent to distinguish the case where it is the EC that has the whole responsibility of the CPDLC dialogues from the case where the PC may also decide and send CPDLC messages autonomously.

b. Limits

Globally we have reservations about any extrapolation of those experimental data in an operational environment. The data collected among the controllers at the time of the break are not necessarily representative of the ones that would be collected in a real world.

Nevertheless, the use of these data for comparative purpose between different experimental organizations has been completely consistent.

c. Perspectives

Several perspectives for further studies can be considered.

The most promising perspective concerns the design of a cooperative environment for the CPDLC, especially on the HMIs level. A «standard» HMI that only displays occurrences of CPDLC dialogues proved not to be sufficient to maintain an acceptable level of shared representations between EC and PC nor sufficient individual representations [2].

The availability of CPDLC functions for both controllers may be effective on the condition that the HMI allows an accurate understanding about the ground-air dialogues. Not only is it a matter of making the dialogues more perceptible, but also of conveying more meaning, allowing the controllers to have a better understanding of the intentions and building more easily consent on their mutual actions.

Although the HMI will not be used without mutual communications on the CWP, the language and the visual support should be complementary.

It shall be necessary therefore to work on the concept of a common workspace supporting mutual awareness, i.e. the mutual perception and comprehension of the controllers' activity.

In conjunction, further studies on the controllers needs shall be conducted. The results of this study are currently more deeply analyzed with the assistance of controllers in order to highlight events that could have generated incidents.

REFERENCES

[1] Expérimentation CAMELIA - Rapport d'analyse ergonomique côté sol, CENA, Toulouse, France, R98_005, Oct. 1998
[2] M. Beaudoin-Lafon, (1999). Computer Supported Co-operative Work. Available : http://insitu.lri.fr/~mbl/Trends-CSCW/fulltext.html
[3] P Salembier. M. Zouinar, L. Rognin, "Cooperation, reliability of socio-technical systems and allocation of function ", International Journal of Human-computer Studies, 52(2), 357-379, 2000
[4] Bolstad, M. R. Endsley, "The effect of task load and shared displays on team situation awareness", in *Proc. 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA: HFES, 2000
[5] Eurocontrol, "Operational Requirements for Air Traffic Management (ATM) Air/Ground Data Communications Services", Edition 1.0, Dec 2004. Available : http://www.eurocontrol.be/projects/eatchip/odiac/
[6] Eurocontrol, "The development of Situational Awareness measures in ATM systems", Edition 1.0, Jun. 2003.
[7] M. Zouinar, "Contribution à l'étude de la coopération et du partage d'informations contextuelles dans les environnements de travail complexes", Ph.D. dissertation, Université Paul Sabatier,Toulouse, France, Jun. 2000.
[8] S. Dusire, "Naviguer dans un espace verbal : la construction de la conscience de la situation", Ph.D. dissertation, CNAM, Paris, France, 2000.

AUTHORS' BIOGRAPHIES

Vincent Kapp has an engineering degree in aeronautics. Since he joined the research center of DSNA (Direction des Services de la Navigation Aérienne, formerly DNA), he is working in a team responsible for ATC En Route HMI evaluations and was in charge of the DAP En Route study.

Nicole BOUDES received a Ph.D. in Ergonomics from Toulouse University, where she studied cognitive ergonomics, psychology and linguistics. She has co-authored a number of scientific papers on cognitive activity in Air Traffic Control, and evaluations of ATC performance in new environments. She is currently in charge of the SOFREAVIA team working for Human Factors projects at the research center of DSNA, which main ones are CPDLC, current French En Route CWP, and a European project.

Céline CLASSE is an Ergonomist. She has a degree in psychology and ergonomics from Toulouse University. She has gained practical experience through industrial projects and telecommunication projects. She is presently employed by SOFREAVIA were she mainly works on ATC projects including CPDLC project at the research center of DSNA. She is a specialist in the ergonomic approach on operational project, for the design, the specification and the evaluation of Humans-Machines Interface

Investigation of Time Critical Concept and Impact in Mixed Voice/Data Link Environment

Tanja Luchkova (1, 2)

Abstract—**The increment of air traffic in the recent years is leading to a point where ATC system is becoming increasingly overloaded. For that reason, there is a need to change from existing ATM (Air Traffic Management) System towards new forms of ATM, which have the potential to provide the required capacity for the long term future, whilst maintaining or increasing safety.**

Data link communication could be seen as an enabling technology to achieve this. This kind of communication has been used to relieve the great number of deficiency of voice communication that usually appear, such as channel congestion, misunderstandings between pilots and Air Traffic Controllers (ATCO) and corruption of the signal due to simultaneous transmissions.

Data link will also allow down linking from an aircraft of various preferences and parameters. Aircraft's preferred flight level or top of decent can be down linked, meteorological conditions can be reported (wind direction and speed, turbulence, air temperature and pressure etc.). Carrying into effect this type of data communication will importantly change the way pilots and ATCOs communicate.

The goal of this study is to define a time criticality that appears in case of potential conflict (i) and to define a time needed for decision making in switching from Data link to voice channel (ii).

This paper describes the preliminary results from the interviews with ATCOs that have been made during the last Data link simulation in EUROCONTROL CEATS Research, Development and Simulation Centre and mathematical model used to achieve above mentioned objectives of the project.

Index Terms - Data-link, time critical situation, mathematical model, decision making.

I. Introduction

THE EUROCONTROL Link 2000+ Programme is coordinating the implementation of Controller – Pilot Data link Communications (CPDLC) in the core area of Europe. CPDLC is a pre – canned message – based communication system between pilots and controllers and has much the same advantages over voice radio communication system as e – mail has over the use of the telephone.

CPDLC is only one of the possible ways to exploit data link communications between pilots and controllers. It is the first step and as such requires more effort to be implemented, as the basic infrastructure; computers and software have to be installed on board of the aircraft. Today's air traffic control radio frequencies are often saturated. This leads to the need to repeat messages, thus wasting controller's valuable thinking time. This saturation means that controllers might be unable to deliver timely instructions that would provide optimal flight profiles for all the aircraft flying in their sectors. At the same time, pilots are prevented from making requests that would ensure a more efficient flight.

CPDLC can relieve the workload by providing silent, easily readable messages to the correct recipient. It cannot replace all voice communications, because for the moment it is slower than the radio. It can, however, take care of all routine, non – time critical communications [1].

So far in the ATC Data Link Manual for Link 2000+ services there has been mentioned only one principle concerning time critical situation. Principle B: CPDLC (Controller/Pilot Data Link Communication) shall only be used in the context of non – time critical situations. Time – criticality is determined by the following factors: ATC traffic situation, end – to – end performance (systems and flight crew/controller response time) and recovery time [2].

Using both, voice channel communication and Data Link communication raises different issues, such as: operational procedures, phraseology, and task sharing in mixed environment. It's a completely new environment for the ATCOs using a new technology comparing with the one they are using for the time being.

In this mixed environment there are certain points where a time critical situation might appear. This time criticality can be caused by different factors.

Until now only few studies have been made raising this question. Those are: Vocalise, which is a project of CENA and it's based on 60 hours of French En – Route traffic, spread over 6 types of sectors in a situation of heavy traffic. It's a comparison between the occupations in voice environment versus Data Link environment [3]. Another study is the FAA study report (Controller Evaluation of Initial Data

[1] Manuscript received March 2, 2006.

This work is funded by the CEATS Research, Development and Simulation Centre (CRDS), Budapest, Hungary.

Tanja Luchkova is an undergraduate student attending the final year at Civil Aviation Engineering at the Technical Faculty in Bitola. Currently she is a trainee at EUROCONTROL CEATS Research, Development and Simulation Centre (CRDS), Hungary. Address: Technical Faculty in Bitola, Ivo Lola Ribar b.b., 7000 Bitola, Macedonia; EUROCONTROLCRDS Ferihegy1, "A" Porta, H-1185, Budapest, Hungary..E-mail:luchkova@yahoo.com

Link Air Traffic Control Services). The objectives of this study are: (1) to evaluate and refine DL controller procedures and displays for the Altitude Confirmation, Transfer of Communication and En route Min Safe Altitude Warning services and (2) to solicit initial opinions from controllers regarding the general utility of the Mode S Data Link [4].

The question that has been raised in my study is the question about what time critical situation means. What kind of time critical situation appears during the usage of voice and Data Link communication?

In mixed environment ATCOs are using Data link communication and at certain point they need to switch to voice communication, due to the appearance of some kind of time critical situation. Time critical situation can be caused by many factors, such as: technical failure, operational failure or in case of potential conflict. Another question that has been raised in this study is when should ATCOs revert to voice channel?

If a time critical situation appears, caused by technical or operational failure, then the controller reverts to voice immediately without any questioning. But, in a case of potential conflict, there are many factors that influence controllers decision when to revert to voice. Next section will briefly describe the approach of the study.

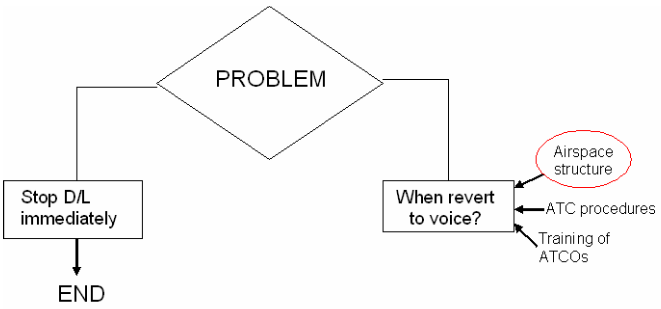


Fig. 1. Time criticality evaluation

II. Parameters

Based on the interviews that were made with ATCOs who participated in the previous Data link simulation and the literature study that has been made concerning Data Link project, the most important dimensions were discovered. Those are five dimensions that influence the decision making when to revert from Data Link to voice channel:

1) *Airspace structure*
2) *Training of ATCOs*
3) *Technical failure*
4) *Operational failure*
5) *ATC procedures*

The first dimension is consisted of few important characters: sector shape, sector size, traffic level and vertical movements. The shape of each sector is of great importance, according to the entering and exiting point of the aircraft in each sector. If the shape of the sector is such that the over flying time of an aircraft is very short, than the reversion to voice in case of potential conflict would be earlier. Sector's size results in the same way. Bigger sector – longer over flying time and later revert to voice, smaller sector – shorter over flying time and earlier revert to voice.

Traffic level and vertical movements in each sector as a parameters are of greatest importance to most of the ATCOs. Complexity of each sector was expressed by the number of crossing points inside the sector.

The second dimension, training of ATCOs is not taken in account in this study, but it's an assumption from the interviews that it has a big influence. Due to this fact it can be considered in further study.

All the ATC procedures can be found in the ATC Data Link Manual for Link 2000+ services, version 3.0. It should be noted that for some of the ATC procedures formal endorsement is still required from ICAO – Air Navigation Committee.

A. Approach

This January a Data Link simulation was held in CEATS Research, Development and Simulation Centre. One of the objectives of this simulation was to validate procedures to be used for reverting to voice communication according to LINK 2000+ manual using the new RTF phraseology from the point of view operational acceptability and usability, clarity, correctness and consigns. The requirement was that revert to voice message must be simulated for the following cases: inappropriate messages, time constraint, correct unintended data link clearance and error generated by the system. During this simulation, as a methodology of my hypothesis, a questionnaire and an interview were made with the ATCOs. The questionnaire consists of six questions which where filled by each controller after completing the exercise during the simulation. The idea is to receive their opinion and knowledge and to confirm the assumptions about the four above mentioned parameters in the environment they worked. They gave explanation about sectors D7 and D9 (described later in the paper) in which they worked with different traffic level together with the decision when to switch to voice channel in case of potential conflict. In the questionnaire they gave a value (from the lowest to biggest importance) to each of four above mentioned parameters.

The idea of the whole project is to give a relative value taking into account the four parameters. This value would give an advice to the controller when to switch from data-link to voice. It should result in a formula which will be tested and validated in the ten most significant CEATS sectors. Afterwards it can be applied in any sector where the Data-link is applied to predict the approximate time for switching from Data link to voice channel in case of potential conflict.

As a second part of this kind of methodology, an interview was made with ATCOs as mentioned above. Ten sectors were given to ATCOs, on which they had to answer one question, taking in account three different situations. Each controller had the opportunity to see all those ten sectors with the sector complexity inside (number of crossing points). The first situation takes into account fair traffic level and changing of the vertical movements from fair to high number. The second situation takes into account medium traffic level and the third one, high traffic level and progress of the vertical movements from fair to high number.

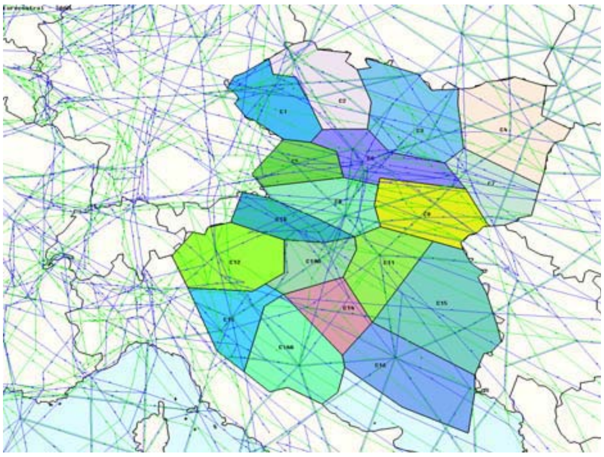


Fig. 2. CEATS Sectorisation

Sectors that have been used in this study are CEATS sectors: D10, D9, D8, D7, D3E, D3SW, D3NW, D5, D1 and D11. Letter D stands for a last version of the sectors in CEATS Airspace.

B. Preliminary results

For each sector and each of the three situations the average has been calculated representing the controllers' answers. Twelve controllers participated in the interview as well answering the questionnaire. As it can be seen in the table above shown as an example, there was one question in three different situations, with three values for traffic level (L) and three values for vertical movements (V). That's exactly nine values for L and V. Taking into account that these parameters were investigated in ten sectors, 90 answers were received. ATCOs answered on the question in each situation giving a value from 0 to 1 (0-revert to voice immediately before loosing separation between the aircrafts; 1 revert to voice long time ahead before the potential conflict appears.

TABLE I
FIRST SITUATION IN SECTOR D10

ATCO's	question1	question2	question3
C1	0	0,5	1
C2	0,2	0,5	0,7
C3	0,5	0,7	0,7
C4	0,2	0,4	0,5
C5	0,2	0,4	0,6
C6	0,2	0,2	0,3
C7	0,2	0,4	0,7
C8	0,3	0,4	0,5
C9	0,2	0,2	0,4
C10	0,2	0,2	0,3
C11	0,4	0,4	0,6
C12	0,5	0,6	0,7
Average	0,27	0,52	0,57

A lot of data was collected during these interviews. First initial raw results confirm the assumptions that the four parameters: S, C, L and V are most important in this study. The table above shows preliminary results of the interview which was made to evaluate one of the parameters. In this first situation it is given a low traffic level and the vertical movements are changing from fair to high. Then an average was made for each situation and each value of the parameters L and V. As the table shows the assumptions were verified and controller's answers are in between the values that were given. It was also noticed that there are some recoils in their answers. This will be taken into consideration for further studies.

The second part of the study deals with a linear regression. The average values for all sectors as explained above will be imported into a formula which will be tested in ten sectors. The result would be a relative value of time giving an advice to the controllers when would be the most appropriate time to revert to voice according to all the information they have for each sector. In the next paragraph this mathematical model is explained. The plan is to develop the model until the end of this project.

C. Mathematical model

All the results from the interviews will be transformed into appropriate values and together with all the other parameters a mathematical model will be created. Each sector has its own shape, complexity, traffic level and vertical movements inside. According to that, four dimensions were included in the mathematical model giving a function:

$$T = T(S, C, L, V) \qquad (1)$$

T is a non dimensional value in a range between 0 and 1. It gives the time criticality that needs to be discovered, S is shape of the sector, *C* is complexity of sector including crossing points inside, *L* is traffic level inside the sector which is a value that is already known for a certain sector and *V* is vertical movements.

It is important to mention that the influence that the four variables (S, C, L and V) make on T it might not be linear. If T has a bigger value and it's closer to 1, ATCOs should revert to voice later. In the contrary, if T has smaller value and its closer to 0, ATCOs revert to voice earlier.

This approach is based on defining and calculating each parameter independently. During the work it has been attained that first two parameters, sector's shape and complexity, are constant values for one sector and traffic level and vertical movements are changeable variables.

Sector's shape can be calculated in different ways. The most appropriate one chosen for this study is to calculate the area of a polygon with the dimension of square kilometres. This number is then divided by the number of sector's corners. After calculating the area of each sector, next step is to define the value S:

$$S = \frac{A}{N} \tag{2}$$

A is the area of the sector in square kilometres and *N* is the number of corners in each sector. All the information for the coordinates of the corners were taken from RAMS Plus ™ (the Reorganized ATC Mathematical Simulator) and updated from the last version of CEATS sectors. It has been noticed that this variable has a linear influence on T. With the increment of S, T increases as well.

Second parameter in the formula is *C* – complexity of a sector expressed in numbers of crossing points in the sector. Complexity of each sector varies from one point to another and it is not connected to the shape or size of a sector. The complexity parameter will be expressed as well with mathematical formula. Initially the calculation of the distances between each crossing point in sector have been made and then divided by the number of crossing points inside. After calculating the crossing points and their position in the sector, their distance is measured and divided with the number of crossing points inside. The hypothesis is that this number can be after divided with the longest air way inside the sector in order appropriate value to be received. But the preliminary results show that this variable has also a linear influence on T. With the increment on C, it is expected that the value of T will increase as well. This part is still under investigation.

The hypothesis is that when one sector is taken into consideration, its shape and complexity are considered as constants. They don't change their value during the variations of the other two dimensions L and V.

Traffic level and vertical movements are considered as variables as mentioned before. They can be changed according to the number of aircraft over flying the sector. Both of the dimensions can vary for different situations. Another assumption has been made concerning these two dimensions during the interviews. Values defined as fair, medium and high traffic level and vertical movements were given in the interview, due to the assumption that ATCOs will take into consideration same values.

In this case L and V do not influence linear on T. Due to that fact, they might be taken as L^{-1} and V^{-1}. It's also important to mention that L and V are dimensions that are taken into account at the time when the potential conflict appears.

III. CONCLUSSION

Today's air traffic control radio frequencies are often saturated. This is partly due to the high traffic figures, but also partly due to a large number of misunderstandings and simultaneous transmissions on the voice channel. In order to minimize the voice communication and to prevent potential conflicts that might appear as result of conflicts, CPDLC has been applied.

This study explains the linear regression and the approach and methodology dealing with examining time criticality to switch from data-link to voice channel. The mathematical formula explained in the paper consists of both: constants and variables. Constants in the formula are: shape and complexity of a sector and variables are: traffic level and vertical movements. The methodology for gathering the relevant information were questionnaires and interviews with ATCOs participating on simulation taking place in CRDS, Budapest. Next step is to evaluate and validate this formula for certain sectors.

IV. FUTURE WORK

Due to time shortage this study will be finished with creating and evaluating a mathematical formula about time critical situation with respect to the selected parameter of airspace structure. Further studies should consider also other parameters that influence this question such as: training of controllers, operational failure, and technical failure. The approach could be to deal first with each parameter independently and later investigate to interdependences and the influence on time criticality as one factor for more compatible results. Also as a future work can be seen evaluating the other parameter - training of ATCOs, because during the interviews were noticed differences in their answers. These differences can be caused by many factors and it is seen as an opportunity for further analyses and investigation, which might influence on T.

In order to validate the formula for time critical situation, an indicator can be made and tested during the Real – Time Simulation (RTS) in EUROCONTROL CRDS, Budapest. This indicator will be warning to ATCOs when they are close to a time critical situation. This also can be seen as an future work.

ACKNOWLEDGMENT

The author is grateful to the discussions, and knowledge transfer, with all the personnel of CRDS Budapest, Hungary. Especially to Philippe Debels, Sophie Dusire, Renee Schuen – Medwed, Lenka Dravecka for their advices, time and continuous help.

REFERENCES

[1] SSRTS Data Link 2000+, EUROCONTROL CRDS Budapest, 2005
[2] ATC Data Link Manual for Link 2000+ services, EUROCONTROL, Brussels 2005
[3] [3] VOCALISE'00 Centre d'etudes de la navigation aeriennne, CENA 2000
[4] N. J. Talotta "Controller Evaluation of Initial Data link Air Traffic Control Services", Natioanl Technical Information Service, Springfield, Virginia

[5] SSRTS5 Data Link Simulation, EUROCONTROL CRDS, Budapest 2004

Tanja Luchkova is undergraduate student at EUROCONTROL CEATS Research, Development and Simulation Centre (CRDS), Hungary, under the supervision of Stefan Oze. She is attending the final year at Civil Aviation Engineering at the Technical Faculty in Bitola. Address: Technical Faculty in Bitola, Ivo Lola Ribar b.b., 7000 Bitola, Macedonia; EUROCONTROL CRDS, Ferihegy 1, "A" Porta, H-1185, Budapest, Hungary. E-mail: luchkova@yahoo.com

"Did you give the clearance? No! You?" The influence of the working method on controller performance in Data link Environment

Renée Schuen-Medwed[1]

Abstract— **A data link simulation (SSRTS 5) was conducted at the CRDS (CEATS Research and Development Simulation Centre) Budapest in 2004. The traffic in central Europe has increased dramatically over the last years. Data link can be seen as a possibility to increase capacity by offloading the voice channel. Indeed, one of the bottlenecks of a sector's capacity using the voice communication channel in certain traffic conditions is the Air Traffic Controller's working capacity. The overall results of the simulation are summarized in [6]. This study consisted of two phases. As no working method for data link environment did exist so far the first phase aimed at developing a working method and the second phase aimed at assessing the impact of the working method on controller's performance in a data link environment. The working method describes in detail how the controllers should work as a team together using the possibilities of Data link communication. A multi-step approach was used to develop the working method that was accepted by all controllers. In the second phase different levels of Data link equipage and different levels of traffic load were simulated in order to validate the working method. The results outline the final working method that was considered as operable and acceptable by all controllers. The operability was supported by a good situational awareness created by the working method. Situations with 50% of Data link equipage resulted in the highest workload in the high traffic sample. The higher workload was recorded on the Executive Controller's (EC) working position. Those results indicate that the working method for Data link environment has still to be improved in order to better balance the workload between Executive and Planning Controller (PC).**

Index Terms— **Air traffic control, Human factors, Controller-Pilot Data-link communication, Working method model**

I. Introduction

In recent times, air traffic has increased to the extent that Air Traffic Control (ATC) systems have become increasingly overloaded. Saturation of the voice communication channel is now recognized as one of the major bottleneck in a sector's capacity. Porterfield 1997 [7] showed a high correlation (0.88) between time spent communicating and a controller's subjective assessment of workload.

[1] CRDS-EUROCONTROL, Budapest, Hungary

For the reason of the benefit to switch from voice communication to data link communication the use of controller pilot data link communication is not just only a discussion but a topic for planned implementation.

Several projects such as EATCHIP IIIb (European Air Traffic Control Harmonisation and Integration Programme), EOLIA (European Pre-Operational Data link Applications), PETAL (Preliminary Eurocontrol Trial on Air/ground Data link), DOVE (Data link Operational Validation Experiments) and LINK Programme [4] developed and tested Data link and evaluated the potential benefits which might be achieved by that technology in Air Traffic Control. The objective of the LINK 2000+ Programme is to plan and co-ordinate the implementation of operational air/ground data link services for Air Traffic Management (ATM) in the core area of Europe in the timeframe 2000 – 2007. So far none of the studies was concentrating on the development of a working method for data link environment. Therefore this study focuses on developing and validating a working method to be applied in Data link environment.

II. Introduction of New Technologies

Introduction of new technologies for which performance does not conform to controllers' expectations, although unavoidable, may therefore have an unanticipated but fundamental impact on controllers' working method, strategies, and performance (Hopkin, 1995 [3]; Wickens, Mavor, Parasuraman, & Mc Gee, 1998 [9]). Implementation of data link communications will significantly change the way pilots and ATCOs communicate and therefore the effect on operations must be carefully studied before deciding the extent to which voice can be replaced or supplemented by Data link. To carry out their tasks, controller rely on a mental picture comprising static information of the airspace layout, the rules and standard procedures regulating the conduct of flight, aircraft flight plan information and performance characteristics, and the dynamic traffic situation changing from moment to moment (Rantanen 1994) [4]. Wickens (1998) [9] stressed the importance of shared situational awareness and mental models.

The main difference - that is of interest for this study - between voice environment and Data link environment is that in voice environment only one controller in a sector has the possibility to communicate with the air crew and in Data link environment both controllers on a sector have the possibility of communication. The experience from previous simulations indicated a general lower workload for the PC than for the EC. This naturally posed the question if there is any possibility to delegate responsibilities of communication and control from the controller using voice communication, the EC, to the other controller, the PC. Hence transferring some of the communication tasks, using Data link, to the PC will contribute to a more balanced workload. With giving the opportunity to the PC to perform tasks to off load the EC gives the possibility to balance he workload between EC and PC more.

The theoretical possibility of this delegation raises the need of defined procedures and rules concerning the use of Data link communication.

The two communication means behave technically differently. Controllers have to become accustomed with the present technological limitations of Data link. According to the technology used in air and ground, the delivery time of an ATC instruction can be different for each Data link equipped aircraft. A typical value of delivery time of a Data link message is between 2-12 s. Any new function, such as Data link, must be evaluated with human-in-the-loop simulations. The definition of an appropriate working method related to that function and tool is a pre-requisite to a reliable evaluation of the benefits and of the impact (e.g. safety) into the ATCO's environment. Therefore a new working method was developed and assessed in a Real-Time simulation.

III. METHOD

A. Subjects

12 Controllers from 8 different air service providers (Austria, Croatia, Czech Republic, Maastricht Upper Area Centre, Hungary, Italy, Slovakia, and Slovenia) participated in 42 exercises within three weeks. One exercise lasted one hour and 20 minutes. The first ten minutes were used to build up the traffic in the sector and the last 10 minutes were used to run down the exercise.

B. "Modelling" of the working method

A working method is defined as the description of the way controllers are working according to their position. Therefore the tasks the PC and the EC have to execute in a Data link environment are described explicitly.

The working method was build with the help of a hierarchical model that describes each task of each controller in detail with reference to the state of the aircraft. A base line of the working method was provided to the controllers. The "modelling of the working method" was an interactive process. The controllers were working with this baseline and after some experience this baseline was modified with the help of a multidisciplinary expert group. The development of the working method was a prerequisite for the validity of the data and the results.

This approach can be summarized as an approach in two steps:

1) *Create a baseline about which the controllers would have discussed during the simulation instead of starting "from scratch". The initial working method was not supposed to be a "best one" but a working document to support the discussions and interviews during the simulation.*
2) *Identify risks or weaknesses to be solved or mitigated later, e.g. by accentuating the effort in improving some parts of the working method.*

C. Independent variable

According to the objective of the simulation two experimental variables have been manipulated during the simulation, which are relevant for the operability of the concept:

1) *Traffic load (100% 120%) and*
2) *Data link equipage (25%, 50%).*

D. Dependent variable

The Instantaneous Self Assessment (ISA) method was used to assess controller workload. The participants are asked to respond to a prompt every two minutes by pressing one of five buttons appropriate to their perceived workload at that moment: Very Low, Low, Fair, High and Very High

Situational Awareness (SA) is a frequently discussed concept in Air Traffic Control. It is often defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" [1]. In this study the investigation of SA is based on SASHA (Situational Awareness for Solution for Human-Automation Partnerships in European ATM) [2].

Controller's co-ordination is defined in this study as the flow and quality of the internal EC – PC communication. Coordination includes in this study also the direction of the communication flow.

Controller's task performance was recorded via inputs that the controller made on the Human-Machine Interface (HMI).

The study was defined in a way that some of the exercises were used to give the controller the possibility to try the working method and modify the working method and some exercises were simulated with the final defined working method.

This modification ended in having the "type of working method" as an additional independent variable.

Indeed, three different working methods (WM) were simulated.

WM 1 (the proposed working method) was outlining that all Data link aircraft are handled by the PC.

WM 2 (conventional working method) was outlining that the EC is performing all clearances –whatever type of clearance- to the a/c, whereas the PC is planning the flow of the incoming traffic into a sector.

WM 3 (task sharing light) outlined that the PC is able to perform Data link clearances under certain circumstances.

The differences in the different working methods can be displayed as follows.

	EC	PC
WM1	Handles clearances to a/c without data link equipage.	Handles clearances to all data link a/c, fulfils the conventional planning tasks*
WM2	Handles clearances to all a/c independent of the communication means	Fulfils conventional planning tasks
WM3	Handles data link clearances and clearances to a/c without data link equipage (see description in result part)	Handles data link clearances under certain conditions (see **Fel! Hittar inte referenskälla.**)

Table 1: Difference in Working Methods. * Conventional planning tasks are not further defined as they are not main interest of the study

Interviews and observations were conducted with the controllers after working on a working position with 120% of traffic and, 25% and 50% of Data link equipage applying the final working method. The interviews were analysed according to Mayring [5].

IV. Results

The presentation of the results is divided into two phases as each phase answers a different question.

A. Phase I

The multiple step approach of defining the working method was very much appreciated by the controllers. The way how the working method was defined was a reason for the commitment of the controllers and therefore for a final, acceptable version of the working method. The key points of the working method accepted and agreed at the end of the simulation can be described as follows:

- -Responsibility stays with the EC.
- -The EC is assuming all the a/c
- -The EC can delegate to the PC (no delegation of conflicts)
- -The PC can transfer and auto-transfer Data link a/c
- -The PC can communicate with the EC for whatever reason (i.e. asking for delegation, informing the EC)
- -The PC can silently (without informing the EC) issue Data link clearances if the following condition is met:

The clearance does not affect the traffic configuration,

The information concerning the clearance is "not useful" for the EC.

Examples are:

Routine clearances; Clearances depending on sector constraints.

Table 2: Description of WM 3

Following graph shows the distribution of the tasks "transfer" and "assume":

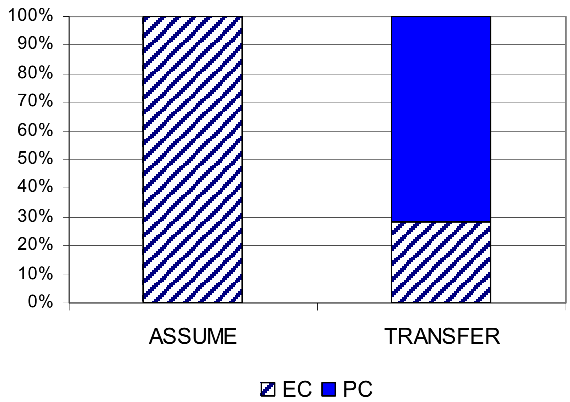


Figure 1: Frequencies of Tasks

This graph shows one aspect of the working method and confirms that the controllers were following the working method and that the EC was assuming all aircraft and the PC was mainly transferring the Data link aircraft as proposed in WM 3.

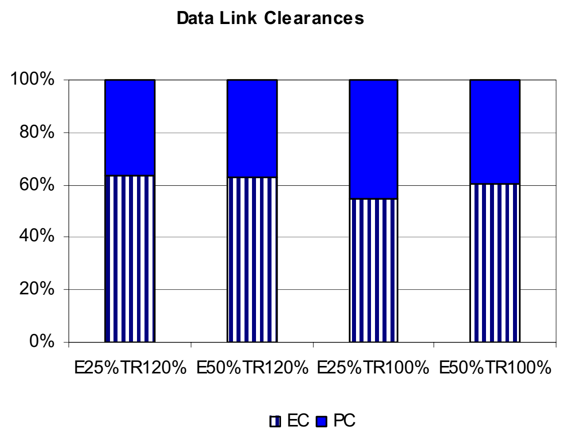


Figure 2: Data link clearances given by EC and PC.

Figure 2 shows that depending on the organisation the PC was performing between 40% and 50% of the Data link clearances when applying WM3.

The question: "Which working method is best supporting the use of Data link by reducing and balancing the workload of the EC and the PC?" was investigated by recording the workload of the controllers applying all three working methods.

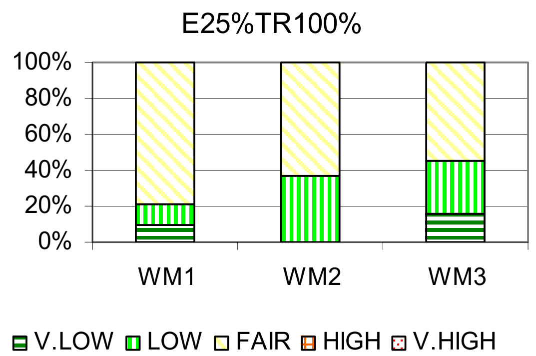


Figure 3: Sector workload with 100% traffic and 25% of data link equipage

In the scenario with 25% Data link equipage and 100% of traffic load no high or very high workload was recorded.

A one way ANOVA performed on the workload showed an effect approaching significance F (2, 147) =2.800, p=0.64.

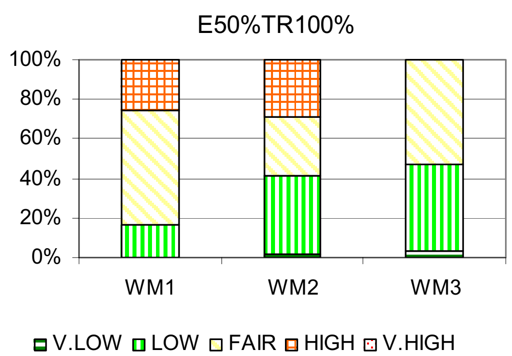


Figure 4: Sector workload with 100% traffic and 50% of data link equipage

In the scenario with 50% of Data link equipage and 100% of traffic load, up to 30% of the controllers assessed the work load as "high" when they were applying the proposed (WM 1) and the conventional working method (WM 2). The agreed working method was the working method with the least recorded workload, WM 3.

A one way ANOVA performed on the workload showed a main effect, F (2, 159) =7.936, p=0.001.

The post hoc comparison between the agreed working method (WM3) with the conventional working method (WM2) as well as with the proposed working method (WM1) were significant under the condition of 50% of data link equipage and 100% of traffic load.

B. Phase II

The question of the "Influence of Data link on controller's performance" was investigated in step II. The final working method, WM 3, was accepted by all controllers and further results of the study are based on this working method. The graph below shows that the working method was accepted by 100% of the controllers. Only 16% of the total considered that the acceptability was dependent on some circumstances such as traffic load.

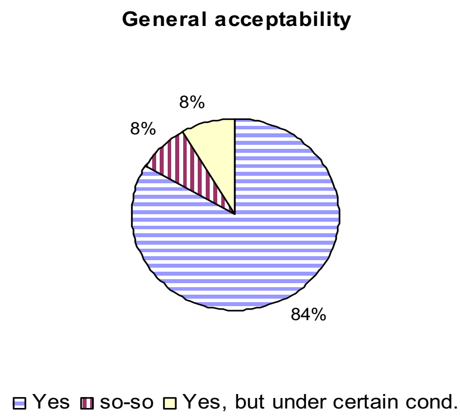


Figure 5: Acceptance of the working method

Comparing WM 3 with WM1 and WM3 controllers mentioned in the interviews following points for their acceptance of WM3.

- PC can perform his planning task and if needed he/she can help the EC
- EC is released from transferring the aircraft
- Controllers experienced a more balanced workload
- Controllers experienced more internal communication
- Controllers experienced a better situational awareness
- Controllers experienced a better team work.
- Operability of the working Method in the mixed Data link environment

Factors which are an indication for the level of operability are the impact on individual working activity like workload and the impact on the collective working activity like communication.

1) Controller's workload

The assumption is made that the controller's workload is one of the contributing factors that influence the operability and the acceptance of a concept. The following graphs indicate the workload of EC and PC separately in different Data link equipages.

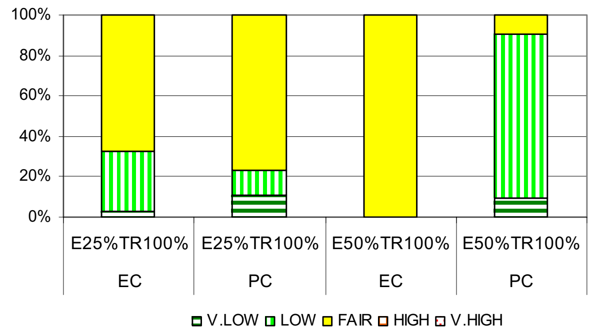


Figure 6: Workload in TR100%

Independently from the Data link equipage, an occurrence of high workload was not once recorded in the 100% traffic sample. The lowest workload was recorded in the 50% equipage on the PC position.

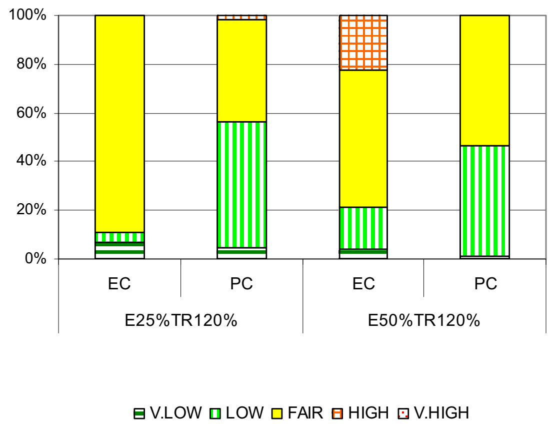


Figure 7: Workload in TR120%

On the EC position with 50% Data link equipage, 22% of the occurrence of high workload was recorded. This reflects the answers of the controllers regarding acceptability of the workload where 8% of the controller emphasized that the final working method is operable under low traffic load. Furthermore it is indicating that the switching from data link to non data link aircraft is demanding. In a high traffic situation the switch between communication means has an impact on the workload of the controller.

2) Situational awareness

Following results are gained via interviews with the controllers which took place after the controller was working on a measured position.

The majority (89%) of the controllers mentioned that they were always or most of the time able to anticipate the situation and always able to predict the traffic situation:

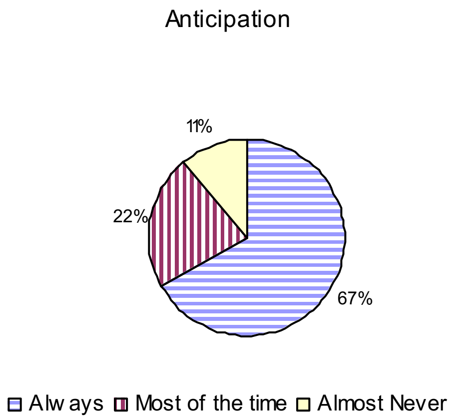


Figure 8: Anticipation

Controller workload and the planning task of the PC were mentioned as influencing factors in all situations regarding anticipation. Indeed, when the controller was able to anticipate the situation, it resulted in less workload which was caused by the appropriate planning done by the PC. Conversely, when the controller was not able to anticipate the situation, he gave as a reason that the PC was supporting the EC and therefore more focusing on the sector and not on his planning task.

In that case trust was also considered as an issue, because as soon as the EC could trust more and expect efficient planning tasks by the PC, he experienced less workload. This trust, being supported efficiently by the PC via the defined planning tasks was one of the main strengths of the working method (WM3) agreed by the controllers.

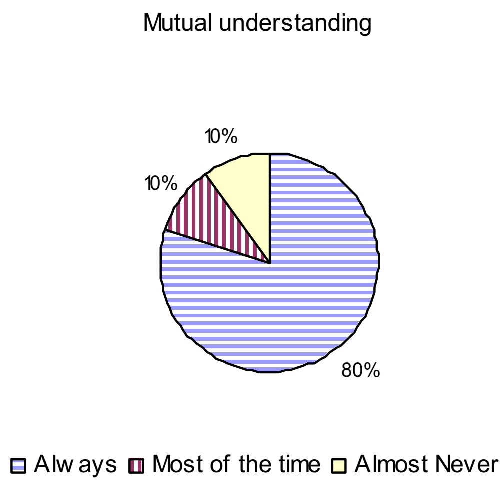


Figure 9: Mutual understanding

Additionally, a high internal verbal communication was mentioned as a reason for mutual understanding. This mutual understanding is one of the key elements of situational awareness. The accepted working method generates this mutual understanding.

It must be noted that this mutual understanding was also supported by the HMI which supported both controllers to be aware of each others' actions. Both controllers saw on their screen for instance when their team member was up linking a clearance. Furthermore the shared availability of extra data via CAP figures (Controller Access Parameter: Parameters of the aircraft which are automatically displayed on the controller working position.) was experienced as very useful in planning situations as well as in conflict situations, where the CAP figures were used as a decision support.

3) *Co-ordination*

With teamwork, co-ordination of tasks is one of the key elements for efficiency. Efficient co-ordination generates flexibility in team work and therefore adaptability to e.g. traffic increase. One important part of the co-ordination process relates to the delegation of tasks between the EC and PC. The EC could either delegate a task to the PC or the PC could ask for a task to be delegated by the EC.

The study showed that the working method supported the co-ordination process between the PC and EC, but the mechanisms were not balanced among the PC or EC regarding who initiated the delegation of a task. Indeed, the EC almost never offered to delegate a task spontaneously to the PC (see figure 10) but the PC was the one mainly asking for delegation from the EC. This result indicates the problematic of individual working cultures. In a multicultural environment the way of the team communication and co-ordination is different according to the working culture of the controller. The influence of this multicultural background has to be taken into account and further investigated as an efficient working method must support any co-ordination process between the controllers working in a team.

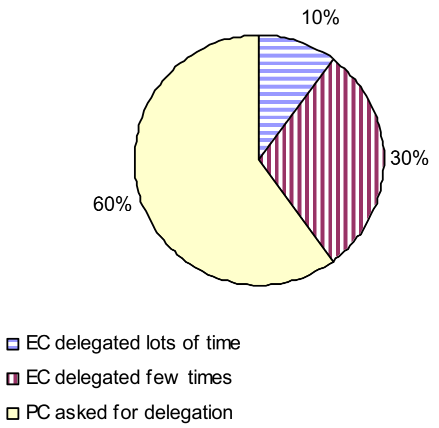


Figure 10: EC Delegation

Before taking action, the PC was asking the EC (a few times), but he was also communicating with the EC after he completed his action, and a few times he was coordinating silently.

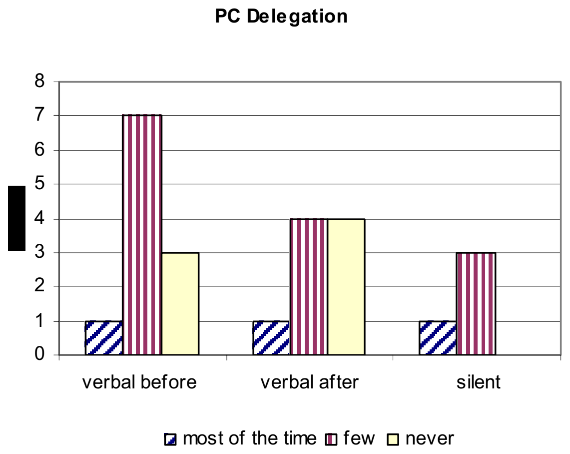


Figure 11: PC Delegation

V. Discussion

The final working method was developed in a multiple step approach. One step was performed before the simulation and the second step was performed throughout the simulation. This multiple-step method of developing the working method in an interdisciplinary expert group has turned out to be successful.

The final working method (WM3:"Task Sharing Light") resulted in less workload compared with the proposed (WM2) and the conventional working method (WM1).

Both PC and EC were almost equally performing Data link clearances whereas the EC was assuming Data link aircraft and the PC was transferring Data link aircraft. The final working method gave the EC the opportunity to delegate some tasks concerning Data link clearances to the PC. The result showed that the EC did not delegate a lot but that the PC regularly asked for delegation. The PC asked mainly before he performed the task for the delegation from the EC. Although the PC had the possibility to perform a Data link clearance without informing the EC it happened very rarely.

The final working method was considered operable and acceptable by the majority (84 %) of the controllers. The agreed working method (WM3), supported by the designed HMI contributed to a good situational awareness of the controllers. The situational awareness was measured via the anticipation and the mutual understanding of the controllers. The result of a good mutual understanding was reached partly due to a good collaboration between PC and EC and a good HMI support. In WM3 the workload of the EC was still recorded higher than the workload of the PC but controller experiences a more balanced workload.

In 50% of data link equipage with high traffic (120%) was the highest workload recorded for the EC. This allows assuming that still some additional tasks could be delegated to the PC or that the working method has to be further developed to offload the EC.

VI. Recommendations

The development of the working method has to be further improved. Initially it has to be focused on the fact that a working method and procedure have to be developed before the simulation. The first step would be to define different working methods and than to compare the working methods in a study with a bigger sample of data. Just after this the impact of the most operable working method on controller's performance in different data link equipages can be assessed to be able to make a conclusion on the operability of the working method. The work on the co-operation within the team has to be further enhanced and developed and the influence of different working cultures has to be taken into account. Furthermore the effect of the level of Data link equipage on the controller performance has to be investigated. As the present study indicated the switch from Data link to non Data link (voice) communication has to be further investigated. The outcome of such an investigation can have a deep impact on training and the related Team Resource Management as the controller might have different cognitive strategies in different Data link equipages. Furthermore the dependency of the procedure and the working method has to be investigated in connection with the system support (tools) and the airspace complexity, as the use data link itself is very much influenced by the characteristics of the airspace.

References

[1] Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In: Proceedings of the Human Factors Society 32nd Annual Meeting (Vol. 1), 97-101. Santa Monica: Human Factors Society.

[2] EUROCONTROL (2003). The development of Situational Awareness Measures in ATM systems. Brussels, Belgium by O. Straeter and M. Woldring.

[3] Hopkin, V.D. (1995) Human Factors in air traffic control. London: Taylor & Francis.

[4] LINK 2000+ programme. LINK Baseline-Pioneer Phase; EATM DAP/SPR Eurocontrol 2005

[5] Mayring, P.H. (2000). Mayring, Ph. (2000). Qualitative Inhaltsanalyse. Grundlagen und Techniken (7. Auflage, erste Auflage 1983). Weinheim: Deutscher Studien Verlag.

[6] Oze S., Schuen-Medwed R. CEATS SSRTS 5 Data link, CRDS Note No.13 Eurocontrol 2005

[7] Porterfield, D.H., 1997, Evaluating Controller Communication Time as a Measure of Workload, International Journal of Psychology,7,17,171-182.

[8] Rantanen, E. M. (1994).The role of dynamic memory in air traffic controllers' situational awareness. In R. D. Gilson, D.J. Garland,& J.M. Koonce (Eds) Situational awareness in complex systems (pp/209-215). Orlando,FL: University of Central Florida, Center of Applied Human Factors in Aviation.

[9] Wickens C.D.,A.S. Mavor. And J.P. McGee (Eds) (1998) Flight to the future: human factors in air traffic control, National Academy Press, Washington DC, USA.

Renée Schuen-Medwed has a Masters degree in psychology. Since she joined the CEATS Research and Development Centre (CRDS), Eurocontrol in 2003 she worked on several studies investigating new ATM concepts. She is in charge of the Human Factors Analysis in several simulation projects.

Adoption and Adaptation of Automation: Survey Findings on URET

Tatjana Bolic

Abstract— **The objective of this paper is to present survey findings of the assessment study of air traffic controllers' adoption and adaptation of new technologies, using the User Request Evaluation Tool (URET) as a case study. The importance of this investigation lies in better understanding of changes brought about by the use of Decision Support Tools by sector controller teams.**

The main purpose of URET is to support sector team strategic planning allowing controllers to concentrate on more user-beneficial control actions. Actual improvements depend on the way controllers use the automation tools in their work. A survey was administered in one Air Route Traffic Control Center (ARTCC) to collect the data on URET usage and influencing factors. Survey administration and the results are discussed.

The goal of this research was to draw lessons from the experience with URET that can inform technology deployments in the future. Lessons learned so far are presented.

Index Terms— **Adoption, Adaptation, Air traffic controller, Automation, URET**

I. Introduction

THE objective of this research is to investigate how air traffic controllers adopt and adapt new technologies. The focus is on automation for en-route air traffic controllers and facilities. The deployment of any new automation tool is motivated by anticipated improvements in system performance that are expected to result from its use. But actual improvements depend on the way controllers use the automation tools in their work. In other words, results depend on the adoption and adaptation of new technology, as well as the inherent value of the technology. We study the adoption process (that is comprised of both adoption and/or adaptation) in order to understand, (a) how and why it varies among controllers and (b) how and why the tool's use differs from pre-deployment expectations.

Automation tools or decision support tools (DSTs) have emerged as a means to overcome human operator limitations as they become more significant with growing traffic. Researchers have analyzed post-deployment system benefits of DSTs but there has been little research on the processes of controller adoption and adaptation of the deployed technology.

Each DST is deployed with certain expectations as to its use and the benefits it will provide. Deployment is a part of a diffusion process "by which an innovation is communicated through certain channels over time among the members of a social system [1]." Adoption of innovations is studied in the Diffusion of Innovations (DI) field, investigating various aspects of innovations adoption. Because of similarities found between general diffusion of innovation studies and the current diffusion of technology in the Air Traffic Control, the DI practices were chosen for use in this study.

During the implementation process, which is one of the sub-processes of diffusion, operators adopt, reject or adapt technology. **Adoption** refers to the intended technology use, while **adaptation** describes the modification of an innovation e.g. using the Lotus spreadsheet as a word processor. Any implemented innovation may or may not undergo adaptation by its users. Innovation adoption and/or adaptation in turn influence controller performance.

We employ the case study in our research. The specific DST studied is the User Request Evaluation Tool. Currently, URET has been deployed in more than half of ARTCCs in the US. Three things have been noticed about its usage: first, different teams use it in different ways; second, in many instances URET usage differs from what was intended; and third, the usage differs across centers as well. People involved in its deployment suggest various explanations for the observed differences in usage: inappropriate training, insufficient staffing, differences in center cultures, personal differences and others. A rigorous, carefully designed study of URET adoption and its adaptation variation has not yet been performed.

Our research goal is to analyze controller adoption/adaptation (for brevity termed "adoption" hereafter) of new technology. Here, we present the findings on URET adoption, based on adoption survey of controllers in an ARTCC. Methodology, results and conclusions regarding the URET adoption in the center are presented in the following section. URET functionalities are described first.

II. URET Description

URET is designed as a strategic support tool for the D-side controller of the en-route sector team. A sector controller team has from one to three persons, depending on the traffic-

Tatjana Bolic is with the National Center of Excellence for Aviation Operations Research, UC Berkeley, 107E McLaughlin Hall, Berkeley, CA, USA. Phone: 510-316-5609, tbolic@berkeley.edu

induced workload. Usually, during low traffic one controller works the sector. A second controller, and in rare cases even a third, is added when traffic and associated workload demand it. A two-person controller team consists of [2]:

1) A Radar Controller (R-side), a fully certified controller.

2) A Radar Associate Controller (D-side).

With URET the D-side controller should be able to help the R-side controller to resolve potential conflicts of aircraft that are not yet under the sector's active control, to check if the clearances the R-side controller is issuing are conflict free, and to perform other D-side duties. The implementation expectation was that by using the tool fully the controller team will not only be able to handle more aircraft (because of workload decrease from automation and availability of more accurate information), but will also be able to provide more direct routings and better flight profiles to airlines.

URET has 4 main functionalities. The core functionality (by design) is an automated **"conflict probe"** [3] that detects and displays the potential conflicts (up to 20 minutes in the future) in the **Aircraft List and Plan Display**, which automates flight strip management function. Aircraft List is text based and contains the list of active and incoming flight plans, conflict probe alerts in the dedicated box next to the affected flight plans, and trial planning results. By clicking on a conflict notification controller gets the graphical display of flight routes of involved aircraft and time to potential conflict. Another function, the **Trial Planning**, allows a controller to check a desired flight plan amendment for potential conflicts. If the trial plan is conflict-free, this clearance can be issued to the pilot and at the same time sent to the Host Computer System as a flight plan amendment. Finally, the **Route Amendment** function offers point-and-click entry of aircraft route amendment. A controller can just click on the list of routes or list of fixes or enter its name in order to change the route directly in the Host Computer System.

III. METHODOLOGY

The effort to analyze the controllers' adoption of URET comprises of two phases. Initial phase involved two rounds of interviews with several Subject Matter Experts (designers, controllers, people involved in URET deployment), to learn about the tool itself and its usage. It resulted in 6 hypotheses regarding URET usage [4] of which two that were tested in the analysis are presented in the following section. Consequently, the survey instrument to investigate the hypotheses was designed and refined with the help of SMEs.

The second phase (still in progress) comprised of defining the sample, administering the survey and survey analysis. These are described in subsequent sections.

A. Hypotheses

Hypotheses 4 and 5, that apply to this analysis are defined and described:

Hypothesis 4: Adaptation of URET use will result when controller works the sector alone, and if the controller works alone most of the time, the adaptation will be even more pronounced (adaptation rendering URET almost solely a strip replacement tool). URET is designed as a DST for the D-side controller, which makes it more complex for one controller to use the tool to the full extent and it does not support well the operations of a one-controller team for two reasons: first, when working alone, the controller needs to communicate with the aircraft and generally does not have time to devote to more strategic solutions. Second, in order to use URET, one needs to use its trackball, situated at the D-side, which is cumbersome for the R-side controller. Thus, when controller works alone adaptation is a likely expectation that also subject matter experts have reported.

Hypothesis 5: Strategic environment is more suited for full URET use than the tactical environment, therefore URET adoption in tactical sectors will be either slower or adaptation will take place: using only the strip replacement function. URET primary role should be that of strategic support and the tactical (mostly low altitude) sectors need mostly tactical solutions because through them aircraft climb from the airport and descend to the approach control. Two reasons were pointed out why URET would see less usage of trial plan and conflict alert functions in these sectors. First, the approach and tower controls do not provide information for the URET system; thus the conflict alerts for such flights are not as accurate as for flights coming from airspaces that share the information with the URET system. Second, the environment is very dynamic, and tactical in nature and both controllers are often too busy with other aspects of their job to be able to efficiently use the trial planning and conflict alert function.

B. Survey Design and Survey Administration

The survey questionnaire design was based on the hypotheses and the influencing factors (presented in [4]), identified in the initial phase of adoption analysis. The survey questions cover basic demographic information on controllers (i.e. age, sex), subjective ratings of the traffic situation during the past time-on-position, controllers' ratings of the extent of URET usage in different situations (termed control situations hereafter), and their opinion on positive and negative sides of using URET.

The control situations were classified (see Fig. 1) into conflict, direct-to and altitude change situations. For any situation identified in Fig. 1 that occurred during their past time-on-position, controllers rated to what extent they used a specific URET bundle for it. Each of the four described URET functionalities can be used almost independently. Such loosely bundled tools are easily subjected to adaptation that in case of URET is reflected in not using some or most of the functionalities. Therefore, the exhaustive set of functionality bundles that could be used for resolving each previously mentioned control situation was identified. A bundle is a combination of URET functions that can contain from one to four mentioned functions. Functionality bundles and the related control situations are listed in Table I. Fig. 2 shows an example of the URET usage question in case of altitude change clearances where the levels of usage are: did not use it, used in for some cases, and used it for most cases.

The survey was administered in one ARTCC, that was chosen to represent the third of the 4 groups of centers that are fairly homogeneous internally with regard to two characteristics: the length of time they have been using URET and the training received.

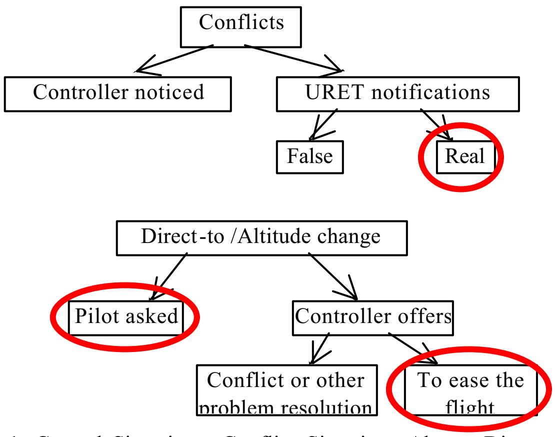


Figure 1. Control Situations: Conflict Situations Above, Direct-to and Altitude Changes Below.

Controllers have been using URET for two and a half years. The traffic count is not very high, but the airspace is very complex: many general aviation airports and many restricted airspaces. A fair percentage of transitional flights exist and combined with a particular mix of aircraft types contribute to complexity. After the start of URET daily use, an increase in instances of controllers working alone has been noted, attributed to the strip handling workload decrease. There is no mandate for working in two-controller teams unless workload dictates it. Two-controller teams are mandated at all times only for a few sectors where operational errors happen. Moreover, the center culture is more individualistic, and the R-side position is perceived as being more important than D-side as well as being a "command position" that issues tasks to the D-side position.

To facilitate altitude changes issued to pilots on controller's initiative, I used the following combination of URET FUNCTIONALITIES (CHECK ONE BOX IN EACH ROW AS APPLIES TO YOUR TIME-ON-POSITION):			
	Did not use it	Used it for some cases	Used it for most cases
Aircraft List display only		√	
Aircraft List display and Trial Plan	√		
Aircraft List display and Route Amendment		√	
Aircraft List display, Trial Plan and Route Amendment	√		

Figure 2. Altitude Change URET Usage Question.

All the center controllers received 36 hours of training for URET use. Four hours of it was the training in the control room with the training cadre going through examples of how to use URET in specific situations. We learned that during the months needed to complete the training of all controllers the peer pressure has been used to help introduce URET usage faster, and an incentive procedure saying that if both R-side and D-side are trained, then they can use URET and disregard paper strips was set.

Survey took place on August 22nd, 23rd 2005 from 10 am to 2 pm (ARTCC busy hours), when 36 surveys were collected. All the controllers that worked directly with URET (R-side working alone, or the D-side) during those 8 hours represented the target population. Each day about 80 controllers were working during the 4 hours. Every controller leaving the control room was asked to fill-out the survey. If they agreed, their break was extended for 15 minutes (time to complete the survey), according to the air traffic controllers' union regulations (NATCA).

TABLE I. SUMMARY OF FUNCTIONALITY BUNDLES AS THEY RELATE TO THE CONTROL SITUATIONS.

Bundle	Control Situations
Aircraft list only	Conflicts, Direct-to pilot asked, Direct-to, controller offered, Altitude change pilot asked, Altitude change controller offered
Aircraft List, Trial Plan	Conflicts, Direct-to pilot asked, Direct-to, controller offered, Altitude change pilot asked, Altitude change controller offered
Aircraft List, Route Amendment	Direct-to pilot asked, Direct-to, controller offered, Altitude change pilot asked, Altitude change controller offered
Aircraft List, Trial Plan, Route Amendment	Direct-to pilot asked, Direct-to, controller offered, Altitude change pilot asked, Altitude change controller offered
Aircraft List, Clicked on Conflict Notification	**Conflicts**
Aircraft List, Clicked on Conflict Notification, Route Amendment	Conflicts
Aircraft List, Clicked on Conflict Notification, Trial Plan	Conflicts
Aircraft List, Clicked on Conflict Notification, Trial Plan, Route Amendment	Conflicts

C. *Survey Data Analysis*

This section reports the analysis of the part of the collected data: URET usage in different control situations (circled, Fig. 1). A quantitative analysis - testing of hypotheses is presented first, followed by the qualitative analysis of the extent of URET usage.

a) Quantitative Analysis: Hypotheses Testing

Controllers were asked to rate how much they used different bundles of URET functionalities in each of the identified control situations. A group of models was estimated for each of the control situations; group comprising of related URET functionality bundles (Table I). Here, a model for combined usage of Aircraft List and Clicked on Conflict Notifications in case of conflict resolutions, and the hypotheses tested are presented.

Discrete choice model was used to estimate parameters of controllers' choice functions. Estimated parameters were then used to test the hypotheses. The ordered probit model was employed to account for the ordinal nature of the dependent variables that rate the extent of usage of a specific URET functionality bundle. The dependent variable has three ordered levels: did not use it, used it for some cases, and used it for most cases; then choice probabilities are (1):

$$P(did\_not\_use\_it) = 1 - \Phi(\beta' x)$$
$$P(used\_it\_for\_some\_cases) = \Phi(\mu - \beta' x) - \Phi(-\beta' x) \quad (1)$$
$$P(used\_it\_for\_most\_cases) = 1 - \Phi(\mu - \beta' x)$$

The set of explanatory variables used in controllers' choice functions is shown in Table II. Presented variables were used to estimate a full model. Consequently, removal of highly non-significant variables resulted in reduced, better fitting (preferred) models.

TABLE II. INDEPENDENT VARIABLES USED IN ESTIMATION OF ORDERED PROBIT MODEL.

Variable	Values
Rd	=1, if the controller worked as an R-side =0, if the controller worked as a D-side
Age1	=1, if the controller's age<35 =0, otherwise
Age2	=1, if the controller's age<35 and <=45 =0, otherwise
Fl1	=1, if the flights in the sector were mostly level flights =0, otherwise
Fl2	=1, if the flights in the sector were mostly transitional flights =0, otherwise
Trd1	=1, if the traffic density in the sector was moderate and heavy =0, otherwise
Trd2	=1, if the traffic density in the sector was light =0, otherwise
Tra13	Controller rating of agreement with: "Tool use for two-controller team manning a sector was explained well and examples were provided" (1- disagree, 5- agree)

Table III shows the estimates and the significance levels of both the full model and the preferred (reduced) model.

Estimated choice probabilities are:

- P(did not use it)=0.666
- P(used for some cases)=0.319
- P(used it for most cases)=0.013

This functionality bundle is less likely to be used than not, which is consistent with the findings of initial interviews [4].

All variables included in the preferred model are significant at the 5% level. However, the marginal effects of the independent variables on the probabilities in the ordered probit model are not equal to the estimated coefficients. To analyze the marginal effects of dummy (or categorical) variables we compare "the probabilities that result when the variable takes its two different values, with the other variables held at their sample means" [5]. The effect of changing the value of one of the variables while holding the others constant is equivalent to shifting the distribution slightly to the right (or left). Fig. 3. shows the change in predicted usage probabilities with the controller position change.

TABLE III. ESTIMATED COEFFICIENTS FOR AIRCRAFT LIST DISPLAY AND CONFLICT NOTIFICATION USAGE (FULL AND PREFERRED MODELS)

	Full Model		Preferred Model	
Parameter	Estimate	Significance	Estimate	Significance
Intercept	3.9010	0.0665	-0.6276	0.2118
Intercept2	2.1398	0.0058	**1.7876**	**0.0031**
RD	**1.9229**	**0.0201**	**1.3053**	**0.0332**
Age1	-2.3277	0.0986		
Age2	0.8075	0.3332		
Trd1	**-3.2535**	**0.0212**	**-2.3319**	**0.0296**
Trd2	-0.7723	0.5214		
Fl1	**-2.8285**	**0.0407**	**-2.1113**	**0.0302**
Fl2	0.4756	0.6327		
Tra13	0.0029	0.9917		
Tra4	-0.0606	0.8346		
	LR_R	0.9679	LR_R	0.9883

Controllers working alone (R-side) are much less likely to use Aircraft List and Click on Conflict Notifications for resolving conflict situations. On the other hand, when controllers are working as a D-side they are more likely to use these functionalities. The effect of this variable on the choice probabilities is such that we cannot reject the hypothesis 4 (the part of it):"Adaptation of URET use will result when controller works the sector alone". In this case, the adaptation is reflected in the fact that the most of controllers working alone do not use Conflict Notifications and/or Aircraft List display.

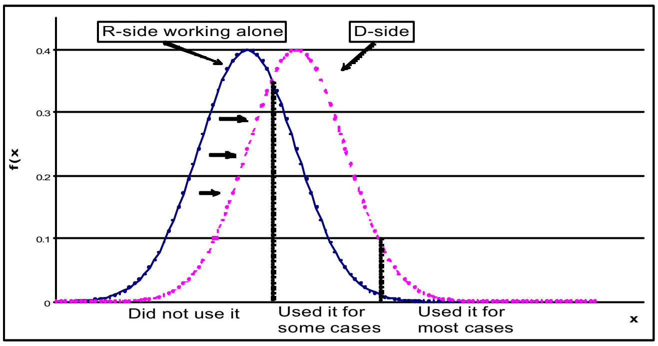


Figure 3. Effects of Change in RD on Predicted Probabilities.

Hypothesis 5 states: "Strategic environment is more suited for full URET use than the tactical environment." A sector is in a strategic environment when most of the flights in it are level flights. Fig. 4 displays the effect the environment (level flights, climbing/descending flights, mix) has on controllers' choice probabilities. When flights are level, it is more likely that controllers will use this URET bundle than not. On the other hand, for climbing/descending flights types it is more likely that the controllers will not make use of Conflict Notifications, and/or Aircraft List display. These results support the hypothesis 5. The estimated effects of the traffic density show that when the traffic is moderate (Trd1=1) controllers are more likely to use this bundle. When the traffic is light they are least likely to use it. This is to be expected since the likelihood of conflict situations is lower in light traffic then moderate/heavy traffic.

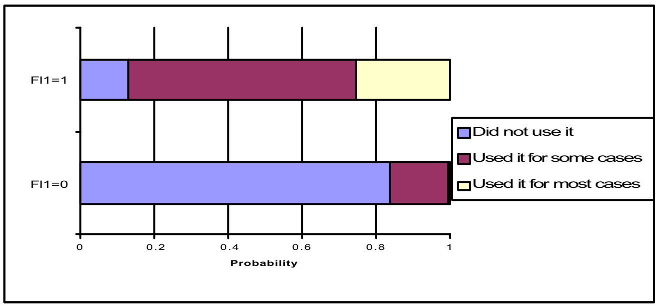


Figure 4. Effects of Change in Fl1 on Predicted Probabilities.

b) Qualitative Analysis

Rating of usage of each URET bundle was collected during the survey. Here, these ratings are presented, with explanation of the reasons behind such usage utilizing surveyed controllers' opinions and the findings of the first interview rounds [4].

Fig. 5 shows the distributions of usage ratings for conflict resolution control situations. The most used bundle is Aircraft List only (A), which is to be expected since they need it to be aware of situation, thus being aware of conflict situations and the ways of their resolution. Initial interviews indicated that this is the most used bundle in any control situation (see also Fig. 6 and 7). This is further corroborated by the controllers' opinions on URET: seventeen out of thirty six said that URET is "great" as a flight strip replacement, and a great help in easing the amendment input. Further confirmation of it is the fact that the second most used bundle is Aircraft List, Trial Plan and Route Amendment (A,TP,RA).

Surprisingly, the third place by usage is taken by "Aircraft List and Clicked on a Conflict Notification" bundle (A,C) and "Aircraft List and Trial Plan" (A,TP). Initial phase interviews in this and other centers revealed that conflict notifications are used very little, and about 20% of the controllers that filled-out the survey stated that they do not use conflict notifications at all, and are relatively indifferent towards trail planning. Here we have evidence that the conflict notifications and trial plans are being used, more than expected. This might be because the new controllers are utilizing this bundle and it needs to be investigated further.

The least used bundle is Aircraft List, Conflict Notification and Trial Plan. This is consistent with the initial findings where the controllers of this center say that they are trained to resolve the conflicts in their own sector and not to worry about the influences their actions will have on neighboring sectors, on one hand. On the other hand, their airspace has many restricted areas that are very dynamic in nature, and URET is still not able to capture this, thus making the Trial Plan of little use.

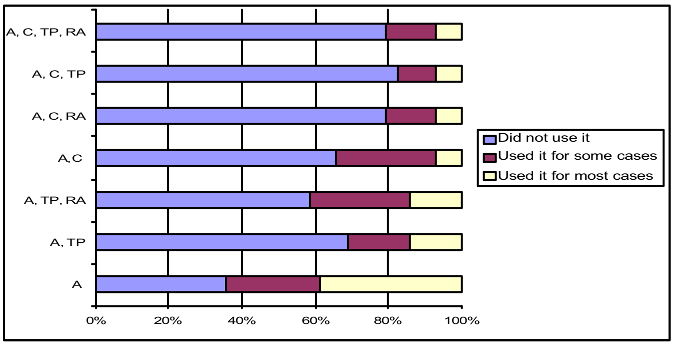


Figure 5. URET Functionality Bundle Usage for Conflict Situations

The fact that the controllers are more likely not to use any of the URET functionalities can be explained by the controllers' statements: "9 times out of 10 I will see a conflict before URET does", or "I do not completely trust "conflict probe". It is wrong about 50% of the time." The experienced controllers say that most of the time they already know if and how the clearance is going to work. Meaning that in the current state of air traffic control practices they do not perceive a clear advantage of using URET for conflict resolutions.

Direct-to and altitude change situations are divided into ones requested by pilots and the ones that controllers offered. Upper part of Fig. 6 and 7 show the usage distributions for situations when pilots asked for a direct-to clearance, and the lower for the situations when controllers offered direct-to to pilots.

An interesting finding about direct-to control situations is that some of the bundles, containing trial plan function are less likely be used, while others (A and A, RA) are more likely to be used (see Fig. 6). Overall, all the bundles are used more for direct-to control situations resolution than for conflict resolution. Controllers are almost equally likely to use "Aircraft List and Route Amendment" (A, RA) bundle at any of the 3 levels. The Route Amendment automatically amends the flight plan so that all the upstream controllers are informed, and thus lowers the amount of coordination needed. Decrease of communication workload probably accounts for the high usage of this bundle. In other words, using it makes issuing direct-to clearances easier so the controllers are more likely to embrace its usage. High usage of Route Amendment function in this center was reported in the findings from the initial interviews.

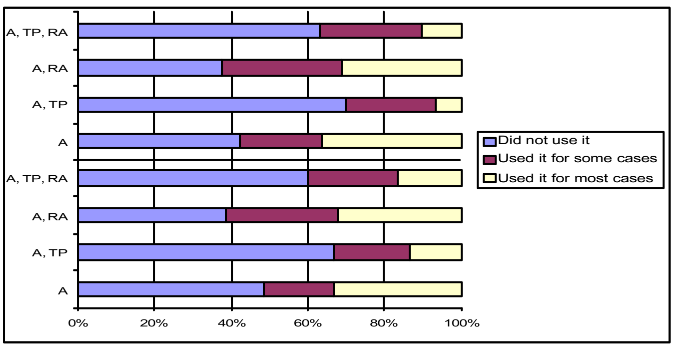


Figure 6. URET Functionality Bundle Usage for Direct-to Situations.

Altitude change control situations are also being resolved using mostly Aircraft List only. Overall, URET functionalities are rarely used when providing the altitude change clearances (similar to the conflict resolution situations). During the initial phase of this research controllers reported that it is easier to input altitude change amendments through the R-side console than through URET. This is the likely reason for such a low usage.

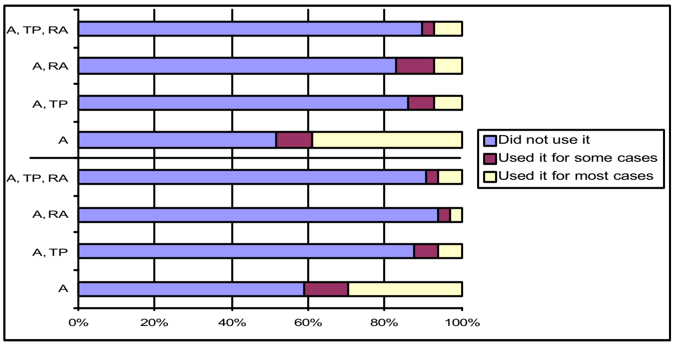


Figure 7. URET Functionality Bundle Usage for Altitude Change Situations.

IV. CONCLUSIONS, RECOMMENDATIONS AND FUTURE WORK

Data collected in one ARTCC allowed us to test two hypotheses. Parameters estimated using the ordered probit models show that there is a difference in URET usage depending on the position the controller works on. The D-side controllers are more likely to use URET functionalities to at least some extent, since it is not as complicated to use in this position. Controllers' opinions collected and the findings of the initial interviews support this explanation.

Having mostly level flights in the sector increases the probability of using URET functionalities, while transitioning flights lower it. This adaptation is the consequence of the design characteristics of the tool and the environment in which it is applied.

What is more interesting is that results show differences in URET usage for three categories of control situations: conflict resolutions, direct-to changes, and altitude changes. Usage is the highest for providing direct-to clearances, likely because for this type of clearances using URET reduces communication workload, offering a clear advantage over the previous methods.

Conflict situation resolutions see less usage of URET, since the tool does not offer a more advantageous method over the current practices of experienced controllers. In years to come, this is likely to change since the new controllers are being trained to use URET from day one. Moreover, some controllers say it is useful for a controller-in-training to learn the effect of the issued clearances by using Trial Plan function. URET is least used for facilitating altitude changes. The likeliest explanation is that it is easier to amend the changes using the R-side console. Again, when there is no clear advantage to using automation, adaptation takes place; in this case low usage.

To sum up, when advantage of completing the tasks in the current tactical control practices by using automation exists, the automation tool is used.

Findings reported in this and [4] exposed some lessons that can be used in the future technology deployments. Technology advantages are highly susceptible to influences of both the physical environment and social structure of the center (i.e. center culture). These need to be taken into account prior to initial deployment phase so that the training and/or new operational procedures can be tailored to address integration issues.

Changes of work methods caused by the use of automation need to be addressed through training and changes in related "operating procedures" on the institutional level. It has been hypothesized that using URET for strategic planning will result in better system performance and more user benefits in terms of shorter flight times and better altitude profiles. However, current "operating procedures" focus on the tactical control as was noted in this center. Not even the continuing practices are addressing the envisioned change to the strategic control.

The changes in controllers' job tasks with the introduction of the new technology should be tracked, so that the integration of other new tools (or changed procedures) can be as smooth as possible.To sum up, the fast adoption depends not only on the characteristics of the technology. A well thought out implementation plan can increase the rate of adoption.

Future work will focus on administering surveys in couple of other ARTCCs, which will enable investigating differences in URET usage and the reasons behind it across the centers, and testing of other hypotheses.

REFERENCES

[1] Rogers, Everett, 2003, Diffusion of Innovations, New York, The Free Press

[2] Nolan, Michael, 1990, Fundamentals of Air Traffic Control, Belmont, California, Wadsworth Publishing Company

[3] Kirk, Daniel, October 2002, Enhanced Trial Planning and Problem Resolution Tools to Support Free Flight Operations, Report MP00W0000293, MITRE, MITRE

[4] Bolic,Tatjana and Hansen, Mark, 2005, Automation Adoption and Adaptation: User Request Evaluation Tool Case Study, Transportation Research Record: Journal of Transportation Research Board, Issue 1915, pp 62-68

[5] Greene, William, 2003, Econometric Analysis, Upper Saddle River, N.J. : Prentice Hall

Tatjana Bolic is currently a PhD Candidate at the University of California at Berkeley. She obtained a Bachelor's degree in 1999, from Faculty of Transport and Traffic Engineering, University of Belgrade, Serbia and Montenegro. She obtained a Master of Science degree in Transportation Engineering in 2002, from UC Berkeley. Her research interests include air traffic control, human factors, and air traffic management.

Multi-Agent Automation in Arrivals Management

Italo R. Oliveira and Paulo S. Cugnasca

Abstract— **The demand growth in air transportation, without an adequate evolution in its management systems, leads to more delays and more incidents in flight operations. The delays must be bounded with an increase of the traffic system efficiency, and the incidents must be bounded with an increase in the system safety. Within this context, the arrivals management appears as a key-point to the achievement of these goals. This paper presents the risks related to aircraft separation loss, gives examples of existing tools for arrivals management, and describes the multi-agent approach for Air Traffic Management, which strongly relies on FMS and data link communications.**

Index Terms— **Air Traffic Management, Arrivals, Airspace, Capacity, Conflict, Multi Agent, Runways, Safety.**

I. Goals

The main goal of this work is to analyze the efficiency of new scenarios of arrivals management in air traffic Terminal Maneuvering Areas (TMAs) with high traffic density. On the other hand, there is a strong safety requirement in Air Traffic Management (ATM), evaluated as the protection level that the same system offers to maintain the integrity of its users and operators.

II. Introduction

The air transportation demand increase expected to the next years is remarkable. FAA (Federal Aviation Administration, USA) works with a forecast of 2% increase per year, in the activities of its control facilities [7]. To the total air transport with origin or destination in USA, FAA previews an average growth of 4.7% [7]. Eurocontrol forecasts 3.1% per year growth, taken in average [21]. ICAO (International Civil Aviation Organization) previews that the air traffic is going to increase 5% per year, in average, up to the year of 2010 [10]. Doing an extrapolation of the growth, using the above rates, we obtain the results present in Table I.

Manuscript received February 26, 2006. This work is supported by the Foundation of Technological Development in Engineering (FDTE), São Paulo, Brazil, and the Coordination for the Improvement of High Education Students (CAPES), Brazil.

I. R. Oliveira is Ph.D. student at Polytechnic School of University of São Paulo SP, Brazil, 05508-910 (Phone: +55 11 3091 5734; fax: +55 11 3091 5294; e-mail: italo.oliveira@poli.usp.br).

P. S. Cugnasca is Assistant Professor at Polytechnic School of University of São Paulo (e-mail: paulo.cugnasca@poli.usp.br).

TABLE I
PROJECTIONS OF AIR TRAFFIC GROWTH

Source	Nature / Coverage	Mean yearly growth	Projection in 10 years
FAA [7]	Workload at FAA and subcontractors control facilities	2.00%	21.90%
Eurocontrol [21]	PAX-Km with Origin/Destination in Europe	3.10%	35.70%
FAA [7]	PAX-Mile with Origin/Destination in USA	4.72%	58.60%
ICAO [10]	World PAX-Km	5.00%	62.89%

The traffic increase, close or beyond the current capacity limit of the system, generates delays and greater frequency of incidents. This fact brings the necessity of a real capacity increase, to be obtained through the adoption of new procedures. The approach developed here was first presented in the reference [25].

III. Accidents Related to Separation

Aeronautical accidents may have as causes a vast diversity of factors. However, the accident classes in focus in this study are those ones related to separation between aircraft, or between aircraft and airport vehicles. The separation related accidents can be distinguished in some types, shown in Table III. In that table, one can notice the expected fact that accidents involving two airborne aircraft have the greatest risk.

Due to the great number of fatalities in the accidents involving two airborne aircraft, the most important risk factor to this work is the airborne-airborne collision.

TABLE III
STATISTICS OF ACCIDENTS RELATED WITH LOSS OF SEPARATION FROM 1980 TO 1999 [22]

	Fatalities	Fatal accidents	Accidents
Collision with vehicle	1	1	64
Collision with standing aircraft on ground	4	1	42
Collision with moving aircraft on ground	147	6	39
Collision with aircraft, both airborne	620	14	22
Collision with aircraft, one airborne	11	1	7
Aircraft encountered vortex/wake turbulence	0	0	8
Near collision with aircraft, both airborne	0	0	3

ICAO establishes risk target-levels to Civil Aviation, distinguished by causing factors. For example, if we consider the accidents caused by vertical separation loss between aircraft, the accepted level risk is 5.10^{-9} accidents per each performed flight hour [8]. This level must be guarded, even though the criteria for ensuring separation become modified.

The criteria of minimum separation between aircraft, commonly practiced until few years ago, were established based on the accumulated practical experience in air traffic, and this experience is dependent on sets of technologies used since the 1950's (Radar, RF voice channels, flight instruments, etc.) without structural alterations. Some of these criteria are physically and mathematically justified, as in the case where the wake turbulence is the decisive factor. However, many of them have empirical nature, as in the case of lateral separation [15, 16]. For the case of vertical separations, the low accuracy of altimeter was the most relevant factor.

With the advent of more accurate equipment, it is possible to demonstrate theoretically that the vertical separation practiced en route, that was formerly 2,000 ft, was far greater than a reasonable minimum, and besides, that 1,000 ft is more than sufficient to ensure the ICAO risk target level [19]. Because of this, ICAO started to promote the aircraft certification to use the RVSM criteria, which establishes 1.000 ft as a safe separation to be practiced en route.

IV. Arrivals Management

The aircraft arrivals management needs to ensure a safe separation between arriving aircraft and, meanwhile, it needs to have efficiency for serving the required demand of flights, along the time, which may present high demand peaks.

As the longitudinal separation criteria standardized by ICAO are upheld, a runway hardly allows more than one landing per minute. It is desirable that an arrivals programming is made, in order to sequence aircraft, in such a manner that an optimal runway usage is achieved. The updated arrivals programming, with half an hour of anticipation, could have excellent results for runway optimization [2; 14].

If there is need of a greater number of flights, it is necessary to build new runways, with adequate separation between them [16]. As a consequence, the complexity of circulation inside the TMA increases, causing greater workload for traffic controllers. Because of this, the use of automation tools is seen as an unavoidable means to increase air traffic capacity

A. Some existing systems for arrivals management

Currently, there are several automation tools that help the arrivals management, being used in several airports around the world. The most commonly term to name this kind of tool is AMAN (Arrivals MANager). Some examples are:

MAESTRO [18; 20]: with French origin, it is used in airports as Paris CDG, Paris Orly, Copenhagen Kastrup, Helsinki Vantaa and Sidney. It does the arrivals sequencing, using data from flight data processing system and from the radar data processing system. It assigns a runway to each arriving aircraft and issues advisories to the controller, in order to obtain the resulting landing sequence as close as possible to the planned sequence. It has also algorithms for distributing delays among the aircraft involved in the sequence. The delay absorbing may be done through speed decrease, or by path stretching, before, as the last measure, to perform waiting orbits, that spend considerably more fuel;

TMA-SC and McTMA [5; 13]: tools developed by NASA and other companies in USA. These tools help in the arrivals management work, but differ from MAESTRO due to the air traffic management architecture in USA.

The U.S. air traffic management architecture has many common elements to the ones present in other air traffic regions around the world, having same basic entities. However, the main difference between the USA system and most of other systems in the world is that the USA system is a homogeneous and highly integrated system to a wide airspace area.

Within this context, there is a key agent named TMC (Traffic Management Coordinator). The TMCs, which are placed in the ARTCCs (Air Route Traffic Control Center), control aircraft that have destination airports located in his Center's region, and also control aircraft that leave these airports. From the present and the predicted flow of traffic, the TMCs create plans to safely deliver aircraft to the TRACON (Terminal Radar Approach Control), in the frequency that may fill completely the TRACON and the destination airports capacities, but without exceeding them. The TMC's plans consist of aircraft sequences with scheduled times (STA - Scheduled Time of Arrival) for those aircraft to reach metering fixes (Fig. 2). These fixes are points located, generally, in the border of the terminal area. On the other hand, the controllers of the Centers give instructions for the aircraft under their control, in order that they cross the metering fixes in the times specified by the TMC (STAs). Close to the TRACON's border, the controllers of the Center transfer the arriving aircraft to the TRACON's controllers.

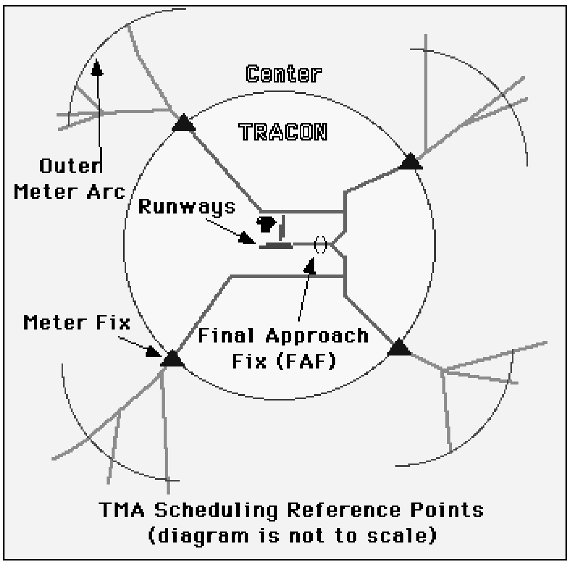


Fig. 2: References used to manage airport arrivals, in the U.S. system [5]

The TMA-SC tool (Traffic Management Advisory - Single Center) assists the TMCs and the controllers of traffic in several ways. One of them is to increase the situation awareness through graphical interfaces and alerts (Fig. 3).

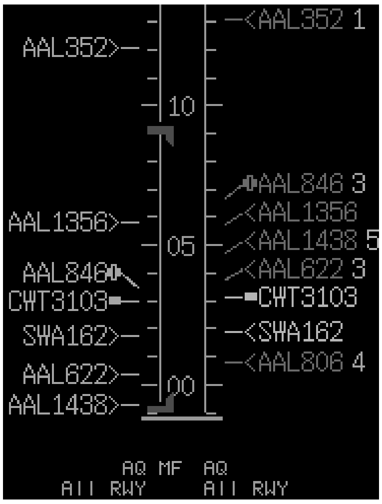


Fig. 3: Traffic peak alert, provided by TMA-SC. On the left side and on the right side, the sequence of arriving flights is displayed [5]

Another functional characteristic is to generate statistics and reports on the traffic flow. Further else, TMA-SC calculates the estimated times of arrival (ETA) without delay, for the external arcs of measurement (Fig. 2) of the Centers. In addition to the TMA-SC, the McTMA tool (Multiple-Center Traffic Management Advisory), assists the TMCs to preview the future demand, involving multiple Centers, and to regulate the traffic flow in the involved terminal areas.

TMA-SC was installed initially in the Ft. Worth Center (ZFW) and, after operational testing period, was installed in six other Centers¸ within the context of Free Flight Program Phase 1 [3]. For the Free Flight Program Phase 2, it is foreseen that it will be installed in four other control facilities. Comparative data, extracted from the Centers where TMA-SC was installed, indicate that it allows from 3% to 5% of capacity increase in the traffic. This percentage, though sounding small, means a greater fluidity in peak schedules, making possible a significant reduction of delays. Meanwhile, the McTMA was implanted in experimental phase in the Philadelphia TRACON and will integrate operations of New York, Cleveland, Boston and Washington Centers.

B. FMS utilization in the arrivals management

From the moment that an aircraft has taken off, its FMS (Flight Management System) is able to calculate an estimate of its landing time [14]. Until these days, the estimate of the landing time is sent by ACARS (Aircraft Communications Addressing and Reporting System) to the Airline Operations Center, but this information still is not available in the Aeronautical Telecommunications Network (ATN). If that estimate were available in the ATN, it could be communicated to the downstream control centers. If this landing time estimate was known by Arrivals Manager (AMAN) at this advanced moment, AMAN would be able to calculate with greater precision the required arrivals schedule, taking into account several concurrent aircraft. The possibility of this FMS use is an important step to improve the estimation of the time each aircraft will effectively block on the gate. As an alternative to ACARS, it is possible to use ADS-B to send a 4-D estimated trajectory for the aircraft, calculated by FMS. In one way or another, it is possible to use the trajectory preview to generate efficient and feasible arrivals sequences [1].

Moreover, it is possible to use the CPDLC system (Cockpit-Pilot Data Link Communication) to make the aircraft FMS to be automatically programmed for executing the established sequence. In this case, the traffic controller instructions would be elaborated through some tool that, by means of an interface inside the aircraft, makes automatically the FMS reprogramming (see Fig. 4). This procedure would substitute, with advantages in many cases, the current radar vectoring.

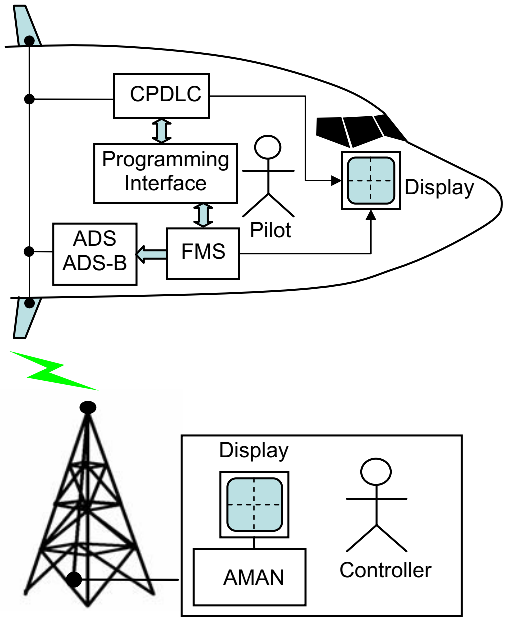


Fig. 4: Use of FMS in the arrivals management

V. ATM SYSTEM REQUIREMENTS

ATM Systems need real time information flow and processing, once the air navigation has its proper timing and dynamics, on which depend the integrity of human lives and properties of people and institutions involved in aviation activities. The real time information flow can be achieved by means of safe, high reliability, and high performance telecommunications equipment, including data links and routers.

Besides the telecommunication and data processing infra-structure, it is necessary to provide user applications that allow acting with knowledge of an adequate set of information. User actions must be constrained by some coordination rules and, once started, must be immediately divulged to other users under influence of these actions or, at least, be accessible to all users in a circle of interest. These working properties turn ATM System an integrated management environment that requires the use of tools enabling Collaborative Decision Making (CDM).

The view of integrated information management should guide the CDM applications developers to consider as essential the requirements of ATM information generation and distribution. Such requirements include [4]:

- *Precision and accuracy*: information must be sufficiently precise / accurate, and user must be informed about the confidence levels or intervals for the validity of information.
- *Stability*: information must be sufficiently stable and not subjected to spurious oscillations. Even valid oscillations should be filtered in many cases.
- *Coverage*: information must be sufficiently complete in terms of including all elements relevant to a given situation.
- *Opportunity*: information must be available at the necessary moment.
- *Consistency and Coherence*: there must be no conflict in information, which have to make sense to user (as an example, an ATC entity has to receive a flight plan not after the respective flight execution). Being the information the basis for decision, there must be synchronization in the information updating, which therefore allows coherence in the decision making process.
- *Security and Confidentiality*: non-authorized interventions in the system have to be avoided. Confidential information, mainly in special operations, must be protected.
- *Availability*: lack of availability may generate inefficiency and safety problems, as traffic inertia cannot be suppressed neither in nominal nor in degradated states.

The challenge to the integrated information management in CDM is not simply the distribution of relevant information but, overall, is to have the engagement of maintaining and providing task performance abilities in different scopes of the system. And it is in this point that CDM concept associates itself with the multi-agent interaction paradigm [24]. In a multi-agent environment, the problem resolution is decentralized and, since the agents actuation powers are balanced, this way of working can obtain high quality results in systems where the centralized resolution would have too high complexity.

VI. THE MULTI-AGENT APPROACH

Jennings [12] shows that, due to multi-agent systems characteristics, they have significant advantages over monolithic problem solvers in several situations. These characteristics are:

- Exploitation of parallelism, which may lead to greater velocity in solving problems;
- Smaller amount of data traffic by exchanging high level partial solutions instead of raw data to a central place;
- More flexibility, following the fact of having agents with different skills interacting to solve a same problem;
- Increase of availability by allowing active agents to assume responsibilities of faulty agents.

The definition of agent used here is from [24], which states that an agent is a computer system, located in some environment, and capable of performing autonomous actions in this environment, seeking to reach its proposed objectives. Originally [23], an agent was understood as an entity with embedded problem solving ability.

Due to the distributed nature of air traffic, and to its complexity, the multi-agent systems characteristics seem to be adherent to this application field. One of the objectives of this research work is to define basic agent properties, methods and communication schemes, in such a way that the introduction of a new automation tool, the modifications in existing tools, or the absence of some of them do not require structural changes in the system nor in the interaction rules [17]. Missing functions of faulty tools may be supplied at least partially by interaction of other agents.

A. Agents interaction

A set of specialized agents was defined in order to identify common characteristics for agents in air traffic. This is not a closed set, but represents the most significant agents for the ATM decision making and communication process. This set is composed by agents depicted in Fig. 5 and Fig. 6.

Some agents in Fig. 5 and Fig. 6 correspond to existing air traffic actors, such as:

- ATMS is the integrated ATM automation system, for a national or continental system.
- ATFM is the Air Traffic Flow Management, which performs strategic system planning to avoid traffic excess in some critical space-time regions.
- ATCS is the Air Traffic Control System, which handles executing flights, giving clearances directly to aircraft.
- AOC is the Airline Operations Center, which coordinates a set of aircraft in order to reach the business objectives of an airline.
- Aircraft agent is the aircraft including all its subsystems and pilots.

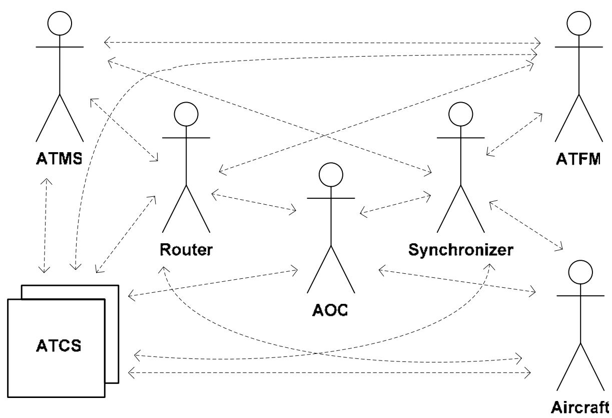


Fig. 5: Agents in ATM and its communication lines.

Other important agents are Router and Synchronizer. These agents were introduced to help in the agent communication process. The Router agent is basically a server for finding data paths and connections between agents through the available networks. The Synchronizer agent provides universal time reference, data flow control reference and data conflict management. Notice that all agents in Fig. 6 are, except Router and Synchronizer, part of ATCS.

Besides the agents, Fig. 6 presents other elements, such as:

- Meteo data sources, which provides weather data;
- AMS - Agent Management System - it manages the entrance and the exit of agents in the system, and provides references to agents to know the abilities of each other.

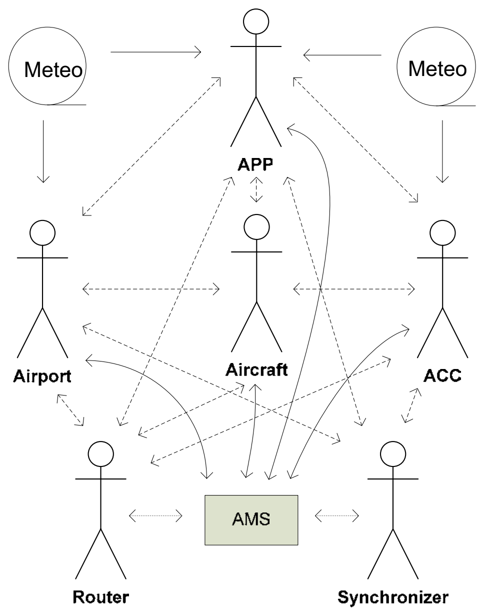


Fig. 6: Interaction diagram of ATCS agents, including data sources and communication elements

The communication between agents is performed through aeronautical messages. The messages standards needed in this application environment seem to be compatible with the message set currently normalized by ICAO (see [9]), which includes ADS, ADS-B and CPDLC, and with the well-established ACARS protocol.

VII. SIMULATION SCENARIOS

The Multi-Agent environment described in the previous section is designed to be simulated. When putting all agents to interact in this simulation, two opposite scenarios, and some intermediate combinations, may be applied:

- **Scenario I**: Approach Control (APP) and Area Control Centers (ACC) give a series of local instructions, through the time, for each aircraft arriving into the TMA, in order to get an arrival sequence in an iterative feedback fashion;
- **Scenario II**: ATMS provide a global and complete solution of trajectories, free of conflict, for all aircraft crossing the outer meter arc. In this case, manual adjustment instructions would be exceptions.

It is important to mention that both scenarios are subject to random disturbances, which make them more similar to the real system's behavior, which by its turn presents unexpected occurrence of events, such as changes in the atmospheric conditions. The agents intercommunicate by means of standardized data link messages, as defined by ICAO [9], in distinct protocols, following the purpose of the specific communication.

A. Uncertainties Models

A real air traffic scenario has many uncertainties. Two basic types of uncertainties may be identified in the system:

- **Uncertainty Type I**: they occur before the take-off, due to adverse meteorological situation in the airports, security procedures, logistic factors, equipment failures, and cascade delays. All these factors cause unexpected delays in the take-off schedule. This type of uncertainty may be modeled by an aggregate time distribution composed of many distributions such as passenger delay, maintenance time, etc;
- **Uncertainty Type II**: they occur during the flight, due to the winds or bad weather cells, and also cause unexpected variations in the speed or trajectory of the aircraft. This type of uncertainty will be taken into account in the conflict detection, and in the determination of the optimal arrivals sequence, by the AMAN. It can be modeled as a Gaussian distribution for the aircraft speed vector [6].

B. Computing resources

A set of software agent modules is being developed to perform the simulation of the scenarios above described. Following the Multi-Agent approach, each agent may run in a separate hardware, as it is usual in the air traffic system. As there will be a great volume of information to be exchanged between agents, and from databases and data sources to agents, the simulation will require high-performance data network linking the agents. A suggested architecture for the architecture of this simulation environment can be found in the reference [11].

VIII. FINAL REMARKS

Simulation has an important role in the definition of technologies and concepts to be used in control systems, and in the way of using them. In the Air Traffic System, the simulation is needed in an even greater extent, due to its high complexity, due to the high cost of doing experiments with a great number of human participants, the high cost of spending aircraft time and fuel, and due to the risks related to the experiments. Besides that, the simulation environment organization in itself may provide concepts that help in the real world air traffic efficiency increase.

This research work is ongoing in some parallel research tasks, such as agents interaction, database development, efficiency and safety metrics, which are planned to converge and provide qualitative and quantitative results in next stages of this project.

ACKNOWLEDGEMENT

The authors thank the support being provided by FDTE and CAPES in form of scholarships to the students involved in this project.

GLOSSARY

4-D	Four-Dimension
ACARS	Aircraft Communications Addressing and Reporting System
ACC	Area Control Center
ADS	Automatic Dependent Surveillance
ADS-B	ADS in Broadcast
AMAN	Arrivals MANager
AOC	Airline Operations Center
APP	Approach Control
ARTCC	Air Route Traffic Control Center
ATCS	Air Traffic Control System
ATM	Air Traffic Management
ATMS	Air Traffic Management System
ATN	Aeronautical Telecommunications Network
CDG	Charles de Gaulle Airport
CDM	Collaborative Decision Making
CPDLC	Cockpit-Pilot Data Link Communication
ETA	Estimated Time of Arrival
FAA	Federal Aviation Administration
FMS	Flight Management System
ICAO	International Civil Aviation Organization.
McTMA	Multiple-Center Traffic Management Advisory
PAX-Km	Passenger-kilometer
PAX-Mile	Passenger-nautical-mile
RVSM	Reduced Vertical Separation Minimum
STA	Scheduled Time of Arrival
TMA	Terminal Maneuvering Area
TMA-SC	Traffic Management Advisory – Single Center
TMC	Traffic Management Coordinator
TRACON	Terminal Radar Approach Control
ZFW	Ft. Worth ARTCC

REFERENCES

[1] CARE/ASAS Activity 3: Airborne Separation Minima - Final Report. CARE/ASAS/Sofréavia/01-018 - Version 2.1 - April 3, 2002.

[2] CDM Scenarios - European CDM. Web site: http://www.euro-cdm.org/scenarios.html. Accessed in July, 2004.

[3] CIP 2005. National Airspace System - Capital Investment Plan Fiscal Years 2005 - 2009. Federal Aviation Administration, 2004.

[4] Correa, M.; Camargo Jr., J. B.; Siewerdt, E. Collaborative Decision Making - Uma Abordagem Multi-Agentes para Tomada de Decisão em Cenários Complexos. III Simpósio Brasileiro de Pesquisa em Transporte Aéreo. São José dos Campos, December, 2004.

[5] Center TRACON Automation System. Web site: http://www.ctas.arc.nasa.gov/project_description. Updated in February, 2003. Accessed in October, 2004.

[6] Dimarogonas, D. V.; Kyriakopoulos, K. J. Inventory of Decentralized Conflict Detection and Resolution Systems in Air Traffic. European Comission - WP6, Hybridge Project Deliverable Number D6.1. 2003.

[7] FAA Aerospace Forecasts - Fiscal Years 2004-2015. U.S. Department of Transportation. Federal Aviation Administration, Office of Aviation Policy & Plans, March 2004.

[8] ICAO International Standards and Recommended Practices - Air Traffic Services, Annex. 11. ICAO, 12-th edition, 1998.

[9] Manual of Air Traffic Services Data Link Applications. First Edition - ICAO Doc 9694-AN/955, 1999.

[10] ICAO Circular 281. Outlook for Air Transport to the Year 2010. International Civil Aviation Organization, 2000.

[11] Oliveira, I. R., Cugnasca, P. S., Lubacheski, F. A., Sato, L. M., Correa, M. A., Siewerdt, E. Uma Arquitetura Lógica para o Auxílio à Tomada de Decisão em Gerenciamento de Tráfego Aéreo. In the proceedings of IV Simpósio de Transporte Aéreo, São José dos Campos, Brazil, 2005.

[12] Jennings, N.R. Coordination techniques for distributed artificial intelligence. In: O'Hare, G. M. P.; Jennings, N. R. (Eds.). Foundations of distributed artificial intelligence. New York: John Wiley & Sons, 1996, p.187- 210.

[13] Landry, S. et al. Distributed Scheduling Architecture for Multi-Center Time-Based Metering. AIAA Aircraft Technology. Integration and Operations (ATIO) Conference, Denver, CO, November 17-19. 2003.

[14] LEONARDO. Identification of Indicators and Metrics. Linking Existing ON Ground, Arrival and Departure Operations, Deliverable D2.1. Eurocontrol, 2003.

[15] Nova Circulação Aérea Geral, Fase 2. Comando da Aeronáutica, Depto. de Controle do Espaço Aéreo - DECEA. 10 de junho de 2004. Web site: http://www.decea.gov.br. Accessed in October, 2004.

[16] Ogata, P. H. Avaliação do Perigo de Colisão Entre Aeronaves em Operação de Aproximação em Pistas de Aterrisagem Paralelas. Master Dissertation, Escola Politécnica da Universidade de São Paulo, 2003.

[17] Oliveira, I. R. et al. An Evolutive Environment for the Development of Free-Flight. In: Spitzer, C., Schmocker; U., and Dang, V. N.: Proceedings of PSAM - International Conference on Probabilistic Safety Assessment and Management, 2004, Volume 1, pp. 513-518. Springer-Verlag, Berlin, 2004.

[18] Piduch, J. M. Orchestrated Approach. Air traffic management directory 1999-2000. Euromoney Institutional PLC, 2000.

[19] Review of the General Concept of Separation Panel, Sixth Meeting, Montreal, 28 November, ICAO Doc 9536, RGCSP/6, Volumes 1 and 2, 1988.

[20] System Development, ATC Systems. Web site http://www.sofreavia.com/pages/projects/index.html. Accessed in October, 2004.

[21] Medium-Term Forecast of Annual Number of IFR Flights (2002 - 2009). Eurocontrol - Air Traffic Statistics and Forecasts (STATFOR), Volume 1, March 2002.

[22] VAN ES. NLR's Air Safety Database - National Airspace Laboratory, The Netherlands, 2001. Apud. BLOM, H. A. P. Hybrid stochastic dynamical modeling of accident risk in air traffic. Summer Pedagogical School in Hybrid Systems, Lecture 3. University of Patras, Greece, 22-24 September 2004.

[23] Wooldridge, M.; Jennings, N.R. Intelligent agents: Theory and practice. The Knowledge Engineering Review, 10 (2):115-152, 1995.

[24] Wooldridge, M. Intelligent Agents. In: Weiss, G. (Ed.) Multiagent Systems - A Modern Approach to Distributed Artificial Intelligence. MIT Press, 1999.

[25] Oliveira, I. R.; Cugnasca, P. S. Arrivals Management Using Anticipated Information and Automated Re-Programming of FMS. In: Proceedings of the 6th Air Transportation Research Society World Conference, Rio de Janeiro, 2005.

An Application for Detecting Potential Traffic Conflicts in Areas with Unreliable ATC

Siegfried Bauer, Nima Barraci, Uwe Klingauf
Institute of Flight Systems and Automatic Control
Technische Universität Darmstadt, Germany
Email: {bauer,barraci,klingauf}@fsr.tu-darmstadt.de

Mark Hartmann
Deutsche Lufthansa AG
Frankfurt am Main, Germany
Email: markhartmann11@web.de

***Abstract*— The quality and capability of Air Traffic Services varies from region to region. The International Air Transport Association (IATA) as well as the International Civil Aviation Organization (ICAO) both address this situation with the *Inflight Broadcasting Procedure* (IFBP) and the *Traffic Information Broadcasts by Aircraft* (TIBA) respectively. Both procedures require the flight deck crew to listen to and perform regular transmissions which include – amongst other – information about their position, direction and flight level. These transmissions have to be used by all flight deck crews in the vicinity to derive a mental picture of the surrounding traffic, as well as a picture of the future flight paths of other aircraft in order to prevent hazards to ensure safe operations. This has to be done "by hand".**

This paper discusses an approach to reduce the effort for making this kind of estimations by supporting the flight deck crew with a computer based "stand alone" application, which can be used independently of the aircrafts current equipage and does not require the establishment of new procedures. In the course of this discussion different conflict-detection approaches as well as different depiction possibilities will be presented and discussed.

I. Introduction

In order to cope with „*areas of the world where poor or unreliable ATS communications present a hazard to safe operations*" [1], the International Air Transport Association as well as the International Civil Aviation Organization established the Inflight Broadcasting (IFBP) and the Traffic Information Broadcasts by Aircraft (TIBA) procedures. Since both procedures have been harmonized in 1987 this paper will refer to IFBP without limiting its scope.

A. Allocation of duties

The duties of Air Traffic Control (ATC) and flight deck crew are usually clearly separated.

ATC's responsibility is to guarantee the safe and economic *air traffic flow*. As a consequence *strategic planning* is in the foreground of ATC's work. In order to fulfill his duty, an Air Traffic Controller shall have a *complete* overview of the airspace, which is the main focus.

On the other hand, the responsibility of the flight deck crew is to guarantee a safe, on-schedule and economic *flight*. The primary tasks for the flight deck crew are to *aviate*, *navigate* and *communicate*.

Normally the pilot flying (PF) decides how and where the aircraft has to be steered. The pilot not flying (PNF) gathers the information required and requested by the PF (*support*) and checks (*controlling*) through independent thought-trains the actions and decisions of the PF. To fulfill their duties, the flight deck crew shall have the overview of the *surrounding* airspace and the focus is clearly on the own aircraft.

B. Shifting of duties

The reason for the establishment of the IFBP was to guarantee safe operations in areas, where it is not possible for ATC – be it because of lack of technical equipment or not enough staff – to fulfill its duty, or where the airspace in general is considered uncontrolled. This leads to a situation, where the flight deck crew cannot count on ATC anymore. In these areas, the pilot not flying (PNF) takes over parts of ATC's tasks. This means, that *strategic* planning now is an additional part of the PNF's tasks. To perform his duty – just like ATC would – the PNF now needs a complete overview of the airspace. However with the additional constraint, that he does not have the sophisticated technical means available to do so. Therefore, the PNF has to rely on regular broadcasts, which are required by the IFBP. The scope of the PNF to *support* and *monitor* the PF has shifted from *support* and *controlling* to *support*, *c*ontrolling as well as *monitoring* the airspace within radio range.

This can be considered as undesirable *quasi-autonomous* flight.

II. Current workflow

This section will give an example taking a simple transmitter-receiver scenario as basis. In this scenario, SA123 will be the aircraft broadcasting an IFBP message, while LH456 is the receiving aircraft (one out of several).

A. Transmitting aircraft

IFBP requires the flight deck crew to broadcast a procedure compliant message in a predefined sequence 10 minutes prior to reaching a waypoint. Assuming, that SA123 is southbound from London Heathrow to Johannesburg on airway UR985, a possible IFBP message is:

All stations, this is SA123, south-bound, from Heathrow to Johannesburg on UR985. Position ELO at 2335. Estimate position RIMEL at 2350. SA123, FL290.

Although actually broadcasted IFBP messages do not always comply to the IATA published procedure, the above presented message shall serve as basis for the discussion in this paper. The length of such a message is about 27 seconds, but can be from 20 seconds to one minute and more depending on the current workload in the broadcasting aircraft.

The tasks of the PNF on the transmitting aircraft can be summarized as follows:

- Check the distance to next waypoint
- If the distance to next waypoint is ten minutes or less and the IFBP message has not yet been broadcasted continue with
  - Wait for the opportunity to broadcast the message
  - Collect the information for the IFBP message
  - Broadcast the message

It should be pointed out, that this procedure is supposed to be executed in parallel to the ordinary PNF's tasks as well as to the task of receiving IFBP messages, as described in the next subsection.

The problems that arise for the transmitting aircraft can be summarized as following:

1) Monitor the distance (time) to the next waypoint
2) Collect IFBP relevant information
3) Broadcast the IFBP message according to required format

B. Receiving Aircraft

The IFBP also requires the PNF to constantly monitor all incoming messages and classify them whether they pose a *strategic threat*[1] or an *imminent threat*[2] or no threat at all.

In this scenario, LH456 receives the above IFBP message from SA123. The workflow when receiving an IFBP broadcast can be summarized in the following way:

- Listen to the broadcast and take notes
- If (part of the) message has not been understood
  - Consultation with the PF
  - Communicate with the transmitting aircraft
- Search for the referred airway and waypoint of the transmitting aircraft on the flight chart
- Calculate the probability of a conflict
- In case of conflict: perform corrective action

C. Problems arising

Before going into more detail regarding the possible situations which depend on LH456's position and course, the problems that arise up to now are numerous.

First of all, the amount of data received clearly exceeds the short-term memory capability. In case, the information has been noted properly, the amount of time needed to handle the received information properly is quite long (up to several minutes) which leads to a situation, where the PNF cannot handle additional incoming messages. With rising frequency of messages[3] the capacity and willingness to handle new messages decreases. On flights, where three crew members (two flying, one resting) are mandatory, the problem that no standardized way of denoting the information exists becomes evident. When passing the information of probable conflicting targets to the next crew member, very much effort needs to be spent in order to fully grasp the situation of the surrounding traffic.

The problems that need to be addressed in order to reduce pilot workload when receiving an IFBP message can be summarized as

1) Correct memorizing and noting of the message
2) Search for the airway and waypoint on the flight chart
3) Calculate the probability of a conflict
4) Resolve the conflict

III. Application support

The idea behind a computer based IFBP application is to address the above mentioned problems and difficulties in order to *reduce* pilot workload and *enhance* the situational awareness. The primary target is to assist the *receiving* aircraft's flight crew in coping with their duties.

Thus, a computer based application should

1) Support the flight deck crew in memorizing the IFBP message
2) Ease the search for airway and waypoint
3) Automatically calculate the probability of a conflict.

The secondary target is to ease the duties of the *sending* aircrafts flight crew.

Thus, a computer based application should

1) Automatically derive the aircrafts position from the flight plan and time
2) Advice the flight deck crew to broadcast an IFBP message when required
3) Depict the information on the user interface according to the predefined IFBP message sequence

IV. Application Functionality

To meet the above defined requirements the application addresses some of the basic working steps that have to be performed regularly during the IFBP procedure. The goal of the application is to reduce the time needed for each working step, so as to reduce the possibility of estimation errors and to get more precise results.

The first thing to do when an IFBP message arrives is to understand and memorize its content. This is supported by the automatic audio recording of arriving messages. The crew is then able to get the required information from the recorded message. There is no need to memorize the content. The crew can even replay the recording several times if problems in understanding occur. Another advantage is that the time when

[1]Scope: Several hours

[2]Scope: Up to ten minutes

[3]For the South-African route the highest peak is normally at N'Djamena (the capital of Chad) $\pm$ 10^o latitude depending on departure time on airways A403, G660, A607, M731, G857, G854 where the south-bound traffic meets the north-bound traffic (as well as east/west Haj traffic)

a call is transmitted is decoupled from the time when it has to be worked on.

After the call has been received the crew usually takes some notes and starts to look for the airway and next waypoint on the flight-chart. The duration of this task depends on the knowledge of the crew concerning the current area of operation. A problem arises if the name of a waypoint for example is not understood properly e.g. due to noise during the transmission. The IFBP application presents a list of possible airways to the flight-crew and based on their choice of the crew a number of waypoints are preselected. So instead of searching for positions of airways or waypoints the crew is able to select from a manageable number of items.

To estimate if there could be a conflict with other aircraft the crew has to identify a point where both tracks will cross, then estimate the own flight time to that crossing and compare it to the estimated fly-over time of the other aircraft. These estimations are based on mental arithmetic which can lead to errors, especially under high workload. A computer based application will be faster, more precise and will relieve the workload of the crew, assisting them to return to their normal way of operating.

In order to draw attention to possible conflicts, those aircraft that will pass a common position within a defined boundary will be highlighted.

With this functionality the application eases some basic working steps during the IFBP procedure and additionally gives a complete standardized air traffic situational overview which is necessary at all times as well as when control is handed over to another member of the flight-deck crew.

A. *Ergonomic advantages*

Taken the relief of the PNF of duties which are not in line with his qualification and the more accurate computer based calculation and prediction of possible conflicts, ergonomic advantages can be foreseen.

On the one hand, *flexibility* of flight planning for the flight deck crew grows. It is now possible for the crew to foresee a conflict in the future with more accuracy and take this into account for the own strategic planning. The possibility to record and replay incoming messages, relieves the PNF of the duty, to concentrate on an externally triggered event. Therefore he can still monitor and support the PF and come back to the IFBP task, when the situation allows it.

Since the effort asked of the PNF to perform tasks which are not in line with his duties, the overall PNF's workload can be reduced. This follows a raise in safety.

Summarizing the foreseen advantages of using such an application, three main factors need to be clearly pointed out:

- The attachment of the PNF to IFBP tasks is reduced
- Capacity is being freed for the PNF's core duties
- Instead to react, it is now possible to actively act again

V. Application Architecture

A. *Layers of the Application*

The application is divided into different layers which interact only via interfaces. This way each layer can be developed independently and different presentation and calculation approaches can be implemented easily.

The bottom layer is the data-layer that contains information about the airway structure (waypoints, legs and airways). On top of the data-layer the calculation-layer resides which is responsible for all calculations e.g. time differences at waypoints or intersections. The information that results from this calculation is presented to the flight deck crew through the user-interface / presentation-layer.

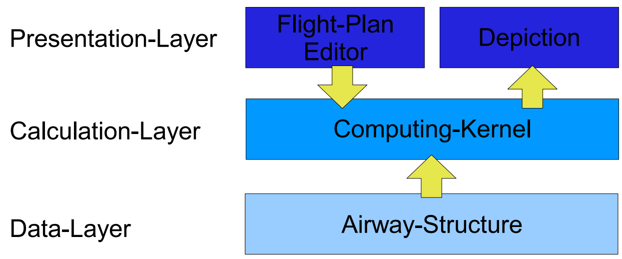


Fig. 1. Application architecture

B. *Data-Layer*

The data-layer contains the data needed for performing calculations to detect possible conflicts. Therefore information about waypoints (position and name), legs (connecting which waypoints, length) and airways (name, containing which legs) are included. This data is made accessible to the calculation-layer through specified function calls.

C. *Calculation-Layer*

Based on the own flight-plan which is entered through the user-interface as well as the information about other aircraft (see subsection V-D) potential conflicts with other aircraft are detected.

At first a check is performed if the own track will cross the track of the currently monitored aircraft. If so, the point of crossing is calculated. This may either be a waypoint or an arbitrary position on a leg. Then the own estimated time to that crossing is calculated and compared to the estimated time of the other aircraft. The time difference is passed on to the presentation-layer in order to be depicted. This calculation is performed for every aircraft from which a radio call has been received/entered.

D. *User-Interface / Presentation-Layer*

The user-interface / presentation-layer consists of two parts: a flight-plan editor and an area for depiction of possible conflicts. The flight plan editor is used to insert the own complete flight-plan. This can be done before take-off so that in flight no interaction concerning the own flight-plan is needed.

The computed result after entering a received broadcast is currently represented in a textual way using a table (see Figure 2). Every row presents one aircraft and contains beside call sign and flight-plan information about the position of crossing tracks and the time difference at a crossing.

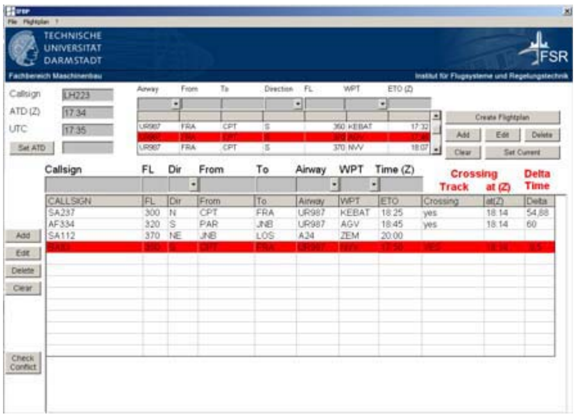


Fig. 2. IFBP Application User Interface

E. Technical Environment

The IFBP procedure is defined by the information that shall be included in the calls and the instants in time when these calls shall be sent out. So the definition includes only the information channel. The following working steps to create a mental picture of the situation are completely left to the flight-crew. As long as the information channel is not involved there are no restrictions. The presented application is a pure onboard supporting tool that concerns only the working steps in the cockpit. The only technical requirement is a computer in the cockpit. No changes in the communication infrastructure are needed. This leads to the advantage that no system wide changes are necessary, which would be costly and would take long time to be implemented as many parties are involved. The operation of the presented application can be decided on the airline level and every airline will get all presented advantages independently of other airspace participants.

A more advanced integration can be achieved, if the application was implemented as an EFB application. This would mainly concern the possibilities of data acquisition and data representation (see section section VI). Through the access to the aircrafts data bus the usability and functionality of the application can be improved. In this case the user interface will have to meet the EFB requirements by then which is easy to achieve as the interface is implemented as a module (see Figure 1).

VI. Future Extension

A. Acquisition of data

As an extension to increase the usability and precision the application can be connected to the aircrafts data bus. This way the application would be able to get “real time“ data directly. So the flight-deck crew does not need to enter the flight-plan via user-interface and flight-plan changes are automatically kept consistent with the FMS. The application will be able to access the current system time which is necessary to continue a consistent tracking of the next waypoint and the estimated fly over time. At last the access to precise position data enables the usage of a moving map depiction.

B. Automatic generation of IFBP calls

All pieces of information that are required in an IFBP call are included in the application. So the application is able to generate a complete IFBP call which can then be represented to the flight-deck crew for transmitting or even a direct voice output can be generated. This would ease workload of the flight-deck as every outgoing call takes about 27 seconds of attention.[4]

C. Calculation

Within the IFBP, in order to reduce the risk of highly accurately navigating aircraft colliding precisely on the airway, offset flying – flying parallel to the airway by one or two miles to the right – is recommended pilot procedure. If only the airway structure and the time differences at points of crossings are taken into account this does not properly represent aircraft which are flying off-track particularly if airways are crossing at acute angles. Instead of using distinct positions during calculation circular areas can be used. This properly models navigation inaccuracies, estimation errors in IFBP calls or aircraft flying off-track. If more than one area around each aircraft is defined this can be used for an automatic classification of potential conflicts.

Currently a crossing point is calculated based on the airway structure and a potential conflict is detected using time estimations to crossing points. To cover lateral deviations and acute angles at crossing points a safety buffer can be added to these calculations. This could be done by using a circle around aircraft positions and using them to detect potential conflicts by doing intersection calculations (see Figure 3).

D. User Interface / Presentation

As the user interface is implemented as an independent module different approaches of data input and data presentation can be integrated easily. The current text oriented data representation can for example be replaced with a moving map depiction that shows the own aircraft position, the airway structure and positions where potential conflicts can occur. This depiction of course requires the connection to the aircrafts data bus from where the own position is acquired and continuously updated. If a reliable source for the position of other aircraft is available, their position can be depicted as well. This would directly depict the current situation together with possible future conflicts, which would increase the flight crews situational awareness.

[4]A typical FRA-CPT flight will have 30-40 wpt’s and there are about 15 major carrier airlines on the north south routes every night.

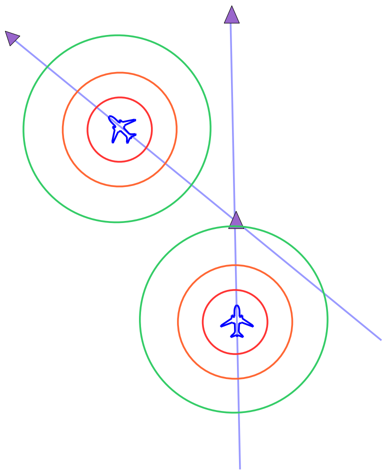


Fig. 3. Conflict detection using position areas

VII. Outlook - Conflict Resolution and ASAS

The IFBP application as presented in this paper leaves the resolution of a detected conflict to the crews of the participating aircraft. Closing the introductory scenario of *SA123* and *LH456*, two possible conflict situations – *crossing conflict* and *merging conflict* – will be taken as an example in order to step towards ASAS (Airborne Separation Assisstance Systems) applications.

A. Crossing conflict

The *crossing conflict* scenario relates (see Figure 4) to the ASAS application Crossing & Passing (ASPA-C&P) as stated in [4, Annex 5.8].

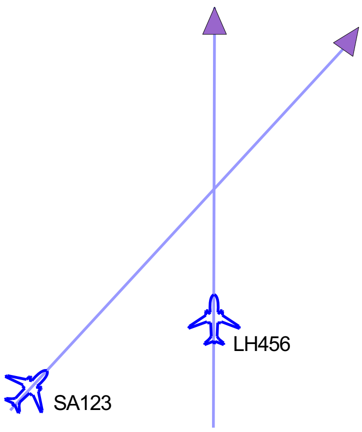


Fig. 4. Crossing & Passing Scenario

The waypoint *KANO* – located at the city of Cano in Nigeria – serves as a good example for ASPA-C&P. From *KANO* four airways – i.e. *UA615*, *UA604*, *UR986* and *UG858* – come together in very close proximity. The area around *KANO* is well known for bad meteorological conditions which forces flight crews to leave their airway and takes them generally at minimum 10 nm off their track. Assuming, that *SA123* – flying south-bound on *UA604* at FL300 and approaching *KANO* – has to take a detour to the west of at least 10 nm. For *LH456* which is on *UA615* at FL290 and also approaching *KANO*. Assuming, that both aircraft are close to *KANO*, a detour taken by South-African could lead to a crossing of Lufthansa's flightpath as illustrated in Figure 4. By applying ASPA-C&P procedure to *SA123*, a possible under-run of minimum clearance distance could be prevented.

B. Merging conflict

Like the preceding scenario, the *merging conflict* depicted in Figure 5 relates to an ASAS application, this time Sequencing & Merging *ASPA-S&M* [4, Annex 5.6]. *LH456* is *north-bound* on airway *UA607* and approaches waypoint *N'Djamena* in *10 minutes* on *FL300* while *SA123* approaches the same waypoint on *UG857*. *SA123* is also *north-bound* on *FL300* and approaches *N'Djamena* in *7 minutes* (i.e. *SA123* made the IFBP call-out 3 minutes in advance of *LH456*). Both aircrafts have filed a flight-plan continuing after *N'Djamena* north-bound on *UA403*. In this scenario the flight paths of both aircraft merge at *N'Djamena*, which constitutes the classical ASPA-S&M. Since *LH456* received South-African's IFBP call-out 3 minutes in advance, the application could automatically trigger an ASPA-S&M procedure which would require *LH456* to merge behind *SA123*.

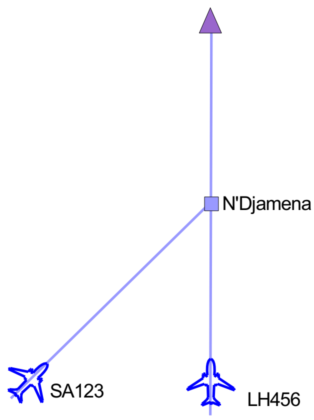


Fig. 5. Sequencing & Merging Scenario

C. Scope

The above drafted examples only take two participating aircrafts into account. Research needs to be conducted in the area of self-separation algorithms with multiple aircraft in order to go this next step towards ASAS. The IFBP application could serve as additional test bed equipment during simulator experiments and support concept validation [5, Section 2.3.5].

VIII. Conclusion

Due to the increasing traffic Africa the IATA IFBP procedure demands a considerably amount of the flight crews' attention and leads to a high workload. The necessary working steps to handle incoming and outgoing IFBP messages are time consuming and recur continuously. Some of them require the intimate knowledge of the area and routes to be executed in a reasonable time. A computer application can eliminate many of these working steps and simplify others. This obviously would lead to a decreasing workload and contributes considerably to aviation safety.

IX. Summary and Forecast

In areas with poor ATC capabilities the pilot not flying (PNF) takes over duties which are usually allocated to air traffic controllers. This increases the flight crews workload. The presented computer application simplifies and eliminates many working steps that have to be induced by this situation. This allows the flight crew to focus on their core duties. To increase the usability and functionality of the application an integration into the flight deck environment is proposed.

References

[1] ATS CONTINGENCY PLANNING ISSUES, *Interregional coordination and harmonization mechanism - Y2K date change preparations*, Third meeting of the ALLPIRG/ADVISORY GROUP (Montreal, 6-8 April 1999)

[2] IATA In-flight Broadcast Procedure (IFBP)
Asia Pacific (ASPAC) Region
IFBP ASPAC Aug 2003
`http://www1.iata.org/WHIP/_Files/WgId_0019/IFBP_ASPAC.pdf`

[3] International Air Transport Association Regional Office Infrastructure - Africa and Indian Ocean
IFBP AFI October 2002
`http://www.ibac.org/Bulletins/IFBP.pdf`

[4] CARE/ASAS Activity 5
Description of a first package of GS/AS applications
Version 2.2, September 30, 2002

[5] Safety and ASAS applications, Action Plan 1
FAA/EUROCONTROL Cooperative R&D
Version 4.1, February 12, 2004

Session 8

Safety

Free flight safety risk modeling and simulation

Henk A.P. Blom, G.J. (Bert) Bakker, Bart Klein Obbink and Margriet B. Klompstra

Abstract— **The basic notion of free flight is that aircrews obtain the freedom to select their trajectory including the responsibility of resolving conflicts with other aircraft. Under low en-route traffic loads there is general agreement that free flight can be safely applied. Under increasing traffic loads, however, the answer to this question becomes unknown. Free flight would change ATM in such a fundamental way, that one can speak of a paradigm shift and that comes with emerging behavior, i.e. novel behavior which is exhibited at the system-wide level and emerges from the combined dynamical actions and reactions by individual systems and humans that affect the operations. Because emerging behavior cannot be predicted from previous experience, we need a complementary approach in estimating the safety of free flight under relatively high traffic levels. This paper explains how recently developed methods in Petri net specification and sequential Monte Carlo simulation can be used to make progress in addressing this outstanding issue. The paper also presents the results of an initial application of these novel methods to a well developed autonomous free flight concept of operations.**

***Index Terms*—Sequential Monte Carlo simulation, Petri net modelling, Safety risk assessment, Safety-critical systems, Autonomous Free Flight**

I. INTRODUCTION

TECHNOLOGY allows aircraft to broadcast information about the own-ship position and velocity to surrounding aircraft, and to receive similar information from surrounding aircraft. This development has stimulated the rethinking of the overall concept for today's Air Traffic Management (ATM), e.g. to transfer responsibility for conflict prevention from ground to air. As the aircrews thus obtain the freedom to select their trajectory, this conceptual idea has been called Free Flight [1]. It changes ATM in such a fundamental way, that one could speak of a paradigm shift: the centralised control becomes a distributed one, responsibilities transfer from ground to air, fixed air traffic routes are removed and appropriate new technologies are brought in. In free flight, each individual aircrew has the responsibility to timely detect and solve conflicts, thereby assisted by navigation means, surveillance processing and equipment displaying conflict-solving trajectories. Due to the potentially many aircraft involved, the system is highly distributed. This free flight concept definition leaves open many challenges in developing adequate procedures, algorithms, equipment performance requirements, and has motivated the study of multiple Free Flight operational concepts, implementation choices and requirements, e.g. [2]-[6]. Crucial in this process is to learn understanding how to optimize free flight designs for safety and capacity [7].

The aim of this paper is to study the safety risk assessment of en-route free flight operations through modeling and Monte Carlo simulation. In [8], such a study has been performed for free flight equipped aircraft that are obliged to remain flying within a conventional fixed route structure. The current paper, however, studies a true free flight concept of operations, i.e. one without using fixed route structure. The free flight concept we identified for this has been developed for air traffic in the Mediterranean area [9]. For short we refer to this operational concept as Autonomous Mediterranean Free Flight (AMFF). We illustrate an advanced model specification and sequential MC simulation approach towards the assessment of collision risk of AMFF operation under relatively high traffic density.

For advanced air traffic operations, [10] gives a nice illustration how statistical data in combination with a fault tree of the functionalities of the advanced operation can serve to predict how reliability of free flight supported systems impact contributions to collision risk of an advanced operation [11]. Through an example it is shown how this allows improvement of the design, such that the reliability-implied contribution to collision risk can be lowered to a desired value.

Also following a fault tree approach, the safety of the AMFF concept of operations has been assessed [12], including real-time simulations of non-nominal conditions [13]-[15]. The results obtained show that application of AMFF seems feasible for accommodating low en-route traffic conditions over the Mediterranean. In order to assess whether AMFF can safely accommodate higher traffic levels, [12] recommends to use a more advanced safety risk assessment approach that considers complex situations involving dynamic interactions between multiple human actors and other systems.

The recommendation by [12] concurs well with the explanation by [16] that the key difficulty of evaluating advanced operations is to include emergent behavior, i.e. novel behavior which is exhibited at the system-wide level and emerges from the combined dynamical actions and reactions by individual systems and humans within the system. This emergent behavior typically cannot be foreseen and evaluated

Manuscript received February 26, 2006 and revised April 26, 2006. This work was supported through the European Commission project HYBRIDGE, IST-2001-32460.

Henk Blom, Bert Bakker, Bart Klein Obbink and Margriet Klompstra are with National Aerospace Laboratory NLR, Amsterdam, The Netherlands, blom@nlr.nl (phone: +31.20.5113544, fax: +31.20.5113210), bakker@nlr.nl; bklein@nlr.nl; klompstr@nlr.nl.

by examining the individuals' behavior alone. [16] explains that agent based simulation is able to predict the impact of revolutionary changes in air transportation; it integrates cognitive models of technology behavior and description of their operating environment. Simulation of these individual models acting together can predict the results of completely new transformations in procedures and technology. Their MC simulations reach up to the level of novel emerging hazardous events. For safety risk assessment however, it is required to go further with the MC simulations up to the level of emerging catastrophic events. In en-route air traffic these catastrophic events are mid-air collisions. The approach described in this paper is an example of the latter approach to estimate such a difficult metric of collision risk between aircraft by the use of advanced approaches in Petri Net modeling and Monte Carlo simulation.

The paper is organized as follows. Section II provides a brief overview of the AMFF concept of operations selected for evaluation on collision risk and explains how this operational concept has been modeled in a specific Petri net formalism. Section III explains the sequential Monte Carlo simulation acceleration approach developed for assessing collision risk for the AMFF simulation model. Section IV presents results of Monte Carlo simulations performed for three AMFF scenarios. Section V discusses the results obtained.

II. DEVELOPMENT OF PETRI NET MODEL OF AMFF

For the development of a Petri net model of an advanced operation, two key challenges have to be addressed: a syntactical challenge of developing a model that is consistent, complete and unambiguous, and a semantics challenge of developing appropriate human cognition performance models. This section explains how the syntactical challenge has been addressed. For the mechanism to manage the semantics challenge, we have followed the approach studied and developed in, e.g. [17]-[21]. The explanation of how this cognitive human performance modeling approach has been applied to AMFF falls outside the current paper's scope.

A. AMFF operation

For a complete description of the AMFF operational concept we refer to [9], [14]. In addition, [22] describes the background of the AMFF design philosophy. A practical implication was to avoid much information exchange between aircraft and to avoid dedicated decision-making by artificial intelligent machines. Although the conflict detection and resolution approach developed for AMFF has its roots in the modified potential field approach [4], there are significant differences. The main difference is that conflict resolution in AMFF is intentionally designed not to take the potential field of all aircraft into account. The resulting AMFF design can be summarized as follows:

- All aircraft are supposed to be equipped with Automatic Dependent Surveillance-Broadcast (ADS-B), which is a system that periodically broadcasts own aircraft state information, and continuously receives the state information messages broadcasted by aircraft that fly within broadcasting range (~ 100 NM).
- In order to comply to pilot preferences, conflict resolution algorithms are designed to solve multiple conflicts one by one rather than according to a full concurrent way that can be handled by the modified potential field approach [4].
- Conflict detection and resolution are state-based, that is: intent information, such as information at which point surrounding aircraft will change course or height, is supposed to be unknown.
- The vertical separation minimum is 1000 ft and the horizontal separation minimum is 5 Nm. A conflict is detected if these separation minima will be violated within 6 minutes.
- The conflict resolution process consists of two phases. During the first phase, one of the aircraft crews should make a resolution maneuver. If this does not work, then during the second phase, both crews should make a resolution maneuver.
- Prior to the first phase, the crew is warned when an ASAS alert is expected to occur if no preventive action would be timely implemented; this prediction is done by a system referred to as P-ASAS (Predictive ASAS).
- Conflict co-ordination does not take place explicitly, i.e. there is no communication on when and how a resolution maneuver will be executed.
- All aircraft are supposed to use the same resolution algorithm, and all crew are assumed to use ASAS and to collaborate in line with the procedures.
- Two conflict resolution maneuver options are presented: one in vertical and one in horizontal direction. The pilot decides which option to execute.
- ASAS related information is presented to the crew through a Cockpit Display of Traffic Information (CDTI).

B. Stochastically and Dynamically Coloured Petri Net

The most advanced approaches that have been developed in literature to model accident risk of safety-critical operations in nuclear and chemical industries make use of the compositional specification power of Petri nets to instantiate a model, and subsequently use stochastic analysis and Monte Carlo simulation (e.g. [23]) to evaluate the model. Since their introduction in the 1960s, Petri nets have shown their usefulness for many practical applications in different industries (e.g. [24]). Various Petri net extensions and generalisations, new analysis techniques, and numerous supporting computer tools have been developed, which further increased their modelling opportunities, though falling short for air traffic operations. In order to capture the characteristics of air traffic operations through a Petri net, [25]-[27]

introduced Dynamically Coloured Petri Net (DCPN) and Stochastically and Dynamically Coloured Petri Net (SDCPN), and proved that there exists a one-to-one relation with the larger class of stochastic processes and analysis techniques needed for air traffic operations [28].

A Coloured PN is a Petri net with a colour attached to each token. Such a colour assumes values from a given set, and this value does not change as long as the token stays in its place. When a token is "transfered" from one place to another place, then the colour moves with the token to the next place and may also be updated. In a DCPN a colour value may evolve as the solution of an ordinary differential equation (ODE). In a SDCPN a colour value may evolve as the solution of a stochastic differential equation (SDE).

The specification of an SDCPN for a complex process or operation is accomplished in a compositional way [29]. It starts with developing relevant Local Petri Nets (LPNs) for each agent that exists in the process or operation (e.g. air traffic controller, pilot, navigation and surveillance equipment). Essential is that these LPNs are allowed to be connected with other Petri net parts in such a way that the number of tokens residing in an LPN is not influenced by these interconnections. We use two types of basic interconnection arcs between nodes and arcs in different LPNs:

- Enabling arc (or inhibitor arc) from one place in one LPN to one transition in another LPN. These types of arcs have been used widely in Petri net literature.
- Interaction Petri Net (IPN) from one (or more) transition(s) in one LPN to one (or more) transition(s) in another LPN.

In addition, a box is drawn around each LPN, and hierarchical interconnection arcs from or to an edge of an LPN box are defined to represent several arcs or transitions by only one arc or transition.

C. *Agents and LPNs to represent AMFF operations*

In the Petri Net modeling of AMFF operations for the purpose of an initial collision risk assessment, the following agents are taken into account:

- Aircraft
- Pilot-Flying (PF)
- Pilot-Not-Flying (PNF)
- Airborne Guidance, Navigation and Control
- Airborne Separation Assistance System (ASAS)
- Communication / Navigation / Surveillance

It should be noticed that our initial model does not yet incorporate Airborne Collision Avoidance System (ACAS), Airline Operations Centre (AOC), Air Traffic Control (ATC) and an environmental model.

Per agent, particular LPNs and IPNs have been developed and subsequently the interactions between these LPNs and IPNs have been specified. The listing of agents and LPNs is:

- Aircraft LPNs:
  - Type
  - Evolution mode
  - Systems mode
  - Emergency mode
- Pilot-Flying (PF) LPNs:
  - State Situation Awareness
  - Intent Situation Awareness
  - Goal memory
  - Current goal
  - Task performance
  - Cognitive mode
- Pilot-Not-Flying (PNF) LPNs:
  - Current goal
  - Task performance
- ASAS LPNs:
  - Processing
  - Alerting
  - Audio alerting
  - Surveillance
  - System mode
  - Priority switch mode
  - Anti-priority switch mode
  - Predictive alerting (of other aircraft)
- Airborne GNC (Guidance, Navigation, Control) LPNs:
  - Indicators failure mode for PF
  - Engine failure mode for PF
  - Navigation failure indicator for PF
  - ASAS failure indicator for PF
  - ADS-B receiver failure indicator for PF
  - ADS-B transmitter failure indicator for PF
  - Indicator failure mode for PNF
  - Guidance mode
  - Horizontal guidance configuration mode
  - Vertical guidance configuration mode
  - FMS flightplan
  - Airborne GPS receiver
  - Airborne Inertial Reference System (IRS)
  - Altimeter
  - Horizontal position processing
  - Vertical position processing
  - ADS-B transmission
  - ADS-B receiver
- Communication / Navigation / Surveillance LPNs:
  - Global GPS / satellites
  - Global ADS-B ether frequency
  - SSR Mode-S frequency

The actual number of LPNs in the whole model then equals $38N+3$, where N is the number of aircraft. In addition the number of IPNs equals $35N$.

D. *Interconnected LPNs of ASAS*

ASAS is modeled through the SDCPN depicted in Figure 1. The ADS-B information received from other aircraft is processed by the *LPN ASAS surveillance.* Together with the information about its own aircraft state information (from AGNC), the *LPN ASAS processing* uses this information to perform conflict detection and resolution functionalities. Subsequently, the *LPN ASAS alerting* and the *LPN P-ASAS processing* informs the PF and PNF through *ASAS audio* *alerting* about any aircraft that is in potential ASAS conflict with the own aircraft, and suggests resolution options including a prioritization. Three complementary LPNs represent non-nominal behavior modes, each combination of which has a specific influence on the ASAS alerting LPN:

- *ASAS system mode* may be working, failed or corrupted (failed or corrupted mode also influences the ASAS processing LPN).
- *ASAS priority switching mode*; under emergency, the PF switches this from “off” to “on”.
- *ASAS anti-priority switch*; this is switched from “off” to “on” when own ADS-B is not working.

Avoidance”, “Conflict Resolution” and “Horizontal Navigation”. For each of these goals, the pilot executes a number of tasks in a prescribed or conditional order, represented in the LPN “Task Performance”. Examples of such tasks are “Monitoring and Decision”, “Execution” and “Execution Monitoring”. If all relevant tasks for the current goal are considered executed, the pilot chooses another goal, thereby using his memory (where goals deserving attention might be stored, represented by the LPN “Goal Memory”) and the Human-Machine-Interface. His memory where goals deserving attention might be stored is represented as the LPN “Goal Memory” in Figure 2.

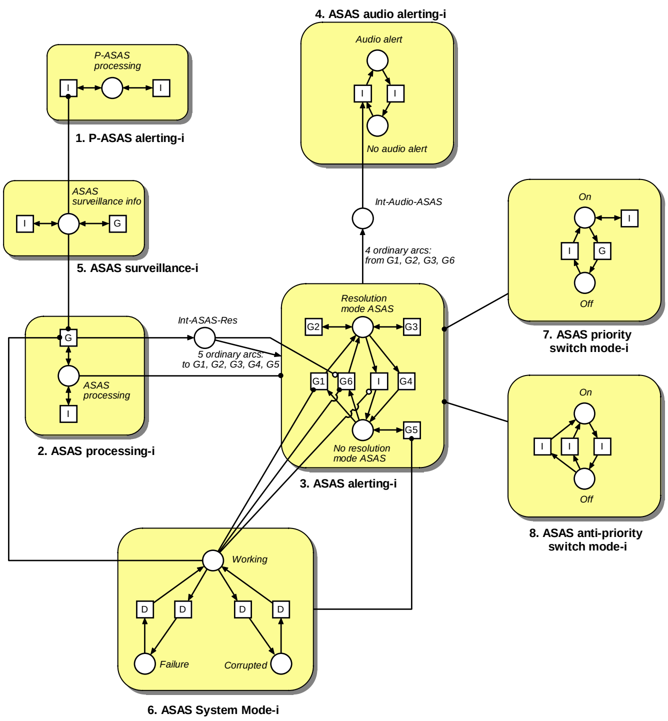


Figure 1. The agent ASAS in AMFF is modelled by eigth LPNs, a number of ordinary and enabling arcs and two IPNs (with one place each).

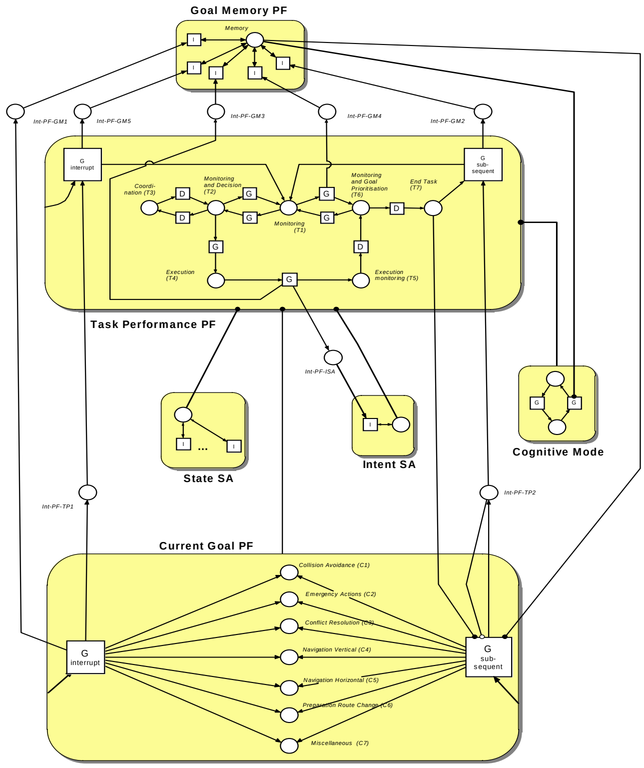


Figure 2. The agent Pilot-Flying in AMFF is modelled by six LPNs, and a number of ordinary and enabling arcs, and some IPNs, consisting of one place and input and output arcs.

E. Interconnected LPNs of Pilot Flying

This subsection illustrates the specific Petri Net model developed for the Pilot Flying. A graphical representation of all LPNs the Pilot-Flying consists of, is given in Figure 2. The Human-Machine-Interface where sound or visual clues might indicate that attention should be paid to a particular issue, is represented by an IPN that is not depicted in the Figure. Similarly, the arcs to or from any other agent are not shown in Figure 2. Because of the very nature of Petri Nets, these arcs can easily be added during the follow-up specification cycle. To get an understanding of the different LPNs, a good starting point might be the LPN “Current Goal” (at the bottom of the figure) as it represents the objective the Pilot-Flying is currently working on. Examples of such goals are “Collision

So, the LPNs “Current Goal”, “Task Performance”, and “Goal Memory” are important in the modelling of which task the Pilot-Flying is executing. The other three LPNs are important in the modelling on how the Pilot-Flying is executing the tasks. The LPN “State SA”, where SA stands for Situation Awareness, represents the relevant perception of the pilot about the states of elements in his environment, e.g. whether he is aware of an engine failure. The LPN “Intent SA” represents the intent, e.g. whether he intends to leave the Free Flight Airspace. The LPN “Cognitive mode” represents whether the pilot is in an opportunistic mode, leading to a high but error-prone throughput, or in a tactical mode, leading to a moderate throughput with a low error probability.

F. Model parametrization, verification and validation

The compositionally specified SDCPN model enables a systematic implementation, verification and validation of the resulting Monte Carlo simulator. This is done through the following systematic steps:

- Software code testing. This is done through conducting the following sequence of testing: random number generation, statistical distributions, common functions, each LPN implementation, each agent implementation, interactions between all agents, full MC simulation;
- Numerical approximation testing. This is needed to identify maximally allowable numerical integration step and minimally required number of particular MC simulations;
- Graphical user interface testing. This is to verify that the input and output of data works well;
- Parameterization. This is done through a search for literature and statistical sources, and complemented by expert interviews. The fusion of these different pieces of information is accomplished following a Bayesian approach;
- Initial model validation through studying MC simulator behavior and sensitivities to parameter changes under dedicated scenarios;
- Overall validation, which is directed to the evaluation of differences between model and reality and what effect these differences have at the assessed risk level [30], [31]. Statistical data collection and analysis, and active participation of operational experts is required.

The last step should start as early as possible, but should also continue throughout the operational concept development cycles. Typically, this way of working allows a significant improvement of the validation level of the simulation model per concept development and safety risk assessment cycle.

III. MONTE CARLO SIMULATION OF COLLISION RISK

The basic idea of assessing collision risk is to perform many Monte Carlo simulations with the SDCPN model, and while doing so, to estimate the collision risk by counting the number of collisions and dividing this by the number of simulated flight hours. Though this idea is simple, in order to make it work in practice, we need an effective way of speeding up the Monte Carlo simulation. This section describes the way we are doing this by extending the sequential MC simulation approach of [32], [33] to collision risk in air traffic.

A. Simulation to first moment of collision

In [27] it has been shown that an SDCPN model represents a stochastic differential equation (SDE) on a hybrid state space, driven by Brownian motion and Poisson random measure. In [34] it has been shown that under reasonable conditions (typically also adopted when specifying an SDCPN) the solution of this SDE is a strongly unique stochastic hybrid process $\{x_t,\theta_t\}$ which has mathematically properties that enable powerful stochastic analysis (semimartingale and strong Markov). In view of this, [35] has extended the approach of [32] to these stochastic hybrid processes. This allows us to study the speeding up of the AMFF simulation in this general setting, and avoids a need to dive into all kind of details of the AMFF model.

For the *N*-aircraft traffic scenario the process $\{x_t,\theta_t\}$ consists of components $x_t \triangleq \mathrm{Col}\{x_t^0,x_t^1,\ldots,x_t^N\}$ and $\theta_t \triangleq \mathrm{Col}\{\theta_t^0,\theta_t^1,\ldots,\theta_t^N\}$, x_t^i assumes values from $\mathbb{R}^{n_i}$, and θ_t^i assumes values from a finite set (M^i). $\{x_t^i,\theta_t^i\}$, $i=1,\ldots,N$, is the hybrid state process related to the *i*-th aircraft, and $\{x_t^0,\theta_t^0\}$ is the non-aircraft related hybrid state process. The total process $\{x_t,\theta_t\}$ is $\mathbb{R}^n\times M$ -valued with $n=\sum_{i=0}^{N} n_i$ and $M=\bigotimes_{i=0}^{N} M_i$.

In order to model collisions between aircraft, we introduce mappings from the Euclidean valued process $\{x_t\}$ into the relative position and velocity between a pair of aircraft (i,j). The relative horizontal and vertical positions are obtained through the mappings $y^{ij}(x_t)$ and $z^{ij}(x_t)$ respectively. The relative horizontal velocity and the vertical rate of climb/descent are obtained through the mappings $v^{ij}(x_t)$ and $r^{ij}(x_t)$ respectively. The relations between these position and velocity mappings satisfy:

$$dy^{ij}(x_t)=v^{ij}(x_t)dt \tag{1}$$

$$dz^{ij}(x_t)=r^{ij}(x_t)dt \tag{2}$$

A collision between aircraft (*i*,*j*) means that the process $\{y^{ij}(x_t),z^{ij}(x_t)\}$ hits the boundary of an area where the distance between aircraft *i* and *j* is smaller than their physical size. Under the assumption that the length of an aircraft equals the width of an aircraft, and that the volume of an aircraft is represented by a cylinder the orientation of which does not change in time, then aircraft (*i*,*j*) have zero separation if $x_t\in D^{ij}$ with:

$$D^{ij}=\{x\in\mathbb{R}^n;\ \left|y^{ij}(x)\right|\le(l_i+l_j)/2\ \text{ AND }\ \left|z^{ij}(x)\right|\le(s_i+s_j)/2\},\ i\ne j \tag{3}$$

where l_j and s_j are length and height of aircraft *j*. For simplicity we assume that all aircraft have the same size, by which (3) becomes:

$$D^{ij}=\{x\in\mathbb{R}^n;\ \left|y^{ij}(x)\right|\le l\ \text{ AND }\ \left|z^{ij}(x)\right|\le s\},\ i\ne j \tag{4}$$

Notice that in (4), D^{ij} still depends of (*i*,*j*). If x_t hits D^{ij} at time τ^{ij}, then we say a collision event between aircraft (*i*,*j*) occurs at moment τ^{ij}, i.e.

$$\tau^{ij} = \inf\{t > 0;\ x_t \in D^{ij}\}, \qquad i \neq j \tag{5}$$

Next we define the first moment τ^i of collision with any of the other aircraft, i.e.

$$\begin{aligned}\tau^i &= \inf_{j\neq i}\{\tau^{ij}\} = \inf_{j\neq i}\{t>0; x_t \in D^{ij}\}\\ &= \inf\{t>0; x_t \in D^i\}\end{aligned} \tag{6}$$

with $D^i = \bigcup_{j\neq i} D^{ij}$. At moment τ^i we can stop the simulation of the process $\{x_t^i, \theta_t^i\}$.

An unbiased estimation procedure of the probability $\mathbb{P}(\tau_m^i < T)$ of collision on $[0,T]$ would be to simulate many times aircraft *i* amidst other aircraft over a period of length *T* and count all cases in which the realization of the moment τ^i is smaller than *T*. An estimator for the collision risk of aircraft *i* per unit *T* of time then is the fraction of simulations for which $\tau^i < T$.

B. Risk factorization using multiple conflict levels

Prior to a collision of aircraft *i* with aircraft *j* a sequence of conflicts ranging from long term to short term always occurs. In order to incorporate this explicitly in the MC simulation, we formalize this sequence of conflict levels through the sequence of closed subsets of $\mathbb{R}^n$, $D^{ij} = D_m^{ij} \subset D_{m-1}^{ij} \subset \ldots \subset D_1^{ij}$, with for $k = 1,\ldots, m$:

$$\begin{aligned}D_k^{ij} = \{x \in \mathbb{R}^n;\ &\left|y^{ij}(x) + \Delta v^{ij}(x)\right| \le d_k \ \text{AND}\\ &\left|z^{ij}(x) + \Delta r^{ij}(x)\right| \le h_k,\ \text{for some } \Delta \in [0, \mathrm{T}_k]\},\ i \neq j\end{aligned} \tag{7}$$

where d_k, h_k and T_k are the parameters of the conflict definition at level *k*, and with $d_m = l$, $h_m = s$ and $T_m = 0$, and with $d_{k+1} \ge d_k$, $h_{k+1} \ge h_k$ and $T_{k+1} \ge T_k$. If x_t hits D_k^{ij} at time τ_k^{ij}, then we say the first level *k* conflict event between aircraft (*i*,*j*) occurs at moment τ_k^{ij}, i.e.

$$\tau_k^{ij} = \inf\{t > 0;\ x_t \in D_k^{ij}\} \tag{8}$$

Similarly as we did for the collision level, for aircraft *i* we consider the first moment τ_k^i that aircraft *i* reaches conflict level *k* with any of the other aircraft, i.e.

$$\begin{aligned}\tau_k^i &= \inf_{j\neq i}\{\tau_k^{ij}\} = \inf_{j\neq i}\{t>0;\ x_t \in D_k^{ij}\}\\ &= \inf\{t>0;\ x_t \in D_k^i\}\end{aligned} \tag{9}$$

with $D_k^i \triangleq \bigcup_{j\neq i} D_k^{ij}$.

Following [35], we can write the probability of collision of aircraft *i* with any of the other aircraft as a product of conditional probabilities of reaching the next conflict level given the current conflict level has been reached:

$$\mathbb{P}(\tau_m^i < T) = \prod_{k=1}^{m} \gamma_k^i \tag{10}$$

where $\gamma_k^i \triangleq \mathbb{P}(\tau_k^i < T \,|\, \tau_{k-1}^i < T)$

With this, the problem can be seen as one to estimate the conditional probabilities γ_k^i in such a way that the product of these estimators is unbiased. Let $\bar{\gamma}_k^i$ denote the value estimated for γ_k^i. Because of the multiplication of the various individual $\bar{\gamma}_k^i$ estimated values, which depend on each other, in general such a product may be heavily biased. The key novelty of [32] was to show that such a product may be evaluated in an unbiased way when $\{x_t\}$ makes part of a larger stochastic process that satisfies the strong Markov property. Hence the resulting sequential MC simulation essentially consists of taking advantage of the nested sequence of closed subsets of $\mathbb{R}^n$: $D = D_m^i \subset D_{m-1}^i \subset \ldots \subset D_1^i$, and then start simulation from outside D_1^i to D_1^i (this yields $\bar{\gamma}_1^i$), and subsequently simulate from D_1^i to D_2^i (this yields $\bar{\gamma}_2^i$), from D_2^i to D_3^i (this yields $\bar{\gamma}_3^i$), …, and finally from D_{m-1}^i to D_m^i (this yields $\bar{\gamma}_m^i$). The estimated risk for aircraft *i* to collide with any of the other aircraft then is $\prod_{k=1}^m \bar{\gamma}_k^i$.

IV. Simulated scenarios and collision risk estimates

The sequential MC simulation approach outlined in section III is now applied to three hypothetical AMFF scenarios. The first scenario has eight aircraft that fly at the same flight level. Based on their flight plans, the eight aircraft are expect to fly through the same point in space at the same moment in time. The second scenario has one aircraft flying through an area of seven randomly distributed aircraft per container of 40 Nm x 40 Nm x 3000 feet. The third scenario is the same as the second, except with a container that is twice as large in width/length. Prior to describing these scenarios and simulation results, we explain the parametrization of the IPS algorithm used.

A. Parameterization of the IPS simulations

The main safety critical parameter settings of the AMFF enabling technical systems (GPS, ADS-B and ASAS) are given in the Table I.

TABLE I. PARAMETER VALUES OF AMFF ENABLING TECHNICAL SYSTEMS

Model Parameter	**Probability**
Global GPS down	1.0×10^{-5}
Global ADS-B down[1]	1.0×10^{-6}
Aircraft ADS-B Receiver down	5.0×10^{-5}
Aircraft ADS-B Transmitter down	5.0×10^{-5}
Aircraft ASAS System mode corrupted	5.0×10^{-5}
Aircraft ASAS System mode failure	5.0×10^{-5}

The IPS conflict levels k are defined by parameter values for lateral conflict distance d_k, conflict height h_k and time to conflict T_k. These values have been determined through two steps. The first was to let an operational expert make a best guess of proper parameter values. Next, during initial simulations with the IPS some fine tuning of the number of levels and of parameter values per level has been done. The resulting values are given in Table II.

TABLE II. IPS CONFLICT LEVEL PARAMETER VALUES

k	1	2	3	4	5	6	7	8
d_k (Nm)	4.5	4.5	4.5	4.5	2.5	1.25	0.5	0.054
h_k (ft)	900	900	900	900	900	500	250	131
T_k (min)	8	2.5	1.5	0	0	0	0	0

B. Eight aircraft on collision course

In this simulation eight aircraft start at the same flight level, some 250 km out of each other, and fly in eight 45 degrees differing directions with a ground speed of 240 m/s, all up to the same point in the middle. By running ten times the IPS algorithm the collision risk is estimated ten times. The number of particles per IPS simulation run is 12,000. The total simulation time took about 20 hours on two machines, and the computer memory load was about 2.0 GigaByte per machine.

TABLE III. FRACTIONS COUNTED DURING FOUR IPS RUNS OF SCENARIO 1

Level	1st IPS	2nd IPS	3rd IPS	4th IPS
1	1.000	1.000	1.000	1.000
2	0.528	0.529	0.539	0.533
3	0.426	0.429	0.424	0.431
4	0.033	0.036	0.035	0.037
5	0.175	0.180	0.183	0.181
6	0.267	0.158	0.177	0.144
7	0.150	0.268	0.281	0.427
8	0.000	0.009	0.233	0.043
Product of fractions	0.0	5.58×10^{-7}	1.67×10^{-5}	4.01×10^{-6}

[1] Global ADS-B down refers to frequency congestion/overload of the data transfer technology used for ADS-B.

For the first four IPS runs, the estimated fractions $\bar{\gamma}_k^i$ are given in Table III for each of the conflict levels, k = 1,..,8, and aircraft i = 1. It can be seen that the first IPS run has zero particles that reach the last (8th level). Hence the first IPS run does not yield a useful estimate, and is not used for estimating the risk.

The IPS estimated mean probability for one aircraft to collide with any of the other seven aircraft equals 2.2×10-5. The minimum and maximum values now are respectively a factor 250 lower and a factor 4 higher than the mean value. We also verified that this risk value was not sensitive at all to the failure rates of the ASAS related technical systems.

In [4] a similar eight aircraft encounter scenario had been simulated many times, without experiencing any collision event. However, at a collision probability value of 2.2×10^{-5}, one needs to run about 6,000 runs to have a 50% chance of counting at least one collision, and this high number of independent simulations with the eight aircraft encounter have not been performed. As such the current results agree quite well with the fact that in these earlier simulations for the eight aircraft scenario no collision has been observed. We also verified that the novel simulation results for the eight aircraft scenario agreed quite well with the expectation of the designers of the AMFF operational concept.

C. Free flight through an artificially constructed airspace

In this simulation the complete airspace is divided into packed containers. Within each container a fixed number of seven aircraft (i=2,.,8) fly at arbitrary position and in arbitrary direction at a ground speed of 240 m/s. One additional aircraft (i=1) aims to fly straight through a sequence of connected containers, at the same speed, and the aim is to estimate its probability of collision with any of the other aircraft per unit time of flying.

Per container, the aircraft within it behave the same. This means that we have to simulate each aircraft in one container only, as long as we apply the ASAS conflict prediction and resolution also to aircraft copies in the neighboring containers. In principle this can mean that an aircraft experiences a conflict with its own copy in a neighboring container. This also means that the size of a container should not go below a certain minimum size.

By changing container size we can vary traffic density. To choose the appropriate traffic density, our reference point is the highest number (17) of aircraft counted at 23rd July 1999 in an en-route area near Frankfurt of size 1 degree x 1 degree x FL290-FL420. This comes down to 0.0032 a/c per Nm3. For our simulation we assume a 3 times higher traffic density, i.e. 0.01 a/c per Nm3. This resulted in choosing containers having a length of 40 Nm, a width of 40 Nm and a height of 3000 feet, and with 8 aircraft flying in such container.

By running the IPS algorithm ten times (+ one extra IPS run later on) over 20 minutes, with 5 minutes convergence time prior to this, the collision probability per unit time of flying has been estimated. The number of particles per IPS simulation run is 10,000. The total simulation time took about 300 hours on two machines, and the load of computer memory per machine was about 2.0 GigaByte. For the first four IPS runs, the estimated fractions $\overline{\gamma}_k^i$ are given in Table IV for each of the conflict levels, k = 1,..,8, for aircraft i = 1.

TABLE IV. FRACTIONS COUNTED DURING FOUR IPS RUNS OF SCENARIO 2

Level	1st IPS	2nd IPS	3rd IPS	4th IPS
1	0.922	0.917	0.929	0.926
2	0.567	0.551	0.560	0.559
3	0.665	0.666	0.674	0.676
4	0.319	0.331	0.323	0.321
5	0.370	0.367	0.371	0.379
6	0.181	0.158	0.162	0.171
7	0.130	0.209	0.174	0.145
8	0.067	0.005	0.094	0.066
Product of fractions	$6.42\text{x}10^{-5}$	$6.76\text{x}10^{-6}$	$1.11\text{x}10^{-4}$	$6.99\text{x}10^{-5}$

The estimated mean probability of collisions per 20 minutes aircraft flight equals 5.22×10-5, which is equal to a probability of collisions per aircraft flight hour of 1.6×10-4, with minimum and maximum values respectively a factor four lower and higher. We also verified that this risk value was not sensitive at all to the failure rates of the ASAS related technical systems.

One should be aware that this value has been estimated for the simulation model of the intended AMFF operation. Hence the question is what this means for the intended AMFF operation itself? By definition a simulation model of AMFF differs from the intended AMFF operation. If it can be shown that the combined effect of these differences on the risk level is small, then the results obtained for the simulation model may be considered as a good representation of the accident risk of the intended operation. In order to assess the combined effect of these differences there is need to perform a bias and uncertainty assessment [30], [31].

In order to better learn understanding what causes the collision risk of the simulation model to be relatively high, we performed an extra IPS run, and memorized in static memory for each particle the ancestor history at each of the eight levels. This allowed us to trace back what happened for the particles that hit the last level set (i.e. collision). There appeared to be five different collision events. Evaluation of these five collision events showed that all five happened under nominal safety critical conditions. Four of the five collisions were due to a growing number of multiple conflicts that could not be solved in time under the operational concept adopted. The fifth collision was of another type: at quite a late moment finally a conflict between two aircraft was solved with a maneuver by one of the two aircraft. However because of this maneuver there was a sudden collision with a third nearby aircraft.

These detailed evaluations of the five collision events of the 11th IPS run also showed that a significant increase of collision risk is caused by the relatively small height (4000 ft) of a container. Because of this small height it happened that an aircraft in one container came in conflict with a copy of its own in a neighbouring container, and in such a situation there was an undesired limitation in conflict resolution options, and thus an undesired artificial increase in collision risk.

The results in this section seem to indicate that the key factor in the increased risk of collision for encounters with homogeneous traffic in the background - as opposed to the eight encountering aircraft scenario - are the multiple conflicts. Under the far higher traffic densities than what the AMFF operational concept was designed for, it is not always possible to timely solve a sufficiently high fraction of those multiple conflicts. On the basis of this finding one would expect that the collision risk would decrease faster than linear with a decrease in traffic density. The validity of this expectation is verified by the next scenario.

D. *Reduction of aircraft density by a factor four*

Now we enlarge the length and width of each container by a factor two. This means that the traffic density is gone down by a factor four. Hence the density is now ¾ of the density counted on 23rd July 1999 in the en-route area near Frankfurt. This still is a factor 2.5 higher than current average density above Europe. At the same time simulated flying time has been increased to 60 minutes (with 10 minutes prior flying to guarantee convergence).

By running four times the IPS algorithm the collision risk is estimated four times. The number of particles per IPS simulation run is 10,000. The total simulation time took about 280 hours on two machines, and the load of computer memory per machine was about 2.0 GigaByte. For these IPS runs, the estimated fractions $\overline{\gamma}_k^i$ are given in Table V for each of the conflict levels, k = 1,..,8, for aircraft i = 1.

TABLE V. FRACTIONS COUNTED DURING FOUR IPS RUNS OF SCENARIO 3

Level	1st IPS	2nd IPS	3rd IPS	4th IPS
1	0.755	0.750	0.752	0.749
2	0.295	0.292	0.286	0.285
3	0.476	0.475	0.497	0.487
4	0.263	0.258	0.266	0.267
5	0.321	0.315	0.300	0.328
6	0.068	0.088	0.082	0.096
7	0.156	0.367	0.290	0.254
8	0.011	0.059	0.021	0.005
Product of fractions	$1.07\text{x}10^{-6}$	$1.61\text{x}10^{-5}$	$4.31\text{x}10^{-6}$	$1.07\text{x}10^{-6}$

The estimated mean probability of collision per aircraft flight hour equals 5.64×10-6, with minimum and maximum values respectively a factor five lower and higher. This is about a factor 30 lower than the previous scenario with a four times higher aircraft density. Thus, for the model there is a steep decrease of collision probability with decrease of traffic density, and this agrees well with the expectation at the end of the previous section.

E. Discussion of IPS simulation results

Because of the IPS simulation approach we were able to estimate collision risk for complex multiple aircraft scenarios. This is a major improvement over what was accomplished by [8] for a scenario of two free flight equipped aircraft that were supposed to fly within a fixed route structure. Inherent to the IPS way of simulation, the dynamic memory of the computers used appeared to pose the main limitation on the full exploitation of the novel sequential MC simulation approach. This also prevented performing a bias and uncertainty assessment for the differences between the simulation model and the AMFF operation. As long as such a bias and uncertainty assessment has not been performed, any conclusion drawn from the simulation apply to the simulation model only, and need not apply to the intended AMFF operation.

The simulations performed for a model of AMFF allow free flight operational concept developers to learn characteristics of the simulation model. Because of the sequential MC simulation based speed up, these simulations can show events that have not been observed before in MC simulations of an AMFF model. Under far higher traffic densities than what the AMFF operational concept has been designed for, the simulations of the model shows it is not always possible to timely solve multiple conflicts. As a result of this, at high traffic levels there is a significant chance that multiple conflicts are clogging together, and this eventually may cause a non-negligible chance of collision between aircraft in the simulation model. It has also been shown that by lowering traffic density, the chance of collision for the model rapidly goes down.

V. Concluding Remarks

This paper studied collision risk estimation of a free flight operation through a sequential Monte Carlo simulation. First a Monte Carlo simulation model of this free flight operational concept has been specified in a compositional way using the Stochastically and Dynamically Coloured Petri Net (SDCPN). Subsequently a novel sequential MC simulation method [32]-[34] has been extended for application to collision risk estimation in air traffic, and has subsequently been applied to an SDCPN model of free flight.

The results obtained clearly show that the novel simulation model specification and collision risk estimation methods allow to speed up Monte Carlo simulation for a much more complex simulation model than what was possible before (e.g. [8], [36]). Moreover, for the simulation model of the free flight operational concept considered, behavior has be made visible that was expected by free flight concept designers, but could not be observed in earlier Monte Carlo simulations: the rare chance of clogging multiple conflicts at far higher traffic density levels than where the particular concept has been designed for. Hence, further attention has to be drawn towards the development and incorporation in the particular operational concept design of advanced methods in handling multiple conflicts. [4] studied a conflict resolution approach that performs better than the one adopted in the AMFF concept. In addition, there are some complementary developments that aim to develop complex conflict resolution solvers with guaranteed level of performance [37], [38], including ways to incorporate situation awareness views by human operators (pilots and/or controllers) in these combinatorial conflict resolution problems [39].

The main value of having performed this collision risk estimation for an initial simulation model of AMFF is that this provides valuable feedback to the design team and allows them to learn from Monte Carlo simulation results they have never seen before. This allows them to significantly improve their understanding when and why multiple conflicts are not solved in time anymore in the simulation model. Subsequently the operational concept designers can use their better understanding for adapting the AMFF design such that it can better bring into account future high traffic levels.

In its current form the sequential MC simulation approach works well, but at the same time poses very high requirements on the availability of dynamic computer memory and simulation time. The good message is that in literature on sequential MC simulation (e.g. [40]-[43]), complementary directions have been developed which remain to be explored for application to free flight collision risk estimation. These potential improvements of sequential MC simulation approach, and their application to free flight collision risk and bias and uncertainty estimation, will be studied in follow-up research.

References

[1] RTCA. Free Flight Implementation, Task Force 3 Final Technical Report, Washington DC, 1995.
[2] NASA. Concept definition for distributed air-/ground traffic management (DAG-TM), Version 1.0, Advanced Air Transportation Technologies project, Aviation System Capacity Program, National Aeronautics and Space Administration, NASA, 1999.
[3] J. Krozel. Free flight research issues and literature search. Under NASA contract NAS2-98005, 2000.
[4] J. Hoekstra. Designing for Safety, the Free Flight Air Traffic Management concept, PhD Thesis, Delft University of Technology, November 2001.
[5] FAA/Eurocontrol. Principles of Operations for the Use of ASAS, Cooperative R&D Action Plan 1 report, Version 7.1, 2001.
[6] ICAO. Airborne separation assistance system (ASAS) circular, Draft, version 3, SCRSP, WGW/1 WP/5.0, International Civil Aviation Organization, May 2003.
[7] FAA/Eurocontrol. Safety and ASAS applications, Co-operative R&D Action Plan 1 report, version 4.1, 2004.
[8] M.H.C. Everdij, H.A.P. Blom, G.J. (Bert) Bakker. Modeling lateral spacing and separation for airborne separation assurance using Petri nets. Journal of Simulation, 2006.
[9] B. Gayraud, F. Nacchia, J. Barff, R.C.J. Ruigrok, MFF operational concept, requirements and procedures, Report MFF D220, October 2005, www.medff.it/public/index.asp
[10] J.W. Andrews, J.D. Welch, H. Erzberger. Safety analysis for advanced separation concepts. In *Proceedings of USA/Europe ATM R&D Seminar*, Baltimore, USA, 27–30 June 2005.
[11] H. Erzberger. Transforming the NAS: The next generation air traffic control system. In *Proceedings of the 24th Int. Congress of the Aeronautical Sciences (ICAS)*, 2004.

[12] MFF. MFF Final safety case, Report MFF D734, ed. 1.0. Available at http://www.medff.it/public/index.asp, November 2005.

[13] R. Gordon, S.T. Shorrock, S. Pozzi, A. Boschiero, Predicting and simulating human errors in using the airborne separation assurance system procedure, Human Factors and Aerospace Safety, 2005.

[14] R.C.J. Ruigrok, N. de Gelder, H.J. Abma, B. Klein Obbink, J.J. Scholte, Pilot perspective of ASAS self separation in challenging environments, Proc. 6th USA/Europe ATM R&D Seminar, Baltimore, USA, 27-30 June 2005.

[15] D. Schaefer, C. Fusai, P. Scrivani, R. Waggin, Pilot evaluation of ASAS spacing applications under non-nominal conditions, Proc. 6th USA/Europe ATM R&D Seminar, Baltimore, USA, 27-30 June 2005.

[16] A.P. Shah, A.R. Pritchett, K.M. Feigh, S.A. Kalarev, A. Jadhav, K.M. Corker, D.M. Holl, R.C. Bea. Analyzing air traffic management systems using agentbased modeling and simulation. In *Proceedings of the 6th USA/Europe Seminar on Air Traffic Management Research and Development*, Baltimore, USA, 27-30 June 2005.

[17] K. Corker. Cognitive Models & Control: Human & System Dynamics in Advanced Airspace Operations, Eds: N. Sarter and R. Amalberti, *Cognitive Engineering in the Aviation Domain*, Lawrence Earlbaum Associates, New Jersey, 2000.

[18] H.A.P. Blom, J. Daams and H.B. Nijhuis. Human cognition modeling in Air Traffic Management safety assessment. In *Air Transportation Systems Engineering*, edited by G.L. Donohue and A.G. Zellweger, Vol. 193 in Progress in Astronautics and Aeronautics, Paul Zarchan, Editor-in-Chief, Chapter 29, pages 481–511, 2001.

[19] H.A.P. Blom, S.H. Stroeve, M.H.C. Everdij, M.N.J. Van der Park. Human cognition performance model to evaluate safe spacing in air traffic. *Human Factors and Aerospace Safety*, Vol. 2, pages 59–82, 2003.

[20] H.A.P. Blom, K.M. Corker, S.H. Stroeve. On the integration of human performance and collision risk simulation models of runway operation. In *Proceedings of the 6th USA/Europe Air Traffic Management R&D Seminar*, Baltimore, USA, 27–30 June 2005.

[21] S.H. Stroeve, H.A.P. Blom and M.N. van der Park. Multi-Agent Situation Awareness Error Evolution in Accident Risk Modelling. In *Proceedings of the 5th USA/Europe Seminar on Air Traffic Management Research and Development*, Budapest, Hungary, 23-27 June 2003.

[22] F. Maracich, Flying free flight: pilot perspective and system integration requirements, Proc. 24th DASC, Washington, 2005.

[23] P.E. Labeau, C. Smidts and S. Swaminathan. Dynamic reliability: towards an integrated platform for probabilistic risk assessment. *Reliability Engineering and System Safety*, Vol. 68, pages 219–254, 2000.

[24] R. David, H. Alla. Petri Nets for the modeling of dynamic systems – A survey, *Automatica*, Vol. 30, No. 2, pages 175–202, 1994.

[25] M.H.C. Everdij, H.A.P. Blom. Petri Nets and Hybrid state Markov Processes in a power-hierarchy of dependability models. In Proceedings of IFAC Conference on Analysis and Design of Hybrid Systems, Saint-Malo Brittany, France, pages 355–360, June 2003.

[26] M.H.C. Everdij, H.A.P. Blom. Piecewise deterministic Markov processes represented by dynamically coloured Petri nets, Stochastics, Vol. 77, pages 1–29, 2005.

[27] M.H.C. Everdij, H.A.P. Blom. Hybrid Petri nets with diffusion that have into mappings with generalised stochastic hybrid processes. In [40], pages 31–64, 2006.

[28] G. Pola, M.L. Bujorianu, J. Lygeros, M.D. Di Benedetto. Stochastic hybrid models: an overview with applications to air traffic management. In *Proceedings of IFAC Conf. Analysis and Design of Hybrid Systems (ADHS)*, 2003.

[29] M.H.C. Everdij, M.B. Klompstra, H.A.P. Blom, B. Klein Obbink. Compositional specification of a multi-agent system by stochastically and dynamically coloured Petri nets. In [40], pages 325–350, 2006.

[30] M.H.C. Everdij and H.A.P. Blom. Bias and uncertainty in accident risk assessment, NLR report CR-2002-137, National Aerospace Laboratory NLR, 2002.

[31] M.H.C. Everdij, H.A.P. Blom, S.H. Stroeve. Structured assessment of bias and uncertainty in Monte Carlo simulated accident risk. In Proceedings of the 8th Int. Conf. on Probabilistic Safety Assessment and Management (PSAM8), New Orleans, USA, May 2006.

[32] F. Cérou, P. Del Moral, F. Le Gland and P. Lezaud. Genetic genealogical models in rare event analysis, Publications du Laboratoire de Statistiques et Probabilites, Toulouse III, 2002.

[33] F. Cérou, P. Del Moral, F. Le Gland, P. Lezaud. Limit theorems for the multilevel splitting algorithms in the simulation of rare events. In Proceedings of Winter Simulation Conference, Orlando, USA, 2005.

[34] J. Krystul, H.A.P. Blom. Generalised stochastic hybrid processes as strong solutions of stochastic differential equations, Hybridge Report D2.3, 2005, see http://www.nlr.nl/public/hosted-sites/hybridge/

[35] J. Krystul, H.A.P. Blom. Sequential Monte Carlo simulation of rare event probability in stochastic hybrid systems. In *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, June 4-8, 2005.

[36] H.A.P. Blom, G.J. Bakker, M.H.C. Everdij and M.N.J. van der Park. Collision risk modeling of air traffic. In *Proceedings of European Control Conference*, Cambridge, UK, 2003.

[37] Lecchini, W. Glover, J. Lygeros, J. Maciejowski. Monte Carlo optimisation for conflcit resolution in air traffic control. In [40], pages 257–276, 2006.

[38] D.V. Dimarogonas, S.G. Loizou, K.J. Kyriapoulos. Multirobot navigation functions II: towards decentralization. In [40], pages 209–256, 2006.

[39] E. De Santis, M.D. Di Benedetto, S. Di Gennaro, A.D.'Innocenzo, G. Pola. Critical observability of a class of hybrid systems and application to air traffic management. In [40], pages 141–170, 2006.

[40] H.A.P. Blom, J. Lygeros. *Stochastic Hybrid Systems: Theory and Safety Critical Applications*, LNCIS series, Springer, Berlin, 2006.

[41] A. Doucet, N. de Freitas and N. Gordon. *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, 2001

[42] P. Glasserman. Monte Carlo Methods in Financial Engineering, Stochastic Modelling and Applied Probability, Vol. 53, Springer, 2003.

[43] P. Del Moral. *Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications*, SpringerVerlag, New York, 2004.

Henk Blom was born in The Netherlands in 1953, holds M.Sc. degree from Twente University (1978) and Ph.D. degree from Delft University of Technology (1990). In 1979, he performed research in forward-looking infra-red picture processing at TNO Physics Laboratory in The Hague. Since 1980 he is at NLR in Amsterdam to perform research on stochastic hybrid systems with potential applications to Air Traffic Management (ATM). In 1988, he was visiting researcher at the University of Connecticut, Storrs, USA. Back at NLR he gives scientific lead in the area of developing stochastic analysis based innovative developments in ATM. The most notable achievements of his approach are ARTAS and TOPAZ. ARTAS is EUROCONTROL's Bayesian multi-sensor multi-target tracking system, which incorporates automated data registration and contingency management. TOPAZ is a stochastic analysis based accident risk analysis methodology and modelling system to provide safety/capacity feedback to advanced ATM designs. Since 2002, TOPAZ has proven its great value in providing safety feedback to the design and introduction of novel air traffic operations at Amsterdam airport. From 2002 till 2005 he was coordinator of the European Commission project HYBRIDGE, which provided top level researchers in stochastic analysis an opportunity to study the modelling and analysis of future ATM. Dr. Blom is (co-)author of over 50 conference papers, 10 journal papers and 10 book chapters, and co-editor of the title: "Stochastic Hybrid Systems: Theory and Safety Critical Applications" published by Springer in 2006. He is recipient of the Dr.Ir. B.M. Spee Award in 2004.

Safety study of Free Flight by Analysis and Experiments

B. Klein Obbink and J.J. Scholte

Abstract—**Free Flight is an aviation concept providing aircrew with the freedom to select their trajectory without conditions imposed by route structure or Air Traffic Control. This freedom is gained at the cost of the crew taking full responsibility for airborne separation management, being enabled by advanced navigation, surveillance and digital communication means. The concept potentially brings several operational advantages but the question whether it is sufficiently safe is unanswered. This paper reports on a broad and in-depth study of the safety of the application within the Mediterranean Free Flight (MFF) Programme. The investigations exist of cockpit experiments and pilot assessment input on one hand and of structured safety analysis by means of functional analysis and elementary probabilistic mathematics on the other hand. After applying the ED78a methodology to derive Safety Objectives from a given Target Level of Safety, it is tried to obtain reasonable safety requirements using Fault Tree Analysis. Several detailed safety requirements were indeed deduced, but some Safety Objectives were not satisfied and further research is required in the field of a) the crew's ability to execute new complex tasks and b) the frequency and nature of multiple conflicts. Despite these unresolved issues, it can be stated that free flight is a promising concept as after several years of critical testing, there turned out not to be any safety issue preventing implementation in commercial aviation.**

Index Terms—**ASAS, Free Flight, Safety**

I. Introduction and Background

In the early days of flying, visual contact with the ground and other aircraft was required in order to navigate and avoid collisions. The introduction of radar allowed flying without outside visual reference and this induced Air Traffic Control to become responsible for separation, using surveillance and communication by radio/telephony via standardised phraseology as most important means. The current Air Traffic Management (ATM) is historically developed and only suboptimal with respect to flight time, fuel consumption, airspace capacity, cost of ATM services and collision occurrence [4]. Recent technological developments make it however possible to reconsider the basics of Air Traffic Control. This has led to a definition of Free Flight as: "… *a safe and efficient flight operating capability under instrument flight rules in which the operators have the freedom to select their path and speed in real time…*". The goal of free flight is to provide more flexibility for aircraft operators at the cost of the aircrew being tasked to avoid conflicts. The ATM system thus changes from a centrally controlled system to a distributed system and may thus overcome some of its current problems.

In order to take the responsibility for separation management, aircrews will be supported by airborne systems as the Airborne Separation Assistance System (ASAS) enabling airborne conflict detection and resolution. ASAS can support several applications as Air Traffic Situation Awareness, ASAS Spacing and ASAS Separation [1]. The Mediterranean Free Flight (MFF) Programme under sponsorship of the European Commission researched several of these applications in broad terms in the Mediterranean Area, addressing basic research issues and exploring the concept rather than prepare its operational introduction [2]. One of the main objectives of MFF was the verification of appropriate operational procedures for flight crews in free flight scenarios, through simulations and flight trials. An important aspect of this verification was the creation of a safety case ([3]). Before findings contained in this safety case are reported on in the sections "Safety Investigations", "Results" and "Conclusions", the free flight concept is first further explained in some detail.

II. Enabling Techniques and Avionics

The navigation system in a free flight enabled aircraft is assumed to include a Global Positioning System (GPS), an Inertial Reference Systems (IRS) and Distance Measuring Equipment (DME). The positioning information is assumed to be broadcast each second in a

Manuscript received February 9, 2006.
B. Klein Obbink and J.J. Scholte are both with the Safety and Flight Operations Department of NLR, National Aerospace Laboratory, 1059 CM Amsterdam, The Netherlands
(e-mail: bklein@nlr.nl and scholte@nlr.nl).

standardised format by means of Automatic Dependent Surveillance Broadcast (ADS-B). ASAS is then fed by navigation data from the own aircraft and by the ADS-B derived information from aircraft in the vicinity. The currently existing ACAS system is foreseen to remain on-board as a safety-net system, and is considered as separated from ASAS.

Being provided by the horizontal and vertical positions and velocities of all aircraft in the vicinity of the ownship, ASAS executes calculations concerning conflicts and potential conflicts. A conflict is a violation of separation standards, which are typically 5 nm horizontal and 1000 ft vertical, corresponding to a cylinder-shaped volume around each aircraft that is meant to be free of other aircraft. A potential conflict is a situation in which a conflict would appear within a certain look-ahead time of typically 6 minutes, if all aircraft maintain the momentaneous velocity. ASAS provides the following functionality:

- Conflict detection: potential conflicts are detected by checking whether the distance between ownship and intruder at the closest point of approach within a look-ahead time is below the separation standards.
- Conflict resolution: in case of a potential conflict, an appropriate heading change and an appropriate track change are calculated that will both prevent that conflict. Appropriate here implies a minimal change corresponding to the “shortest way out of the conflict” such that the conflict is avoided, independent of whether one or both aircraft execute one or both manoeuvres.
- Conflict prevention or Preventive ASAS (PASAS): it is calculated which immediate track changes and which immediate vertical speed changes will result in a potential conflict.

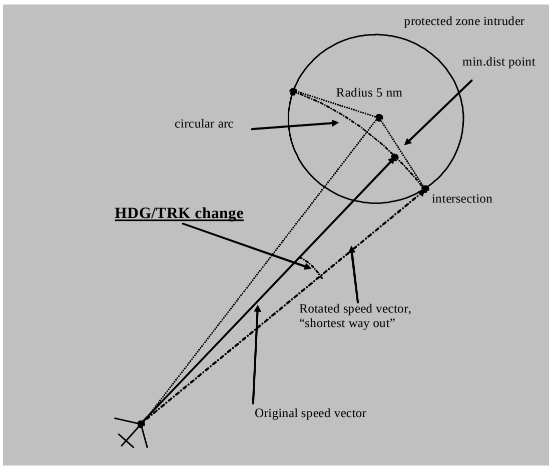


Fig. 1. Conflict detection and resolution. The figure disregards the vertical dimension and shows the ownship and the position of the protected zone of the intruder at the closest point of approach. The rotated speed vector to the edge of the intruder’s protected zone solves the conflict. This will require a heading change as indicated.

The calculations for conflict prevention, detection and resolution correspond to linear extrapolations of the current state vectors for position and velocity, leading to a set of polynomial equations of second degree expressed in no other physical quantities than distances, velocities and time periods. The approach is therefore referred to as state-based, in opposition to intent-based.

The information derived from ADS-B and ASAS is displayed on the Lateral and Vertical Navigation Displays within a format to be selected by the crew members. This traffic information includes:

- Surveillance information of aircraft in the vicinity as call sign, relative horizontal position, relative or absolute altitude, groundspeed or calibrated airspeed, track and vertical speed arrow.
- Conflict detection and resolution information including an indication of the aircraft with which the ownship is in conflict with, the time to conflict and the ASAS derived conflict resolution manoeuvres.
- The results of PASAS: no-go bands on the vertical speed tape of the Primary Flight Display, and on the heading scale of the Navigation Display. A colour distinction is made between conflicts within 0 to 3 minutes away from intrusion and conflicts between 3 and 6 minutes away from intrusion.

In addition to this visual information, the crew is alerted by means of aural signals indicating the appearance and development of potential conflicts.

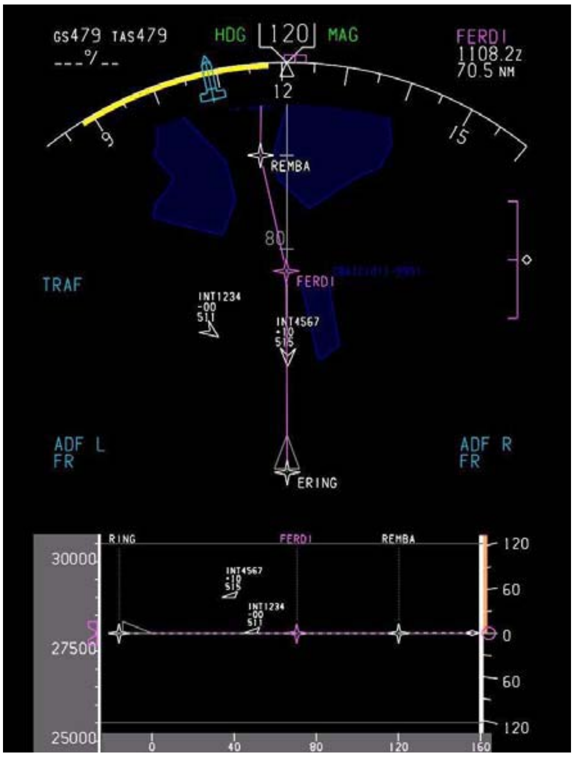


Fig. 2. Surveillance information and PASAS symbology.

III. PROCEDURES AND AIRCREW RESPONSIBILITIES

The displayed traffic information supports the aircrew in its airborne separation management. Several proposals have been made about the design of procedures concerning priority principles, rules for conflict solving sequence determination, exchange of intent information and explicit co-ordination to confirm conflicts. The concept under investigation is quite lean and mean, avoids intent information exchange and complex rules that may reduce the pilot's comprehensibility [4]. The basic procedures can be summarised as follows:

- In case the time to a potential conflict is less than three minutes away −indicated by amber colouring on the display− the crew is obliged to execute a conflict resolution manoeuvre.
- In case the time to a potential conflict is more than three minutes away, there is only one crew obliged to execute a conflict resolution manoeuvre, for which the conflict is displayed in yellow. For the other crew the conflict is displayed in green, meaning that they have priority and that there is no necessity to execute a manoeuvre. The ASAS systems of both involved aircraft automatically determine and indicate priority independently of each other by applying a hierarchic set of specific rules. If for example two aircraft are flying level and only one does obey the Flight Level Orientation Scheme, than this aircraft gets priority.
- There is freedom to select one of the ASAS advised conflict resolution manoeuvres, to combine the manoeuvres or to exaggerate them. It is not allowed to execute a manoeuvre opposite the ASAS indicated resolutions.
- In order to prevent potential conflicts, pilots are not allowed to turn, climb or descend into an amber PASAS no-go band, indicating a potential conflict within 0 to 3 minutes. Moreover, pilots are not allowed to stay in a yellow P-ASAS band, indicating a potential conflict within 3 to 6 minutes conflicts, although crossing such band is allowed. The Flight Management System (FMS) and ASAS are not coupled, and this implies that the aircrew should anticipate trajectory changes in the flight plan when flying on FMS.
- In case of emergency in Free Flight Airspace (FFAS), there are some additional procedures. Depending on the nature and the severity of the emergency, the Pilot-Not-Flying shall for example transmit a distress/urgency message on the FFAS frequency and turn on a priority switch that makes ADS-B broadcast a message such that the aircraft obtains priority in all eventual conflicts.
- Collision avoidance has a higher priority than conflict management. This means that all ACAS Traffic Alerts and Resolution Advisories always overrule ASAS and Free Flight procedures.

It can be expected that a trained and experienced crew assumes the following routine. The surveillance information is monitored regularly. If an aural ASAS alert is given, the Pilot-Flying acquires the traffic information to find out the nature and state of the potential conflict. In case, and only then, ASAS does does not indicate priority, the Pilot-Flying chooses a conflict resolution manoeuvre, probably a level change, executes it, probably by means of the auto-pilot, and monitors the development. After the execution of the manoeuvre, the Pilot-Flying resumes the original flight plan.

IV. SAFETY STUDY

A. Methodology

In order to minimise the contribution to the risk of aircraft accidents and in order to harmonise results obtained for different applications, the MFF Programme decided to apply the relevant parts of the EUROCAE and RTCA methodology referred to as ED78a [5]. This decision is not entirely logic, as this methodology is focussed on allocation of requirements over different stakeholder parties while in the case of free flight there is only one stakeholder party, namely the aircraft operators. Moreover, the methodology is focussed on the aspects of digital communication while free flight comprises other safety critical aspects as well, supports mainly the certification process while basic pre-operational questions concerning ASAS procedures need to be addressed and does only result in a set of requirements while the actual level of safety should be determined (see also [6], [7]).

Nevertheless, the methodology has been applied quite strictly by carrying out the following steps iteratively:

- *Operational Services Environment Definition (OSED)*, providing the basis for an assessment of system failures, procedural errors, and airspace characteristic.
- *Operational Hazard Analysis (OHA)*, where hazards that could adversely effect the aircraft operations are identified and classified according to a standardised severity scheme.
- *Allocation of Safety Objectives and Requirements (ASOR)*, where a safety objective is assigned to each hazard, from which safety requirements are deduced. Safety requirements are here to be interpreted as statements on the maximal allowable frequency of occurrence of certain specific events traceable to one functional block.

B. *Scope and safety criteria*

The main safety questions for Free Flight relate to the crew's ability to prevent and resolve conflicts. In order to focus on that, it was decided to disregard aspects as military traffic and zones, the transitions between Free Flight Airspace and Managed Air Space and the evolutionary intermediate period in which some aircraft are suitably equipped and some crews are appropriately trained while others are not. In due course, it was decided to disregard complex multiple aircraft encounters.

Quantitative safety criteria can be induced and deduced straightforwardly from the numerical overall target level of safety formulated in ESARR 4 for the highest severity class [8] by making the following assumptions:

- The maximum allowable frequency of all occurrences in a severity class is a factor 100 higher for each class that it is less severe.
- Approximately 10 % of the accident risk with a direct contribution of ATM takes place en route (see [9]),
- The mean duration of the en-route flight path is one hour. This is a reasonable assumption, given the averaged duration of a flight and the size of the Mediterranean area.
- The risk that is in scope constitutes 50% of the total risk that can occur during the considered flight segment.

C. *Analysis*

The safety analysis starts with an investigation of the sequence of events in case of an aircraft-aircraft encounter under nominal conditions. This leads to a distinction of the following stages: 1) Situational Awareness and Conflict Prevention 2) Conflict Detection 3) Priority Determination 4) Priority based Conflict Resolution and 5) Co-operation based Conflict Resolution. For each stage, operational hazards were identified by considering the task or function that needs to be performed and by asking generic questions as "What happens if this functionality fails?", "What happens if this task is executed too late?" and "What happens if none of the aircrew is aware of this situation?". This leads to a set of 24 high level formulated operational hazards, such as "Surveillance partially unavailable", "Crew unaware of ASAS-indicated conflict" and "Induction of new conflict by aircraft executing conflict resolution manoeuvre".

The potential consequences of each operational hazard are investigated by safety experts, consulting pilots and considering event sequences known in literature and flight manuals. This leads to a logically structured OHA where an operational hazard propagates to an other one in a following phase or is mitigated such that a final worst credible consequence can be described. The severity and the likelihood of the operational consequences are then assessed for every operational hazard that is not propagated. The severity assignment follows ESARR 4 scheme and guidelines [8], retrieving input from operational experts (see subsection "Safety Validation Workshop" below).

The construction of the ASOR includes the analyisis of the functional architecture, a selection of operational hazards in such a way that the most significant safety issues should be covered and the assignment of safety objectives to operational hazards in accordance with the safety criteria and the severity assessment. Fault trees are constructed to represent the relation between the occurrence of the operational hazards and the elementary causes, in line with the level of detail chosen in the functional architecture description.

It is then tried to identify maximum allowable frequencies of occurrence for the elementary causes, in such a way that the safety objectives are reached. If these values are considered achievable, safety requirements are formulated. If these values are not considered achievable, the safety objective is declared non compliant.

At several points questions arise that can only be answered well by retrieving operational expert input. Experiments and workshops are almost inevitably too limited for statistical significant results but let safety analysts obtain a good understanding of the operational environment, such that they are better qualified to predict human response and to judge human errors. An important and particularly interesting set of experiments and the final safety validation workshop are described below.

D. *Real Time experiments*

In order to examine human behaviour in detail, a set of free flight simulations were organised. A generic flight simulator in a Boeing 747 configuration was used for challenging scenarios within a realistic environment including military operations and significant weather (see also [10]). Twelve hazards that were considered interesting and suited for simulation were selected and injected on purpose, with the aim to observe the participants' reaction and to obtain their feed-back by questionnaires and interviews, see Table I. The subject human operators were eight pilots in service of commercial airlines with an average experience of approximately 5000 flight hours. The effects of failures have been assessed by observing, questioning and interviewing the participants.

TABLE I
REAL-TIME EXPERIMENT SCENARIOS

Military activities	SUA activation, known at start of run. SUA activation, unexpected during run.
Weather	Turbulence Thunderstorm Strong crosswind
Hazards	Other ADS-B transmit failure Own ASAS failure Own ADS-B transmit failure Rude pilot Own aircraft RNP failure Other ADS-B receive failure Incorrect manoeuvre Other aircraft ASAS failure Ownship navigation shift Own aircraft 2 engines flame out / displays black Other aircraft explosive decompression Own ADS-B receive failure

Military activities, weather and hazards of the scenarios used in the Free Flight Real-Time experiments. Due to the nature of flight simulations, ten out of twelve hazards can be considered as direct consequences of equipment failures.

Most events were considered good to handle, except for two scenarios ("Rude pilot" and "Navigation shift") which were considered quite undesirable. Most pilots argued however that precisely those two scenarios were unrealistic.

It turned out that there are considerable individual differences in pilot behaviour with respect to monitoring, solving and avoiding conflicts, choice of routes, awareness of free flight logic, position reports by R/T and reaction to specific failures. Hazards not earlier identified were observed concerning pilots making their own interpretation of conflicts, ACAS/ASAS relation, creating priority, priority switch and other aspects. After examination of the results of a test in which the participants were asked to draw their mental picture of the surrounding traffic after switching off the displays, it turned out that the comprehension of the geometric positions of aircraft in the neighbourhood is limited. Most pilots can reproduce the position of the nearest aircraft in front of them, albeit in some cases rather inaccurately, but not of a second relevant aircraft.

E. Safety Validation Workshop

As an important iteration step, validating and reviewing the intermediate safety case, three full day sessions were organised in which in total thirteen active pilots participated, all but one with experience in flying free flight in the NLR flight simulator. Apart from a quick recapitulation session to refresh the participating pilots' knowledge of free flight, the following sessions took place:

- Task analysis. The aim of these sessions was to identify the aircrew's relevant tasks, goals and activities en route, the perceived priorities of tasks, the ASAS mechanisms that trigger conflict management tasks, the way attention is attracted by means of sounds and lights and to estimate the probabilities to miss such signals.
- Probability assessment questionnaire. A selected set of numeric values for frequencies of occurrence of causes in the fault trees were validated. The questionnaire concentrated on human errors and on human behaviour with respect to detecting and mitigating hazardous situations. To make the pilots less reluctant to give numeric values and to gain knowledge about the degree of uncertainty, use was made of a specific expert judgement technique (see e.g., [11]).
- Severity assessment. The pilots were asked to range a set of 18 operational consequences in an order from least to most severe. Four events were included corresponding to the boundaries between the severity classes [8].
- Hazard brainstorm. Following brainstorm guidelines [12], as many and as diverse as possible ways were identified in which a pilot could abuse the freedom obtained by cunningly making use of the free flight rules.
- Interviews. The subjects were horizontal positioning errors, altitude estimate errors, potential rude pilot behaviour in free flight, the pilot's prerogative to select the four-dimensional path real time and the conditions under which a crew could be unaware of a conflict or its priority, considering the differences depending on aircraft type, airline company and technological development.
- Opinion polling. At the end of each day the extent was measured to which participants could agree with subjective statements categorised into four groups: a) the design of the Free Flight concept, in particular the relation between ACAS and ASAS, b) the logic of Free Flight, and the pilots' understanding of it, c) the required and general availability of traffic situational awareness and d) the overall safety of Free Flight.

V. RESULTS

Once the operational hazards are identified, the safety objectives are assigned, the fault trees are constructed and information on the occurrence of the events in the fault trees is validated, it remains to assign values to the maximal probabilities of occurrence of elementary causes. These values should be realistic (for given elementary causes as "turbulence"), achievable (for variable elementary causes as "ADS-B transmit failure") and such that the safety objectives are fulfilled. It is noted that there is arbitrariness in the assignment of maximal probability values as there are more variables than equations. For most operational hazards reasonable requirements could be formulated, see table II below.

TABLE II
ALL SENSITIVE SAFETY REQUIRMENTS

Elementary cause description	**Maximal frequency less than once per**
Crew without priority does not recover when other aircraft executes incorrect manoeuvre	2 occurrences
Crew without priority, being aware of conflict, priority and ASAS manoeuvre proposal, does not solve conflict	100 potential conflicts
Crew without priority unnecessarily chooses incorrect conflict resolution manoeuvre	100 potential conflicts
Crew misinterprets ASAS conflict alert	100 potential conflicts
Crew with priority unnecessarily executes an incorrect manoeuvre	100 potential conflicts
No R/T communication on malfunctioning aircraft	10 occurrences
No conflict resolution manoeuvre by crew A after communication on malfunctioning aircraft B	100 occurrences
Altitude estimate corruption by pilot input without correction	100.000 flight hours
Altitude estimate corruption by sensor without detection	10.000 flight hours
Flight technical problem aircraft	1500 potential conflicts
Environment forces crew to fly into amber PASAS band	100 Free Flight hours
Potential conflict	20 minutes

The right column presents the safety requirements as maximal values for the averaged frequency of occurrence of the event in the left column. A requirement is called sensitive if an actual frequency of occurrence significantly above the safety requirement would jeopardise the compliance of at least one safety objective.

For three operational hazards the safety objectives are not complied with: no safety requirements were identified that could be considered as reasonably achievable. These three hazards are listed in Table III below but it is noted that their nature does not give a direct clue in which aspects free flight might be improved as hazards propagate into next stages and as several elementary causes contribute to one hazard. The non-compliance is moreover at least partially due to the approach taken in the safety assessment. All this is further discussed in the Section “Conclusions”.

TABLE III
NON COMPLIANT SAFETY OBJECTIVES

Hazard description	**Sev**	**Pe**	**Safety objective**
Induction of new conflict by aircraft executing Conflict Resolution manoeuvre	3	1/5000	$4.4*10^{-3}$
No Conflict Resolution by crew B in co-operative phase	2	1/100	$8.7*10^{-7}$
Incorrect Conflict Resolution by crew B in co-operative phase	1	1/7000	$1.5*10^{-6}$

Crew A represents the crew of the aircraft that caused the co-operative phase of the conflict to be entered; Crew B represents the other crew in the conflict. “Sev” is an abreviation for severity and refers to the ESARR 4 Severity Classification Scheme. The Pe stands for the conditional probability that, given the occurrence of the hazard, the consequences have the severity assigned. The safety objective is a statement of the maximal occurrence of the hazard, in the last column expressed per potential conflict.

In the course of the analysis, the safety analysts identified a number of recommendations for concept designers, which are expected to improve the operation from the point of safety:

- Procedures might be reconsidered for some specific events such as degraded altitude estimates, a failure of the surveillance sub-functionality of ASAS, ACAS/ASAS surveillance inconsistency, and in case the crew detects a conflict that is not considered solvable.
- An automatic ASAS alert might be provided when an aircraft is suddenly not any longer detected by the airborne surveillance (while previously positioned in vicinity of the ownship) or seems to follow an unrealistic trajectory.
- The priority switch and the change of separation standards in case of emergency might be automated. A dedicated hardware priority switch might be implemented instead of the switch on the CDU.
- Non-relevant PASAS information should be suppressed.
- Reduce false ASAS conflict alerts during e.g. turbulence, turns and level-offs. To suppress the latter type of nuisance alerts, it might be considered to leave the pure state-based concept partially and to include the concept of intention by sending out the level-off altitude when climbing or descending, and to take that information into the conflict detection and resolution algorithms.
- The operation might not be acceptably safe in case of rare multiple concurrent conflicts where aircraft encounter in a complicated geometry. In a safety analysis of the same concept executed according to

another methodology, fast time simulation results for multiple aircraft encounters also give ground for hesitations on this point [13]. The operation designers are therefore recommended to reconsider the conflict resolution algorithms with respect to conflicts in which multiple aircraft are involved.

VI. CONCLUSIONS

Although there turned out not to be any safety issue preventing free flight to be implemented, the pre-operational safety case is not finished as some aspects were declared out of scope and as there are safety objectives that could not be complied with by reasonable achievable safety requirements. This non-compliance should not necessarily be interpreted as an indication that the concept is unsafe as the analysis executed so far uses a fault tree model and such approach has limiting analysing capabilities for especially human behaviour, dynamic interaction and aircraft encounter geometry. Apart from that, the kind of analysis applied implies propagation from one stage to another stage, and this makes that it can not be sharply determined which functionality is in doubt or where additional mitigating means are necessary.

It is recommended to step away from the approach of allocating requirements as in the ED78a methodology because it depends on two important assumptions: 1) that hazards, elementary causes and contributing factors can be identified and analysed by means of fault tree analysis supplied with expert judgement and 2) that the combined effect of simultaneous occurrence of hazards can be neglected. The most appropriate step forwards is to extend modelling capabilities and to analyse a) the complex situations involving dynamic interactions of multiple human actors and b) multiple concurrent conflicts where aircraft encounter in a complicated geometry [13].

Although most roadmaps are conservative, the introduction of free flight might soon be expedited, once the airlines understand the economic advantages it potentially offers. The results of the safety analysis in the frame of the MFF programme tend to increase the confidence in the feasibility of free flight.

REFERENCES

[1] Principles of Operations for the use of ASAS (PO-ASAS), Action Plan 1-FAA/Eurocontrol Cooperative R&D, version 7.1, June 2001

[2] MFF D821 final report, MFF820TR01, Ed. 1.0, November 2005; see also http://www.medff.it/public/index.asp

[3] D734- MFF Final Safety Case, MFF730TR034, Working Area 7, proposed Issue 0.5, August 2005

[4] Hoekstra, J.M., 2001, Designing for Safety, the Free Flight Air Traffic Management Concept, NLR-TP-2001-113, Amsterdam, the Netherlands; see also http://www.freeflightatm.org/

[5] EUROCAE, ED78A Guidelines for approval of the provision and use of ATS supported by data communications, December 2000.

[6] Safety And ASAS applications (SAF-ASAS), Action Plan 1-FAA/Eurocontrol Cooperative R&D, versions 4.0, December 2003

[7] P. Brooker, Airborne Separation Assurance Systems: towards a work programme to prove safety, , Safety Science 42 (2004), pp 723-754

[8] EUROCONTROL Safety Regulatory Requirement, ESARR 4, Risk assessment and mitigation in ATM, Edition 1.0, Released issue, 5 April 2001.

[9] G.W.H. van Es, Review of Air Traffic Management-related Accidents Worldwide: 1980-2001, In: “Change: A Safety Challenge”, 15th annual EASS, FSF&ERA, Geneva Switzerland, March 2003

[10] Pilot perspective of ASAS self-separation in challenging environments, R.C.J. Ruigrok et al., proceedings of 6th USA/Europe ATM2005 R&D Seminar, Baltimore USA, June 2006

[11] Cooke, R.M., and L.J.H. Goossens, Procedures guide for structured expert judgement, EURATOM. EUR 18820, 2000

[12] H.H. de Jong, Guidelines for the identification of hazards. How to make unimaginable hazards imaginable?, NLR-CR-2004-094, March 2004; also published as Guidance Material B2 to Chapter 3 of EATMP SAM, Edition 2.0

[13] Sequential Monte Carlo simulation of collision risk in free flight air traffic, H.A.P. Blom at al., Hybridge WP 9 deliverable, June 2005

ABBREVIATIONS

ASAS	Airborne Separation Assurance Systems
ASOR	Allocation of Safety Objectives and Requirements
ADS-B	Automatic Dependent Surveillance -Broadcast
ATC	Air Traffic Control
ATM	Air Traffic Management
FFAS	Free Flight Airspace
MFF	Mediterranean Free Flight
nm	Nautical mile
OHA	Operational Hazard Analysis
OSED	Operational Services Environment Definition
PASAS	Predictive ASAS
RTS	Real-time Simulation

Safety has value! An approach for the assessment of the costs and benefits of safety measures

Robert Piers, Rik Lebouille, Alfred Roelen and Job Smeltink

Abstract— **In the aviation sector, cost-benefit analysis is usually not applied when safety is involved. Decisions on safety related matters are often taken without knowing precisely how safety improvements relate to the costs they bring about. This paper presents an approach to assess the cost benefit effects of safety measures. Key notion in the approach is that safety effects can be expressed in monetary terms, and thus be compared with the associated costs of safety measures. The model consists of three main building blocks: a scenario-based risk model, a cost model that distinguishes severity levels for the applicable heads of costs and the cost-benefit trade-off.**

***Index Terms*—Economics, Risk Analysis, Safety**

I. INTRODUCTION

The application of cost-benefit analysis theory is widespread in the transport sector. Decisions about investments in new infrastructure are often supported by cost-benefit analysis. In few cases, these studies include safety effects in the cost-benefit analysis. Additionally, cost-benefit studies on safety measures are increasingly being carried out. This is the case on both a national level [1, 2, 3], as well as on European level [4, 5]. Examples in the latter respect are a cost-benefit analysis on police enforcement [6] or on measures that reduce the risk of injury to vulnerable and young road users [7].

Various theoretical and practical guidelines exist to strengthen the application of economic principles in assessments of the feasibility of new infrastructure or applications in transport [8, 9, 10]. In the aviation sector, cost-benefit analysis is applied in some subsectors, but certainly not throughout the entire aviation (policy) community. Airport expansions [11] and new technological developments in air traffic management [12, 13] are occasionally supported by cost-benefit studies. However, cost-benefit analysis of safety measures is a relatively new concept in the aviation community. Decisions on safety related matters are often taken without knowing precisely how safety benefits relate to the costs they induce.

Manuscript received February 26, 2006. This work was supported in part by the European Commission as part of their 6th Framework Programme.

Robert Piers, ECORYS Nederland BV, Rotterdam, the Netherlands (e-mail: robert.piers@ecorys.com).

Rik Lebouille, ECORYS Nederland BV, Rotterdam, the Netherlands (e-mail: rik.lebouille@ecorys.com).

Alfred Roelen, National Aerospace Laboratory NLR, Amsterdam, the Netherlands (e-mail: roelen@nlr.nl).

Job Smeltink, National Aerospace Laboratory NLR, Amsterdam, the Netherlands (e-mail: smeltink@nlr.nl).

This paper presents an approach to assess the effects of safety improvement measures. These effects concern both safety and economic effects. The approach offers the opportunity to make a balanced trade-off between safety and investment costs. The main concept in the approach is that safety effects can be expressed in monetary terms, and thus can be compared with the associated costs of safety measures.

The approach has been developed within the ASICBA project [14]. This project is a 6th framework project which builds upon experience gained in previous projects [15, 16]. Foreseen users of the approach, e.g. airlines, airport operators and (safety) authorities, are participating in or contributing to the development of the ASICBA project. The model developed in ASICBA has been set-up as an integrated model and consists of three main building blocks: (i) risk model, (ii) cost model and (iii) cost-benefit trade-off.

II. GENERAL DESCRIPTION OF THE MODEL

The conceptual design of the approach is depicted in figure 1. First, the potential risk reduction of the safety measure is assessed. The output of the risk analysis feeds into a cost module that contains a database of accident costs. In this module the risk reduction is translated into a decrease in accident costs. This can be considered the safety value (benefit) of the safety measure. Other impacts of the safety measures are also assessed. These concern the costs of the measure itself (e.g. investment costs, operational costs, certification costs) and other benefits (e.g. operational benefits, maintenance optimization and fuel costs reductions). All costs and benefits are then combined in a CBA computation.

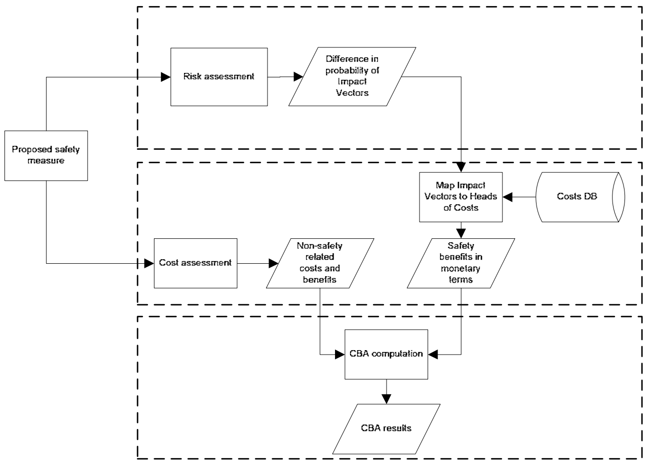


Fig. 1. General outline of the approach.

The following three sections will address the elements of the approach in more detail: the risk model, the cost model and the CBA computation.

III. THE RISK MODEL

In order to assess the impact of certain measures on the level of safety, a scenario-based risk model has been developed. The proposed architecture for this risk model extends conventional risk analysis techniques e.g. fault trees and event trees by introducing a hybrid causal model of event sequence diagrams, fault trees and influence diagrams. Event Sequence Diagrams (ESDs) are used to define the context or scenario within which various causal factors would be seen as a hazard. An Event Sequence Diagram (figure 2) is a flowchart with paths leading to different end states. Each path through the flowchart is a scenario. Along each path, pivotal events are identified as either occurring or not occurring. The event sequence starts with an initiating event, such as a perturbation that requires some kind of response from operators or pilots or one of more systems.

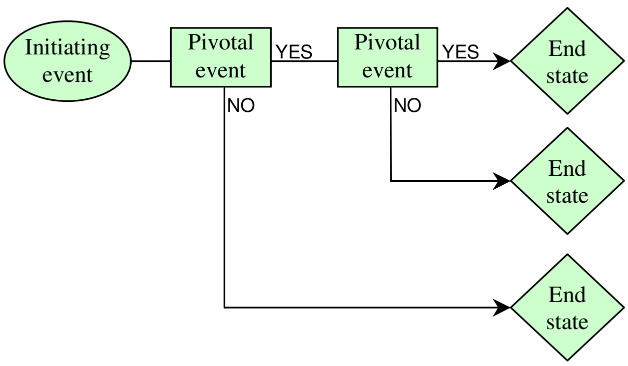


Fig. 2. Event sequence diagram

The ESD represents possible ways in which the effects of an initiating disturbance are handled by the aviation system. Some paths lead to accidents, some paths lead to successful mitigations. A safety improvement is represented in the model by a reduced probability of the occurrence of an accident end state.

Thirty-five generic accident scenarios have been identified, including controlled flight into terrain (CFIT), mid-air collision, runway overrun, etc. These scenarios have been represented in ESDs [17]. An example of such a scenario is provided in figure 3. The development of the accident scenarios was a two-step process. First, a representative sample of accidents and incidents was analysed. Each accident was represented as a detailed sequence of events. These accident scenarios were then generalized and expanded at meaningful branch points following prospective analysis.

Initially the building blocks of the scenarios are kept broad and generic to cover many 'similar' situations. The possible causes and contributing factors of these events are not directly of interest at the scenario level. They are added by means of fault trees and influence diagrams.

For each ESD, probabilities for the initiating events and pivotal events have been derived using the NLR Air Safety Database [18]. The effect of a safety measure on the accident probability is determined by estimating the effect of the measure on the likelihood of occurrence of the initiating events and the pivot events. This will then result in a change in likelihood of the end states.

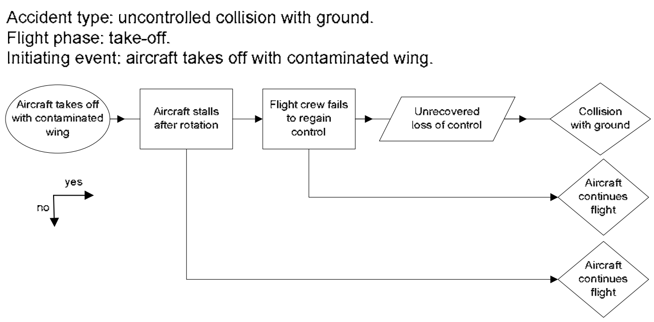


Fig. 3. Example of an accident scenario represented as an ESD

The next step in the approach is to link the risk model to the costs model. The ESDs in the risk model terminate in end states (the diamonds in the diagram). End states describe the consequences of the series of events. To link the risk model to the cost model, the "impact" for each end state is determined. The impact indicates the magnitude of the consequences on all actors involved. Six different groups of actors have been defined:

- Direct
  - Aircraft
  - Occupants
  - Operation of aircraft involved
- Indirect
  - Other airlines/aircraft
  - Airport

- Third party

For each group the severity of the impact is indicated by a number ranging from 0 (no impact) to 4 (maximum impact). Table 1 provides a description for each group and severity level. The impact of an end state is denoted by a vector describing the impact for each group, e.g. (0, 2, 0, 1, 0, 0). The severity levels for the different groups can be independent.

	Direct			Indirect		
	Aircraft	Occupants	Operation aircraft involved	Other airlines	Airport	3rd party
4	>75% damage (beyond economical repair)	80% killed	>4hrs delay (Flight aborted)	Many flights cancelled	Closed for 5 days	Fatalities and damage
3	50% damage	30 % killed	2 hrs < delay < 4 hrs	Many delays	Closed for 2 days	Few injuries and damage
2	15% damage	Some Injuries only	30min < delay < 2hour	Less than many delays	Runway closed for 2 days	Substantial material damage only
1	1% damage	Few injuries	15min < delay < 30min	Little delays	Runway closed for 12 hours	Minor material damage only
0	No Impact	No impact	No Impact (less than 15min delay)	No Impact	No Impact	No Impact

Table 1. Description of the severity levels for each group.

IV. COST MODEL

To assess the costs related to accidents, 24 heads of costs have been identified, including

- Aircraft damage;
- Passenger and crew deaths and/or serious injuries;
- Airline costs for delay and diversions;
- Airport closure;
- Loss of reputation;
- Etc.

For each of the six groups of actors, the relevant heads of costs have been identified. Each group is associated with one or more heads of costs. Next, for each severity level for such a group, the costs are determined. As an example, table 2 shows the aircraft related costs for the different severity levels.

Severity level	Heads of costs		
	Physical damage	Loss of resale value	Loss of use
4	€ 32 m	€ 0	€ 0
3	€ 16 m	€ 4.8 m	€ 3 m
2	€ 5 m	€ 2.4 m	€1.5 m
1	€ 0.3 m	€ 1.3 m	€ 0.4 m
0	€ 0	€ 0	€ 0

Table 2. Aircraft related costs for the different severity levels.

Finally, the estimated costs are allocated to stakeholders. Not all accident costs are borne by the same stakeholder. In order to assess the viability of a safety measure for one particular stakeholder (e.g. an airline) only those costs have to be taken into account that are relevant for this stakeholder. The following stakeholders are distinguished:

- Airline
- Passengers
- Airport operator
- (Safety) authorities
- Third party / Society
- Other airlines
- Insurers
- Aircraft manufacturers

V. CBA COMPUTATION

What results from this whole exercise? The risk model provides an accident likelihood and severity (impact vector) both before and after the application of the safety measure. For each impact vector, the associated accident costs are calculated. By multiplying the difference in likelihood before and after the application of the safety measure with the associated accident costs, the effect of the safety measure is expressed in monetary terms. This can be considered the safety value of the measure, and is a benefit in the CBA calculation.

The safety value is combined with the other required inputs for the CBA calculation. These other inputs are not pre-determined in the model, but need to be provided by the user of the approach. They include:

- investment costs;
- operational costs; and
- operational benefits.

With these inputs, the final CBA calculation can be made. This is usually done over a certain time horizon, e.g. 5 years for airlines, or 25 years for regulators. In the CBA the trade-off between the costs on the one hand and the benefits of the measure on the other hand is made over this time frame. The outcome is expressed in terms of the internal rate of return or the benefit to cost ratio.

VI. APPLICATION TO A CASE STUDY

The approach, as explained above, has been applied to a number of case studies. One of the case studies concerned the cost-benefit assessment of a Flight Operation Quality Assurance (FOQA) programme. A FOQA programme monitors and analysis the in-flight recorded data, in order to improve performance of pilots and maintenance staff, and to improve safety.

The case involved a regional airline with a fleet of 20 aircraft which had to be retrofitted in order to record in-flight data for the purpose of FOQA.

To assess the impact of FOQA on safety, in terms of accident probability per flight, the risk model as described in section III was applied. In potential, FOQA can have effect on almost all elements of the operation. Since the risk model was not completed at this moment, it is decided to limit the scope of this case to the effect of the FOQA program to two accident scenarios describing landing accidents that are a result of an unstable approach. An unstable approach occurs when the aircraft deviates from the proper approach path or approach speed during the approach to landing.

Unstable approaches occur in two Event Sequence Diagrams (ESDs) as initiating events. One of the ESDs is given in figure 4 below. It describes the sequence of events initiated by an unstable approach possibly resulting in a runway overrun or veer-off. The second ESD (not shown in this paper) describes a scenario that potentially results in a collision with ground.

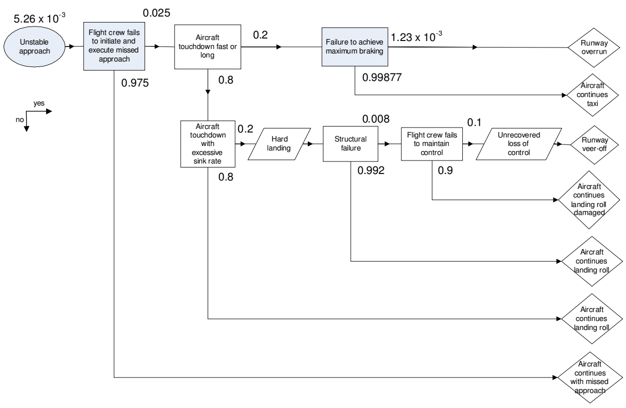


Fig.4. Quantification of an ESD for the FOQA case

The frequency of occurrence per flight of an unstable approach (Initiating Event) and the conditional probabilities of the Pivotal Events for this ESD are shown in figure 4. The values have been derived using NLR's Air Safety Database [18]. The resulting frequencies for the End States are given in table 3.

To assess the effect of FOQA on the probability of occurrences of unstable approaches, in-flight recorded data have been analysed. The data sample comprises aircraft parameter values from before and after the introduction of the FOQA programme. Based on this sample, it was estimated that the introduction of FOQA results in a reduction of the frequency of unstable approaches by 15% (see table 3).

The risk model is linked to the cost model via the impact vector as described in section III. The impact vectors for the relevant end states were estimated using Airclaims data. These impact vectors are given in table 3.

The safety effect of the introduction of the FOQA programme is represented as the difference in probabilities for each End State.

End State	Impact Vector						Probability before FOQA (per flight)	Probability after FOQA (per flight)
Runway overrun	2	1	4	3	1	1	3.23E-08	2.75E-08
Aircraft continues taxi	0	0	0	0	0	0	2.63E-05	2.23E-05
Runway veer-off	2	1	4	3	1	1	1.68E-08	1.43E-08
Aircraft continues landing roll damaged	1	0	3	0	0	0	1.51E-07	1.29E-07
Aircraft continues landing roll	0	0	0	0	0	0	8.42E-05	7.15E-05
Aircraft continues with missed approach	0	0	1	0	0	0	5.13E-03	4.36E-03

Table 3. Estimated safety effect of the introduction of FOQA

Having assessed the safety effect of the introduction of FOQA, the next step towards monetary valuation can be made. For each end state, expressed in the form of the impact vectors, the costs can be calculated. These costs have been estimated based on an analysis by Airclaims of the costs of approximately a thousand accidents [15, 16]. For an airline, these costs concern e.g. the costs of damage to the aircraft in case of a runway-overrun, or the subsequent loss of use of the aircraft when it is damaged. For other actors in society, the costs consist of costs of e.g. injured passengers or the costs of accident investigation.

These costs have subsequently been multiplied with the reduction in end state probabilities and the number of flights of the airline. This resulted in the safety value of the FOQA programme for this airline, amounting to € 0.6 million over a five year period[1]. The safety value for the whole society amounts to € 0.9 million. This value includes the benefits for other stakeholders than the airline itself, like passengers or third parties on the ground.

Subsequently, the costs for the implementation of FOQA have been estimated. This concerns the investment and operational costs, and amount to € 1.5 million for the airline. A large part stems from the investment in retrofitting the aircraft with the equipment, and furthermore in the costs of the application in the airline office as well. The results of the CBA can therefore be estimated, as depicted in the following table.

	Airline*	Society*
Safety benefits	€ 0.6 m	€ 0.9 m
Operational benefits[2]	Yes	Yes
Costs	€ 1.5 m	€ 1.5 m
Cost benefit balance (NPV)	- € 0.9 m	- €0.6 m

* For both the airline and society, the time horizon for the CBA is five years.

Table 4. Case study CBA results

The results presented in this section are intended to illustrate

[1] The value is the net present value over 5 years, at a 4% discount rate. For a description of the principle of discounting and of the notion of net present value, see [19]

[2] It is expected that FOQA will generate operational benefits as well. However, these have not been assessed in this case.

the CBA method that is being developed in the ASICBA project. Due to the limited analysis, the results should not be used to draw any conclusions on the cost-effectiveness of the FOQA programme.

VII. Developments

The development of the concept is largely based on previous studies and experiences [15, 16]. For the quantification, statistical historical data has been used [18]. To determine the validity of the concept and to further specify the model, case studies have been defined in co-operation with users participating in the ASICBA project. In each case study, a measure of interest to the user is analysed according to the method described in this paper. The data necessary for the calculations are largely provided by the user. The user cases play an important role to make the method practically applicable. The computation as provided in section VI, is an example of a case study that has been used to further develop the models and the CBA method.

VIII. Discussion

By applying the approach as described in this paper, the economical viability of introducing a safety measure is determined. The result of a CBA is only one of many inputs in the decision-making process. A cost benefit analysis can not and should not be used to readily provide answers to the question on what investment is appropriate to protect life and health.

In a cost benefit analysis, it is inescapably necessary to compare costs and benefits that can be quantified with relative certainty to costs and benefits that can not. This requires expert judgment on certain matters and consequently includes subjective elements, rather than being rock solid science. The air transport system is incredibly complex and interdependent. A potential danger of cost-benefit analysis can be that it overestimates readily quantifiable variables and disregards variables that are less readily subject to quantitative valuation. Nevertheless, a thoughtful cost benefit analysis, including explanations of the critical assumptions behind the analysis of scientific evidence, will improve the information relied on in the decision-making process. A proper cost-benefit analysis is a useful tool for getting an understanding of the gain and loss associated with safety investments.

It is quite allowable that safety costs money. The method described in this paper enables the decision makers to maximize the effect of the safety investment even if the benefit to cost ratio is negative. By comparing different safety measures, the most efficient alternative can be determined.

IX. Conclusion

In this paper an approach to assess the effects of safety improvement measures is described. The approach offers the opportunity to make a balanced trade-off between safety and investment costs.

Key notion in the approach is that safety effects can be expressed in monetary terms, and thus can be compared with the associated costs of safety measures. The safety effect in monetary terms can be considered the safety value (benefit) of the measure. Combined with other costs and benefits, like investment costs, operational costs and operational benefits, the CBA calculation can be made. Hence, the economical viability of introducing a safety measure is determined, which serves as one of the inputs in the decision-making process.

User cases play an important role to make the method practically applicable.

References

[1] ECORYS, 2003, Cost-benefit analysis of a Safety Culture, Rotterdam (in Dutch).

[2] B. De Brabander, L. Vereeck. 2003. Cost benefit analysis of road safety investments in Belgium, case study for a seat belt reminder system, report RA-2003-16, Steunpunt Verkeersveiligheid bij Stijgende Mobiliteit, Diepenbeek, Belgium

[3] R. Elvik, A.H. Amundsun. 2000. Improving road safety in Sweden, Institute of Transport Economics, Oslo

[4] R. Elvik. 2001. Cost Benefit Analysis of Police Enforcement, working paper 1 ESCAPE project.

[5] ETSC. 2003. Cost effective EU transport safety measures, European Transport Safety Council, Brussels

[6] Austrian Road Safety Board et al, 2005, ROSEBUD: Road Safety and Environmental Benefit-Cost and Cost-Effectiveness Analysis for Use in Decision-Making, Deliverable 6, Testing the efficiency assessment tools on selected road safety measures.

[7] TRL et al., 2001, Cost-benefit analysis of measures for vulnerable road users, PROMISING project deliverable 5.

[8] CPB/NEI, Evaluation of infrastructure projects, guideline for cost-benefit analysis, 2000, The Hague (in Dutch).

[9] European Commission, Guide to cost-benefit analysis of investment projects, 2002, DG Regional Policy, Brussels.

[10] Layard, R., (ed.), Cost-Benefit Analysis, Penguin Books, England, 1997

[11] CPB, The effects of the expansion of Schiphol airport, a key figure cost-benefit analysis, 2002, the Hague (in Dutch).

[12] Helios Technology, Cost-Benefit Analysis of Free Route Airspace - Final Report, 2002, Eurocontrol, Brussels.

[13] European Tripartite Group, EGNOS Multimodal costs and benefits, 1999, Brussels.

[14] The ASICBA website, available at http://www.asicba.org

[15] ASTER project (A.L.C. Roelen, R. Piers, R.J. Molemaker and P. Hayes), Handbook for conducting cost benefit analysis of safety measures in air transport, NLR-CR-2001-609.

[16] DESIRE project (M. Pedrali, A.L.C. Roelen), Causal & Functional Models of the Air Transport System for Cost Benefit Analysis of Safety Measures in Aviation, paper presented at Human Issue in Aviation Safety, HIAS 2001, September 26-28, 2001, Toulouse, France.

[17] A.L.C. Roelen and R. Wever, Accident scenarios for an integrated aviation safety model, NLR-CR-2005-560, NLR Amsterdam, 2005.

[18] G.W.H. van Es, Overview of the most important accident/incident data sources in the NLR Air Safety Database, NLR Memorandum LV-2003-09, NLR Amsterdam, March 2003.

[19] Eurocontrol, Guidelines for the economic appraisal of EATMP projects - the effective use of cost-benefit studies, 2000, Brussels.

Analysis of runway incursion trends at U.S. FAA towered airports

Joseph N.D. Dodoo, Isoken T. Aighewi, Ibibia K. Dabipi

***Abstract*— Our analyses of runway incursion data for all US towered airports over a 21 year period from CY84 to CY04 has shown that during the early 1990's, the number of runway incursions was declining. However, beginning in 1994 that trend reversed, and runway incursions have since increased each year. The range of incursions from 1988 to 2001 was from 186 to 431. We have also studied the runway incursion data for 17 US major airports located all across the continental US. The airports were placed in two groups, namely, Group (1) and Group (2). Data were available for all the airports in Group (1) over the 21 year period, except for Baltimore Washington International where the runway incursion data obtained was from CY87. After analyzing the data for the Group (1) airports it was decided to increase the number of airports as well as broaden the FAA regions. Nine additional major towered airports from various FAA regions were selected. These airports are placed in Group (2). For this group the runway incursion data could only be obtained for CY97 to CY02. For some of the airports in Group (1), the number of incursions was down over a ten year period from CY90. While variations occurred in the numbers the overall trend for these airports was downward until 2001. Other airports in this group showed a high rate of runway incursions over the same period. The Group (2) airports had considerably higher rates of runway incursions during the six years for which the data were analyzed. When extrapolated over 21 years, for the purpose of comparison with those in Group (1), the incidence of runway incursion at each of the airports far exceed any in Group (1).**

***Index Terms*—Deviation, Runway incursions, Runway incursion Trends, Runway safety**

Manuscript received February 2, 2006. This work was supported by NASA-Langley Laboratory through Grant # 123456 under the Chesapeake Information Based Aeronautic Consortium (CIBAC). Joseph N.D. Dodoo is an Assistant Professor of Physics at the University of Maryland Eastern Shore, MD 21853, USA. (E-mail: jddodoo@umes.edu) Phone: 410-651-6033, Fax: 410-651-7739. Isoken T. Aighewi is an Assistant Professor of Environmental/Geospatial Applications, at the University of Maryland Eastern Shore, MD 21853, USA. (E-mail: tiaighewi@umes.edu) and Ibibia K. Dabipi is a Professor and Chairman of Engineering and Aviation Sciences at the University of Maryland Eastern Shore, MD 21853, USA. (E-mail ikdabipi@umes.edu)

I. Introduction

When two aircraft or an aircraft and a vehicle inadvertently get too close together on an airport runway, whether the encounter results in a collision or not, it is known as a runway incursion. According to The Federal Aviation Administration (FAA) such occurrences have been on the increase since 1994. When a suspected runway incursion occurs, it is the responsibility of the air traffic controller to report the incident to the FAA. The FAA investigates the situation and determines whether or not it was a runway incursion and who was responsible for the mistake. An example of a runway incursion would be two aircraft colliding on a runway. However, there are less obvious runway incursions. For example, if an aircraft is on a take-off roll, and another aircraft or vehicle crosses the same runway at that time, then that would also be a runway incursion. Although there was no actual accident, the two aircraft did not have the necessary separation. Here is an account of an actual runway incursion as depicted in an NTSB animation [1]. *On April 1, 1999, just after 2 o'clock in the morning, Korean Air flight 36 and Air China 9018, both Boeing 747s, nearly collided on runway 14 Right at the Chicago O'Hare International Airport. Air China had just landed and was rolling out on runway 14 right when the tower controller instructed Korean Air to taxi into position and hold. After Air China exited the runway at taxiway T10, the tower controller instructed the flight to turn left on taxiway Kilo and cross runway 27 left. The tower controller then cleared Korean Air for takeoff. As the airplane was rolling down the runway, Air China deviated from its assigned taxi route and taxied on to runway 14 Right. The Korean Air captain saw the 747 taxiing on to the runway but it was too late to stop. Instead, Korean Air 36 lifted off earlier than normal and banked left to avoid striking Air China. The two aircraft, carrying 382 people, missed colliding by about 80 feet.*

The first recorded airplane fatality in history was the September 17, 1908 crash of the Wright Flyer III flown by Orville Wright. One of two propellers separated in flight, tearing loose the wires bracing the rudder and causing the loss of control of the aircraft. Since then records of all aviation accidents/incidents worldwide, have been documented. Databases of these accidents abound and are managed by national governments as well as a number of private organizations, [2-5]. While records of accidents have existed for a long time in various formats, The Federal Aviation Authority, (FAA) database on runway incursion began in 1999

[5]. The records indicate that runway incursions at all U.S. towered airports from FY 1999 to FY 2002 increased. However, the records also show that although the number of incursions has an increasing trend, there were periods when the total number of incursions for all FAA regions dropped.

II. Causative factors leading to Runway incursions

The Federal Aviation Authority (FAA), responsible for safe and efficient aeronautical operations within the National Airspace System (NAS), has established appropriate locations at airports for deployment of infrastructure designed with the primary purpose of preventing runway incursions. The FAA has defined runway incursion as: "Any occurrence at an airport involving an aircraft, vehicle, person, or object on the ground that creates a collision hazard or results in loss of separation with an aircraft taking off, intending to take off, landing, or intending to land." While the majority of runway and surface incursions are minor with little or no damage to equipment or injury to personnel, some have resulted in very serious accidents. When a runway or surface incursion occurs it is usually the result of a number of factors. The runway incursion that caused the greatest loss of life in aviation history occurred at Tenerife Airport in the Canary Islands on March 27, 1977. In that accident two Boeing 747 aircraft were involved and a total of 583 people perished. Accounts of that accident have been widely disseminated. Factors such as extremely poor visibility conditions, communication failure, airport knowledge and cockpit procedures for maintaining orientation have been sited as major contributors to runway incursions.

III. Accident and Incident Databases

As an independent Federal agency charged by Congress with investigating every civil aviation accident in the US, the National Transportation Safety Board (NTSB) has investigated over 124,000 aviation accidents since its inception in 1967 [6]. The reports from these investigations are contained in the agency's database and are available to the public via the NTSB website. In order to further its objective of promoting the open exchange of safety information and thus improve aviation safety, the FAA established in 1996, the National Aviation Safety Data Analysis Center, (NASDAC). The NTSB databank follows a prescribed format. By selecting "current synopsis" the reader is taken to a very detailed account of the accident. A report includes information such as the date of the event, event type, damage type, phase of flight, aircraft make and model as well as information on the operator and pilot-in-command. The NASDAC system enables users to perform integrated queries across multiple databases and search warehoused data and display information in useful formats. While the NASDAC database contains the same information as the NTSB database, it is a summarized version extracting the most pertinent elements, and employing its own numbering system. NASDAC also classifies runway and surface incursions by type as follows; OE/D - Operational Error/Deviation, PD - Pilot Deviation, or V/PD - Vehicle/Pedestrian Deviation.

IV. Analysis of Incursions at FAA towered airports

Our study **e**xamined the incursion data for all towered airports over a 17- year period from calendar year 1988 to calendar year 2004 with emphasis on 2001 to 2004. The runway incursion data for all FAA towered airports during the period from 1988 to 2001 is shown in figure 1.

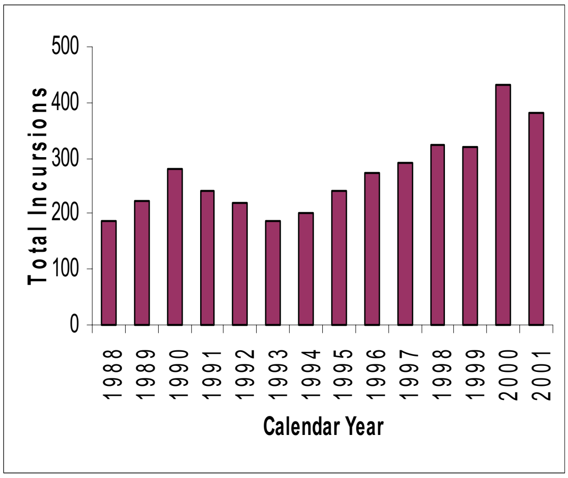


Figure 1, Runway incursions at all FAA towered airports from 1988 to 2001

The data clearly show that in the first three years of record, incursion rose steadily followed by a further three years of steady decline. In 1994 runway incursions began to rise again. The increase continued for another six years until 2000. The slightly reduced value in calendar year 1999 is statistically insignificant with respect to 1998. However, the overall trend in that seven year period is upward. Again in 2001 runway incursion total in all nine regions resumed a downward trend. Further examination of the data shows that the reduction in incursions continued during the period from 2002 to 2004. Figure 2 shows the region by region total runway incursion for this period.

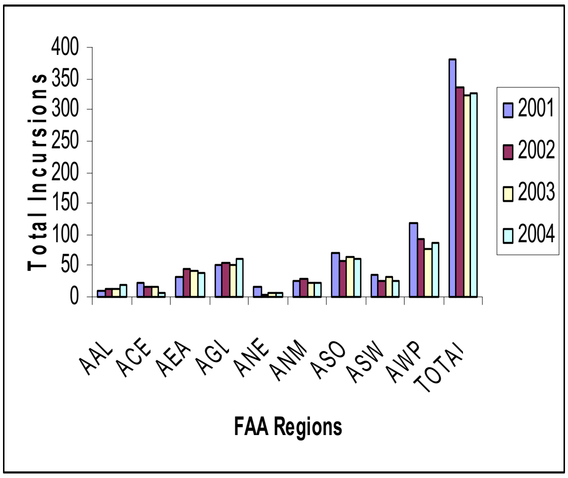


Figure 2 Runway incursions for all nine FAA regions' towered airports during FY2001 – FY2004

Key: AAL - Alaska, ACE - Central, AEA - Eastern, AGL - Great Lakes, ANE -New England, ANM- Northwest Mountain, ASO - Southern AWS - Southwest, AWP - Western Pacific.

The data for 2001 is included in the chart to facilitate comparison. The greatest number of runway incursions occurred in the Western Pacific, (AWP) region. Southern, (ASO) and the Great Lakes, (AGL) regions came in close second and third in the total number of incursions. When compared with the years prior to 2001, however, the overall trend was downwards. When separated by type as defined above, the runway incursions during the period between 1994 and 2004 are shown in figure 3.

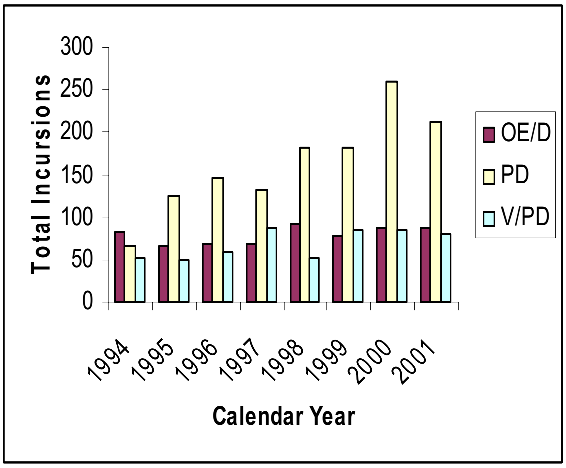


Fig. 3 Runway incursions by type

It is clear that except in1994, Pilot Deviation accounts for the majority of the runway incursions. The number of PD incursions peaked in 2000. This then accounts for the highest total runway incursions for that year as show in Figure 1. Between 1999 and 2001 the deviations due to Vehicle/Pedestrian and Operational Error, were about the same. The picture is very much the same when the incursion data are examined for the period from 2002 to 2004 as shown in figure 4.

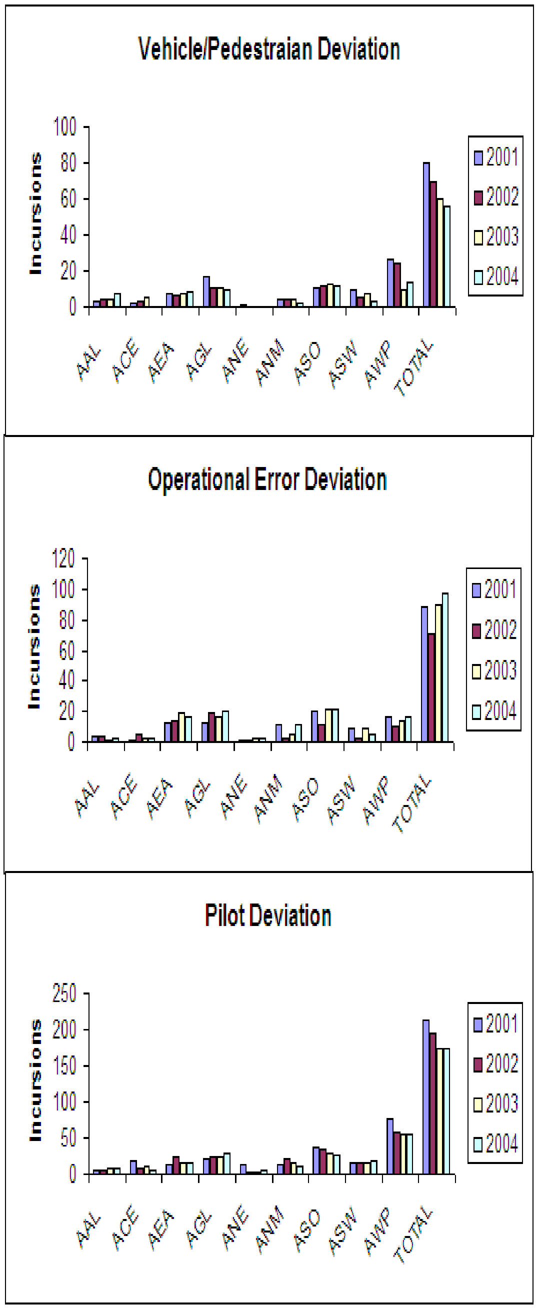


Figure 4, Runway incursions by Type between CY01 and CY04

With this breakdown, it is easy to see how the Western Pacific region has the greatest number of runway incursions since it outranks all the other regions in each of the three types of incursion, namely; Vehicle/Pedestrian Deviation, Operational Error Deviation and Pilot Deviation.

V. A case study of randomly selected airports

We have analyzed the runway incursions data for seventeen randomly selected towered airports. They were selected on the basis of their status as major international airports. In addition all FAA regions except Alaska are represented. Table 1 shows the airports studied listed by their International Civil Aviation Organization (ICAO) codes, FAA region, State and Airport names.

TABLE I
MAJOR FAA TOWERED AIRPORTS STUDIED

ICAO	FAA Region	State	Airport Name
Group I			CY84 – CY 04
ATL	ASO	GA	Hartsfield Atlanta International
BWI	AEA	MD	Baltimore-Washington International
BOS	ANE	MA	Boston Logan International
DEN	ANN	CO	Denver International
JFK	AEA	NY	John F. Kennedy International
LAX	AWP	CA	Los Angeles International
MIA	ASO	FL	Miami International
ORD	AGL	IL	O'Hare International
Group II			CY92 –CY02
DFW	ASW	TX	Dallas/Forth Worth International
FXE	ASO	FL	Fort Lauderdale International
LGB	AWP	CA	Long Beach-Daugherty Field
PHX	AWP	AZ	Phoenix Sky Harbor International
SAN	AWP	CA	San Diego International
SFO	AWP	CA	San Francisco International
SNA	AWP	CA	John Wayne Orange County
STL	ACE	MO	Lambert St. Louis
VGT	AWP	NV	North Las Vegas

The runway incursions data were analyzed for a 21-year period from CY84 to CY04 for all the airports except the Baltimore-Washington International where data were unavailable from 1984 through 1996. In addition the year to year incursion data could not be retrieved for nine of the airports for the period from CY84 to CY99 Accordingly; we have separated Table 1 as Group (1) and Group (2) to accommodate the differences. The total runway incursions for Group (1) for the period studied is shown in figure 5a. The total incursions for Group (2) airports during CY97 to CY02 are shown in figure 5b.

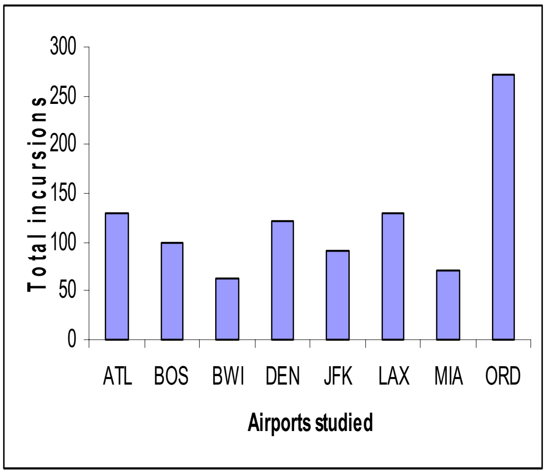


Fig. 5a Total runway incursions for Group (1) Airports from CY84 to CY04

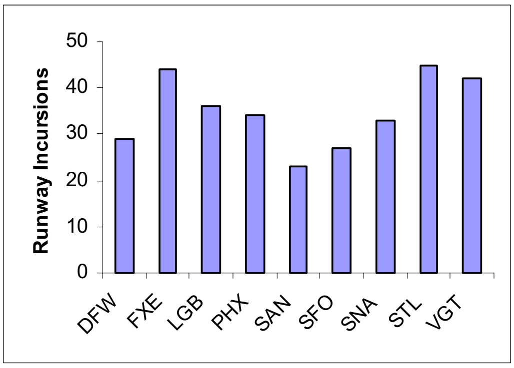


Fig. 5b Total runway incursions for Group (2)

The data for Group (1) show that Chicago O'Hare International (ORD) had more runway incursion during this period with 271 incursions, about three times those at John F. Kennedy airport and about four times the number of incursions at Miami International. Data were unavailable for BWI until 1987. That being so the total number of incursion for that airport is about the same as those at MIA taken over the 21 year period. For the airports in Group (2) Fort Lauderdale Executive (FXE), Lambert St. Louis (STL) and North Las Vegas (VGT) had the highest runway incursions during the period. Considering annual incursion averages, O'Hare International's value was 13. By contrast Fort Lauderdale Executive's average over the six year period was 7. Thus the incursions at FXE were considerably higher since by extrapolation the yearly average could top 25 when taken over the 21-year period for Group (1) airports.

For Group (1), BWI and MIA had average annual incursions of 3. The lowest average value in Group (2) was at SAN with a value of 4. Again by the same reasoning this value could be 14 when extrapolated over 21 years. Thus the runway incursions were considerably higher at all airports in the Group (2) than those in Group (1).

The runway incursions at the individual airports for Group (2) are shown in figure 6.

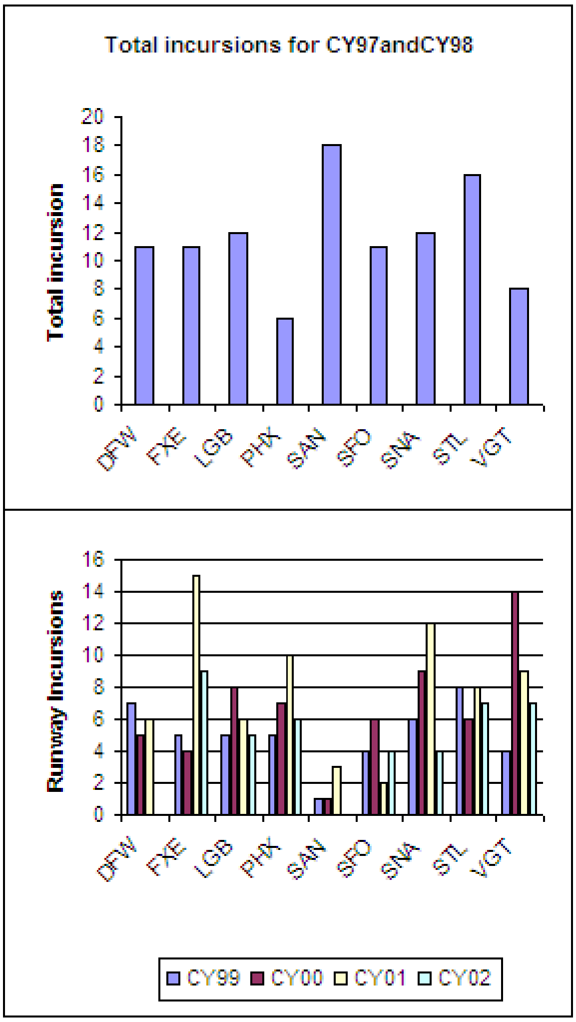


Figure 6 Annual runway incursion data for the Group (2) Airports evaluated for CY97 to CY02

The corresponding data for the Group (1) airports are shown in figure 7.

Nearly all the airports in Group (2) had high incidence of runway incursions during the six year period as shown in figure 6. Six of the nine airports in this group are in the Western Pacific FAA region, AWP. At their peak the annual rate of runway incursions was 8.7 per 100,000 operations. For the airports in Group (1) both Chicago O'Hare and Hartsfield – Jackson Atlanta International airports had a wide range of runway incursion during the 21-year period. However the highest mode for the incursions was at Hartsfield-Jackson Atlanta. Variability in the data occurs for all the airports. BWI airport is almost equal to MIA in runway incursions considering that the data for BWI is for 18 years instead of 21 years for the rest of the airport. However, the greatest number of incursions for MIA occurred in 1997 reaching 18 compared with only 2 for the same year at BWI. In the following year, however, the situation reversed. The runway incursions at MIA dropped exactly 50% to 9. The incursions at BWI increase 100% to 4. After 1999 the incursions at BWI again dropped to 2 and thereafter never rose above this figure. In 1999, MIA reduced its incursion even further to 1. This great improvement could not, however, be maintained. The incursions increase steadily every year after that until 2002 though the numbers in themselves were small. The number for year 2003 and 2004 were 2 and 3 respectively. 1990 was the worst year ATL with a total of 25 runway incursions. After that, however, the number of incursions dropped by more than a third and stayed around that figure for another four years. In 1996 there was again a sharp increase in incursions. That sudden increase must have raised some eyebrows; for in the five years following this there was a steady drop in the numbers. In 2003 the number increased again to 12 followed the very next year by a drop to 4 incursions; a 75% reduction.

The runway incursions at Chicago O'Hare (ORD) peaked in 1989 at 27. There followed three consecutive years in which the incursions dropped, with the third year (1991), experiencing the most reduction. In 1992 the incursions again increased. There were 14 incursions that year followed by two years in which the incursions in which the increase dropped by 50% each. The fluctuation in the number of incursions continued with 22 runway incursions in 1997. After 1997 the incidence of incursions declined but the numbers fluctuated from 9 in 1998 down to 3 in 2004, but not before rising to 10 incursions in 2002.

By far John F. Kennedy (JFK) airport showed a consistently lower numbers in runway incursions. The highest number of incursions occurs in 1990. Before and after 1990, however, the number of incursions was small compared with all the other airports in this study. As can be seen from Table 3 with a range in the incursion of 15 which and mode 3, there were at least three years in which only one runway incursion took place. The number of incursions at Denver (DEN) airport shows more variability than most. Although the maximum for any one year was 11, this figure occurred rather too often. The only two consecutive periods in which reduction was maintained were in 1995 and 1996. At all other times a drop one year was followed by an increase in the following year.

Los Angeles International (LAX) airport had its highest number of runway incursions (16), in 1987. Following that year the number of incursions shows a downward trend to be broken with an increase in 1997. The reduction, however, resumed in 1998 and although fluctuating, it stayed no higher than 6 until 2004. Boston Logan airport data begin with one incursion in 1984, none reported in 1985 or 1986, and then suddenly rises to 10

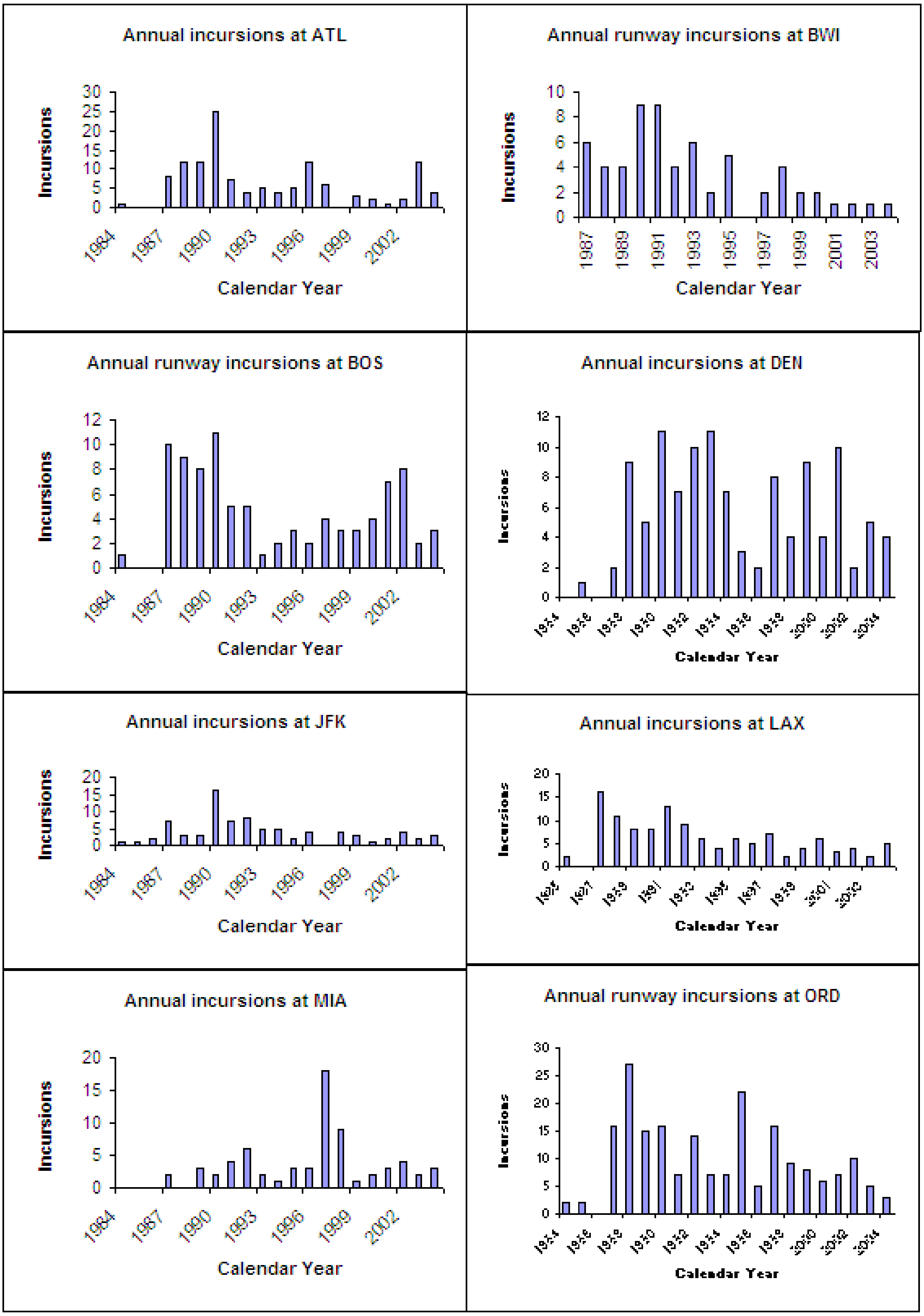


Figure 7 Annual runway incursion data for the Group (1) Airports evaluated for CY84 to CY04

incursions in 1987. There were two consecutive years (1998 and 1990), when the numbers declined. In 1991 the number of incursions again rose reaching a peak of 11. The reduction in the number of incursions continued after for another three years to only one incursion in 1993. However, the upward trend began immediately the following year with variations to 8 incursions in 2002. The numbers reduced for the last two years with three incursions in 2003 and four in 2004.

VI. Discussions and Conclusion

The runway incursions data nationwide from 1988 to 2004 showed that there was a steady rise from 1988 to 1990. This was followed by three years when the number of incursions gradually decreased. However, in 1994 the number of incursions began to climb reaching as high as 431 in 2000. The total number of operations at all the towered airports in 1998 was 62,501,059, [4]. There were 187 runway incursions that year. Two years later in 1990 the figure was 65,506,291 and 281 incursions. These figures correspond to a 4.8% increase in airport operations and a 50% increase in the number of runway incursions. The number of operations in 1999 was 68,672,240. In that year there were 321 runway incursions. Pilot deviation accounted for 182, (57%) of the total. In 2000 the number of operations was slightly lower at 67,480,097. The number of incursions increased to 431. Pilot deviation in 2000 was 259, a 60% share of the incursions. The number of operations though dropped by 1.7% over the previous year did not translate to a reduction in the number of runway incursions. From 2002 to 2004 the number of incursions was lower for each year over that of 2001. Increases in the incursions are greatly enhanced by those due to Pilot Deviation as seen from Figure 4.

The picture for the seventeen randomly selected airports is somewhat different from the national trend. The number of runway incursions at those airports showed a downward trend; at least for the airports in Group (1). Variations were to be found in the data, however, in that at some of these airport, especially at Denver and Boston Logan International airports, one or two years of reduction in incursions was immediately followed by one or two years in which the numbers rose. For Group (2), the airports in the Western Pacific FAA region had the greatest number of incursions in all three categories. A review of a number of the FAA incidence reports on the incursions in these areas showed no more different causes than those at other regions. It is clear, therefore, that regardless of the location of the airport, runway incursions nearly always occur as a result of one or a combination of the well documented factors mentioned in section II. It must be pointed out that in June 2001, the subcommittee on Aviation Hearing on Runway Incursions, Focusing on the Technology to Prevent Collisions proposed some technical initiatives [5] to mitigate runway incursions. It would be necessary to reinvestigate the data at these airports at some future date so as to see if these initiatives have had an impact.

It would be unrealistic to expect that runway incursions can be completely eliminated at any towered airport. One has only to refer again to the FAA definition to realize that it would take a super human effort to completely eliminate runway incursions. However, when FAA recommendations and NTSB initiatives are implemented at all towered airports, and if all operators comply, it will be possible at any rate to reduce runway incursions to some acceptable lowest value. We suggest an acronym “ALAP” to mean AS LOW AS POSSIBLE. The FAA could adopt a number that can be assigned to “ALAP”. If that is done it would provide a target that is not only achievable but could be surpassed at several airports. We anticipate a continuous reevaluation of the data starting from two years after full implementation, to determine whether the runway incursion trends have declined.

ACKNOWLEDGEMENT

We gratefully acknowledge financial support by NASA-Langley under the Chesapeake Information Based Aeronautic Consortium (CIBAC), Grant # NOC1-03033.

REFERENCES

[1] www.ntsb.gov/events/2000/incursion/incur_video.htm

[2] NASDAC Brief Report: www.nasdac.faa.gov/pls/portals/STAGE_BRIEF_REPORT

[3] Federal Aviation Administration Website (2002). Runway Safety. Washington, DC.

[4] Federal Aviation Administration Website (2002). Runway Incursions-A Growing Problem at Controlled Airports. Washington, DC.

[5] Subcommittee on Aviation. Hearing on. Runway Incursions, Focusing on the Technology to Prevent Collisions. www.house.gov/transportation/aviation/06-26-01/06-26-01

[6] NTSB Accident Synopses / National Transportation Safety Board (NTSB) http://www.ntsb.gov

[7] Gero, D. Aviation Disasters : the world's major civil airliner crashes since 1950. - Sparkford : PSL, 1993. - 224 ISBN 1852603798

[8] ICAO Circular Aircraft Accident Digests / International Civil Aviation Organisation (ICAO)[1)]

[9] World Airline Accident Summary (CAP-479) / Civil Aviation Authority (CAA). - London :

[10] Denham, T. World directory of airliner accidents. - Sparkford : PSL, 1996. - 320 p. ISBN 1852605545

[11] Kimura, C. World Commercial Aircraft Accidents 1946-1992. - Livermore, CA : Lawrence Livermore National Laboratory, 1993

[12] www.planecrashinfo.com

[13] World Airline Accident Summary, 1946-1972, British Civil Aviation Authority

[14] International Civil Aviation Organization (ICAO) Adrep Summary 1988 - 1994

[15] International Civil Aviation Organization (ICAO) Circular Aircraft Accident Digests

[16] Air Disaster Volumes 1,2,3,4, Macarthur Job, Aerospace Publications Pty Ltd., 1994-1999

[17] Aviation Disasters, The world's major civil airliner crashes since 1950, Second Edition David Gero, Patrick Stephens Ltd., 1996

[18] Destination Disaster, Paul Eddy, Elaine Potter, Bruce Page, Times Newspapers Ltd., 1976

[19] National Transportation Safety Board, Accident Investigation Reports, various reports
National Transportation Safety Board, Accident Synopses

[20] Civil Aeronautics Board Accident Investigation Reports, various reports

Safe Flight 21 Database Compliance to RTCA Requirements and Implications for Safety

Isoken T. Aighewi, Ibibia Dabipi and Joseph Dodoo

Abstract— **Aerodrome surface maps have been an indispensable component of aviation safety, operations and management as it plays a central role in surface navigation and/or situational awareness (SA), as well as aerodrome operational efficiency. This study evaluated fifty of the eighty two randomly selected U.S airport databases for compliance to the RTCA Requirements, and; evaluated the implications of the RTCA compliance status to aviation safety. The study showed that the compliance at these U. S. airports ranged from a low of 8% to a high of 46% with an average of 31%. This study thus indicates the need to step up efforts at improving the accuracy and currency of the AMDBs in meeting with the free flight concept of the SF-21 program objectives for aeronautic safety and aerodrome operations efficiency.**

Index Terms—**Aerodrome Mapping Database, Compliance, Radio Technical Commission for Aeronautics (RTCA), Surface Navigation, U.S Airports.**

I. Introduction

AERODROME surface maps plays a central role in surface navigation and/or situational awareness (SA), as well as aerodrome operational efficiency. This has become even more significant in recent times as the demand for air travel continues to grow. Worldwide air travel is expected to more than double by 2017 [1]; such growth will undoubtedly put even more strain on existing aerodromes (airports, vertiports, heliports and sea-plane aerodromes) and thus challenges for system designers and the sundry users of Aerodrome Map Databases (AMDBs)-pilots, air traffic controllers, aerodrome planners and managers. There is thus an urgent need to front-end any safety and operations-related problems that may result from the numerous applications of these databases. Presently, most of the existing databases used as guidance materials are primarily applicable to air navigation rather than for surface applications. However, these deficiencies are being addressed in part by the Radio Technical Commission for Aeronautics (RTCA) Aerodrome Mapping Database (AMDBs) requirements developed by its Special Committee 193 (SC-193) and the European Organization for Civil Aviation Equipment (EUROCEA) Working Group 44 (WG-44) in the DO-272 document [2]. This document provides details of the minimum data requirements of immediate priority to aerodrome surface map data originators and integrators. Similarly, the Safe Flight 21 (SF-21) Program-a joint government/industry partnership, has been initiated with a focus on improving aerodrome map databases for the Free Flight concept envisaged to take advantage of the evolving communications, navigation, and surveillance (CNS) technologies. However, while progress is being made to improve the AMDBs for the SF-21 program, so far only eighty two of the more than 520 towered U.S airports are presently available for compliance evaluation to the RTCA requirements.

Manuscript received February 2, 2006. This work was supported by NASA-Langley Laboratory through Grant # 123456 under the Chesapeake Information Based Aeronautic Consortium (CIBAC). Isoken T. Aighewi is an Assistant Professor of Environmental/Geospatial Applications at the University of Maryland Eastern Shore, MD 21853, USA. (E-mail: tiaighewi@umes.edu) Phone: 410-651-8266, Fax: 410-651-7739. I. Dabipi is a Professor and Chairman of Engineering and Aviation Sciences at the University of Maryland Eastern Shore, MD 21853, USA. (E-mail ikdabipi@umes.edu) and J. Dodoo is an Assistant professor of Physics at the University of Maryland Eastern Shore, MD 21853, USA. (E-mail ddodoo@umes.edu)

II. Safe Flight-21 AMDBs Compliance

The RTCA requirements, though not all inclusive, consist of 31 aerodrome features in seven classes of immediate priority, including information of their respective geometry, attributes and attribute coding. The classes are: runways, helipads, taxiways, aprons, vertical structures, construction areas and quality data. Although emphasis on all features and attributes (377 total data elements) of the AMDBs requiring compliance may differ from one aerodrome to another, they are deemed critical and thus require standardization and validation in order to ensure safety and avert human factor issues-the single major cause of accidents in aviation [3]

Our study:

- Evaluated some randomly selected U.S airport databases for compliance to the RTCA Requirements and also;
- Evaluated the implications of the RTCA compliance to aviation safety.

We acquired the RTCA DO-272 document [2] from Alliance Telecommunications Industry Solutions (ATIS) and the SF-21 database from NASA-Langley Laboratory. A template for the aerodrome elements and their attributes were developed in Excel and fifty of the eighty two available towered U.S. airport databases were randomly selected. For each airport, geospatial layers representing each airport features and their corresponding attributes were opened in ArcGIS (ESRI ©) environment and the data fields matched with those in the DO-272 document (See Table 1 for the features required excluding their respective attributes).

TABLE I
AERODROME MAPPING DATA BASE (AMDBS) FEATURES FOR RTCA COMPLIANCE

AIRPORT FEATURES	
Runways	Stop Way
Runway Intersection	Clear way
Threshold	TLOF
Runway Marking	Helipad Threshold
Center Line	Taxiway Segment
LAHSO	Taxiway Shoulder
Arrest Gear Location	Taxiway Guidance Line
Runway Shoulder	Taxiway Intersection Marking
Taxiway Holding Position	Aerodrome Reference Point
Exit Line	Vertical Polygon Structure
Frequency Area	Vertical Polygon Point Structure
Apron	Vertical Point Structure
Stand Guidance Line	Vertical Line structure
Parking Stand Location	Deicing Area and Survey control Area
Parking Stand Area	Construction Area

For each feature, all required fields, values and codes were verified for their presence and correctness. Those features stated as “unknown” or “not applicable” or outside the range required were reported as lacking compliance. The percent compliance was calculated based on the proportion of the total 377 data elements that met the RTCA requirements. Furthermore, the features not in compliance were collated and their frequencies computed and documented.

The result of the fifty airport databases evaluated for compliance is shown in Figure 1. Full name of the airports are available at http://www.photius.com/wfb2001/airport_codes.html.

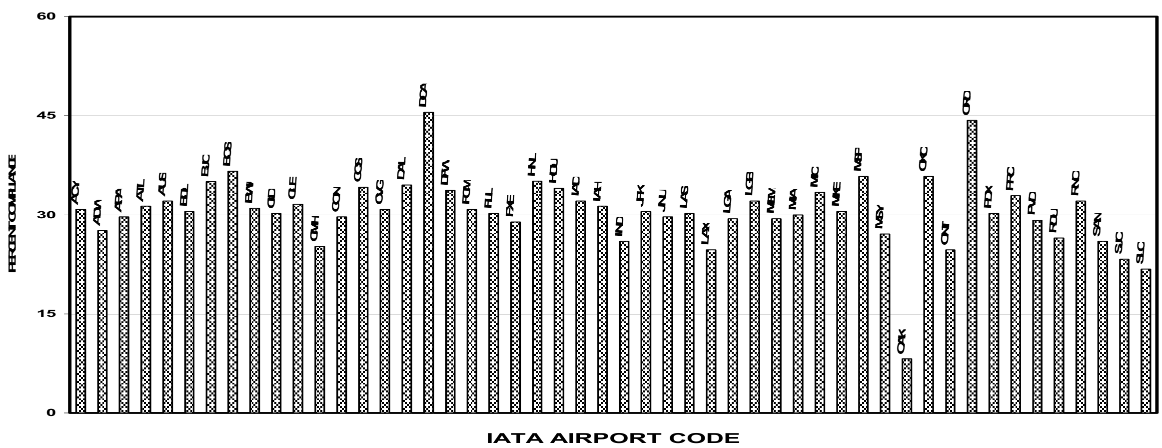


Fig. 1. U.S. Safe Flight 21 (SF-21) Aerodrome Mapping Database (AMDBs) Compliance to RTCA Requirements (n=50).

The result shows that compliance is generally inadequate for all airport AMDBs evaluated. Mean compliance was 31% with a range of 8% for Oakland (OAK) Airport in California to 46% for Reagan National Airport (DCA) in Washington DC. They include: *HRES*-horizontal resolution of coordinates defining the features; *Revdate*-date of last revision or generation of source data, *HACC*-horizontal accuracy of entity at 95% CE; *Arptid*-ICAO aerodrome location indicator; *integr*-integrity of data in the aeronautical data processing chain from origination process to present data manipulation process; *Feattype*-feature type; *source*-name of organization or entity that supplied the data; *VACC*-vertical accuracy of entity at a 95% LE, *VRES*- vertical resolution of coordinates defining the feature; *Objectid*-object identifier, *material*-predominant surface type of Final approach and Takeoff area (FATO); and *status*-permanent state of runway in corresponding takeoff/landing direction. Of the major features, attributes and codes not in compliance, eleven were most frequent (Figure 2)

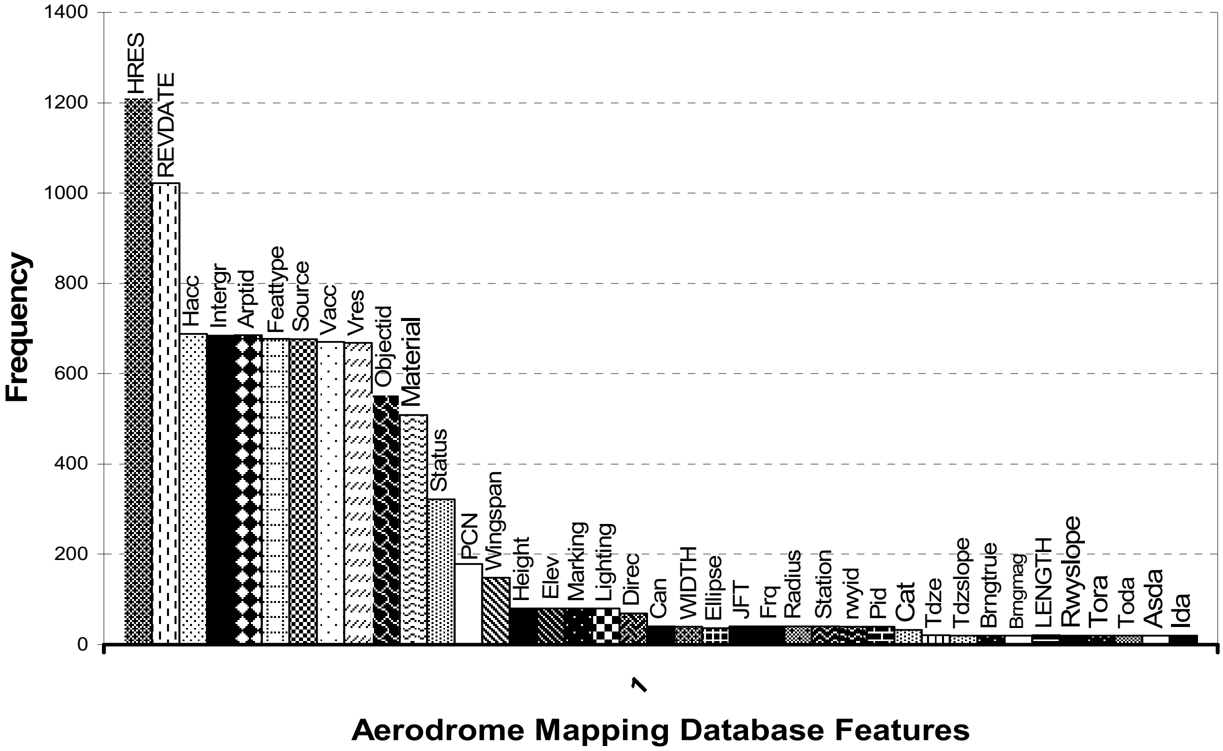


Fig. 2. Frequency of AMDBs features without Compliance to RTCA Requirements

I. COMPLIANCE ANALYSIS AND SAFETY ISSUES

In addition to the low compliance of the AMDBs evaluated, in general the eleven features that were mostly out of compliance have severe safety implications. These data elements fall within four general categories:

Lack of Horizontal and Vertical Accuracy of coordinates defining features (HACC and VACC): This omission affects the geospatial extent of such features as runways, taxiways, blast pad, and vertical structures that can compromise the ability of pilots to maintain safety margin during surface maneuvers. Such omissions compromise data integrity and can pose problems relating to terrain obstacle awareness and avoidance. It could also contribute to failure of system function that could result in catastrophic, hazardous or major failure conditions. While these attributes were found to be the most frequent omission with the airport features, those of horizontal accuracy is of greater significance.

Lack of database currency (Revdate or date of last revision or generation of source data): These data items were the second most frequently missing information from the databases evaluated. They represent a major omission that could mislead database users in several applications with respect to information currency. For example, the lack of information on current construction or modification of the airport terminal can pose major safety problems in surface navigation. In fact, timeliness of AMDBs should be available to users at all time and changes in between standard update cycles provided to users via Notice to Airmen (NOTAM) [2]

Lack of data Integrity (Integr or data integrity in the aeronautical data processing chain from origination process to present data manipulation): The absence of such documentation compromises the database quality assurance requirement in accordance to DO-200A/AE-76 [4]. it prevents knowledge of the presence or absence of critical data that can impact continuous safe operations of aircraft that would be severely at risk with potential for catastrophe.

Lack of horizontal and vertical resolution of coordinates defining features indicated by VRES and HRES: Similar to the accuracy problem previously indicated, the errors resulting from resolution of airport features can also compromise safety margin if users are unaware of such lapses and are unable to identify aerodrome features by the traditional see-and-avoid method.

IV. CONCLUSION

We have shown that the compliance at U. S. airports ranges from a low of 8% to a high of 46%. These values can be attributed to the infancy of the SF-21 program and highlights the need to improve database integrity and currency. Given the projections that worldwide air travel is expected to more than double by 2017, the urgency to comply with the RTCA requirements as a means of standardizing the databases particularly in view of the free flight concept is essential.

ACKNOWLEDGEMENT

This study was supported by NASA-Langley Research Grant # NOC1-03033 under the Chesapeake Information Based Aeronautic Consortium (CIBAC). We thank NASA Scientists-Robert A. Kudlinski and Emory T. Evans, and gratefully acknowledge our consortium Coordinator, Gilbert Haynes and fellow researchers at Morgan State University
and Bowie State University for their cooperation and useful exchanges.

REFERENCES

[1] NASA-FAA Roundtable, 2005. Http:// www.iasa .com.au/folders /Safety_ Issues/others /NASA-FAAsafety.html

[2] RTCA, 2001. User Requirements for Aerodrome Mapping Information, DO-272, Prepared by RTCA Special Committee 193 and EUROCEA Working Group 44

[3] L.J. Prinzel, A.T. Pope T, F. Freeman.2002. Physiological Self-Regulation and Adaptive Automation. The International J Aviation Psychology 12(2):176-196

[4] RTCA, 1998. Standards for Processing Aeronautic Data. DO-200A/AE-76. Prepared by RTCA Special Committee 181.

Session 9

Methods and Processes

The DFS ADAM methodology assessing capacity benefits of ATM systems

T. Bierwagen and A. Tautz

Abstract— **A methodology is presented to assess the initial and long term capacity benefits comparing two different ATM systems. The theoretical background for capacity assessment is given. Objectives, hypothesis and experimental design of an ADAM investigation are detailed. Results of a methodology validation exercise are described. Applicability, comparability and transferability of the ADAM methodology are outlined.**

Index Terms— **air traffic management, ATM system, capacity, real time simulation, benefit analysis**

I. Introduction

THE air traffic system guarantees for economic growth in Europe. A major constituting factor of this system and the intended growth is the European air traffic management (ATM) system. The objective is to ensure the safe, orderly and expeditious flow of air traffic throughout Europe – civil traffic and in some countries also military traffic. For this service, obviously safety is of highest priority. Besides safety, efficiency and productivity gain more importance. In parallel economic pressure on the air navigation service providers is rising due to the European commission single European sky initiative and due to the ongoing privatization process of the air navigation service providers.

This situation is answered Europe-wide with a mixed approach of new operational concepts, new technologies and new responsibilities. In economical terms, the productivity and the capacity of the overall air traffic management system becomes the short to medium term focus of the activities. However, especially the capacity benefits of introducing new technologies or new ATM systems were estimated with a low reliability up to now. In this context, analytical methods based on specification documents were used. Europe-wide, air navigation service providers experienced in the past 10-15 years that this way of securing their major investment is insufficient.

Based on this experience, DFS German Air Navigation Services developed its own new methodology to assess the capacity benefits of new ATM systems in the context of the "Advanced Display Assessment and Validation Measurements" (ADAM) Project. Its target is to provide reliable figures on capacity benefits prior to the operational use of an ATM system. The ADAM methodology uses a comparative approach. It is based on coupled real time simulations with the current and the future ATM system. In addition, a live traffic observation study assures comparability and transferability of the results gained in the simulation environment. This approach can be used not only for capacity benefit assessment, but also for other purposes e.g. in the safety context or in the certification context of ATM systems.

Following this introduction, some methodological issues with regard to capacity assessment in operational ATM systems and ATM systems prior to cutover are discussed. Based on this discussion, the objectives and hypothesis of the ADAM methodology are presented. The experimental design is described. Based on a large number of single observations, the interdisciplinary evaluation of the hypothesis and the synthesis of the results are detailed. An insight into the applicability, comparability and transferability of the methodology to other ATM systems or other locations within the ATM domain is given. The use of this methodology within the European ATM system certification process is proposed. Next steps within DFS conclude this paper.

II. Methodology Review

A. Lessons learned from ATM system introduction

It is well known that in general for introducing a complex technical system (like an ATM system) it is a major difference between the initial and the long term capacity benefits. This is due to training effects of the operators. Starting with some initial training while working with the system they move ahead on the training curve. After a while (i.e. some month or years) they reach a stable position on this training curve: they became an expert on operating the system.

In past introduction processes of ATM systems, the capacity benefits often where estimated based on technical specifications and other documents. The operator interacting with the system was hardly recognized and the learning mechanism was not taken into account. Due to this the estimated capacity benefits where very optimistic. Finally, these estimations could not be implemented. The reality benchmarking the estimations some years later (when the ATM system was in use for some time) always failed –

Manuscript received February 24, 2006.

T. Bierwagen is with DFS Deutsche Flugsicherung GmbH Research and Development division, Am DFS-Campus 5, 63225 Langen, Germany (corresponding author, phone +49-6103-707-5794, fax +49-6103-707-5741, email thomas.bierwagen@dfs.de).

A. Tautz is with DFS Deutsche Flugsicherung GmbH Research and Development division, Am DFS-Campus 5, 63225 Langen, Germany (email andreas.tautz@dfs.de).

sometimes in order of one magnitude.

Resulting from this perception, today human factors aspects like operator's acceptance and optimized training methods are considered as significant contributing factors to the capacity benefits.

Operational concepts and operational procedures are recognized as another important aspect for introducing a new ATM system. Analyzing the operational procedures applied by individual operators within an operational ATM system shows that a significant number of operators stick to the operating procedures they initially learned. Moving from one ATM system to another lets them adapt their operating procedures only in areas where the new ATM system forces them to do so. The fact that very often a new ATM system asks for completely different ways of operating it than the prior one is neglected. It is simply a matter of convenience to act in this way. This approach is possible as long as there are only gradual changes in the way of operating the ATM system and especially the division of tasks between human and machine remain more or less the same. If a larger more evolutionary step appears, this approach of moderate accommodation to the new ATM system fails. As a consequence, the ATM system capacity decreases or does not increase as much as intended.

Taking into account these considerations and according to the DFS' needs the ADAM methodology assesses two different aspects of capacity: (i) the initial capacity benefits of an ATM systems against another ATM system at cutover and (ii) the long term capacity benefits in case a (new) ATM system is in use for some time.

The follow up question on how to reduce the difference between both capacity values is not answered within the ADAM methodology. It has to be picked up as a training issue within the process of introducing a new ATM system or even within the whole system development process. The ADAM methodology does only snap shot the actual capacity benefits in a diagnostic manner, it does not provide means to improve the capacity benefits within these processes.

B. Methods for capacity assessment in ATM systems

How do we measure capacity and capacity benefits within an ATM system? This question keeps a complete unit of the EUROCONTROL organization in motion: the performance review unit (PRU). PRU assesses the current status and future trends in terms of capacity of the individual member states' ATM systems. The assessment works in a bottom up approach based on indications given from the individual area control centers in the member states. Within an area control center, the capacity values are further substantiated down to a sector level with maximum values. Capacity is commonly defined as a number of aircraft per time unit (e.g. one hour) that the sector is capable to handle [1]. Unfortunately a sector capacity value is not a purely objective figure, but results from a subjective process: the value is gained through past traffic figures, which are adjusted by a subject matter expert's judgment on retrospective experiences working in this sector. However, this mixture of objective and subjective methods to find a maximum sector capacity is most frequently used. Results of the PRU assessment are stated in the periodic EUROCONTROL performance review report [2].

Another methodology used for capacity assessment is the capacity analysis method (CAPAN) [1]. It was developed by EUROCONTROL also and is used in several member states. DFS adjusted and further developed this methodology for its own needs. CAPAN determines via fast time simulations for each sector task load values depending on traffic volumes. Underlying is a fixed, task oriented time table providing a duration time for each task to be carried out. The maximum capacity of a sector is defined by passing a 70% task load value (i.e. within one hour of traffic having more than 42 minutes of added task times) (see [1], p. 5). The CAPAN methodology is frequently used around Europe also. It is applicable especially for well known, today's ATM systems with a solid and settled task and time scheme. If the time scheme is not well known and / or not stable, this methodology can not be applied. First, the new time scheme has to be assessed in real time simulations or in live traffic operations. In addition, the new ATM system may come along with new tasks not yet considered in the methodology, whereas others may no longer be needed. So for estimating new ATM systems' capacity the CAPAN methodology is unsuitable.

The integrated task analysis (ITA) is another methodology. It was also developed by EUROCONTROL [3][4]. ITA assesses the cognitive task load by means of observations and interviews. Based on this data, a maximum capacity value is derived. The ITA methodology is not commonly used. ITA picks up the correct idea that in addition to visible tasks activities (as used in CAPAN) the cognitive task load is a relevant part of the controller's work. At the same time, ITA captures criticism by the fact that cognitive processes are neither very good observable nor reportable in interviews.

Again EUROCONTROL developed the INTEGRA methodology [5]. INTEGRA uses fast time simulations. In addition, it follows a similar approach than ITA. INTEGRA distinguishes different human and system actors within the air traffic management task like executive controller, planning controller, MTCD system or data link system. In addition, it considers the operational concept of the individual ATM system by allocating different ATM tasks to different actors. Activities of these actors are triggered based on events which evolve from the actual traffic situation. Based on the actor's activities, for each actor and each time step an information processing load (IPL) is calculated. IPL is in a way comparable to the ITA cognitive task load, but it does not claim to model cognitive processes. The maximum capacity is defined by the traffic situation where the first actor – either human or machine – passes its individual maximum value. Using this definition of maximum capacity, it becomes possible not only to assess the overall system capacity, but also to identify the actor responsible for restricting the capacity. This information allows for optimizing the overall system capacity by means of varying the task-actor allocation. A major disadvantage is the need to explore in calibration experiments the individual actor's maximum values.

Another approach of assessing capacity is a purely subjective one. The idea is to use the psychological concept of task load and work load. Within this concept, task load is the objective loading of the traffic situation, whereas work load is the subjective impression of loading of each individual actors caused by the objective loading. Work load can be assessed with a number of well known questionnaires. Examples are the instantaneous self assessment technique (ISA) used by EUROCONTROL [6] and rating work load on a five step scale or the NASA task load index (NASA TLX) [7] used by DFS and rating on a ten step scale. Both questionnaires follow the idea to assess the actual work load of the operator in discrete time steps like 5 or 10 minutes. By choosing good points in time for the measurements and correlation of the ratings assessed with the traffic situation, a capacity value can be derived.

C. *Selected approach within the ADAM methodology*

The approach to measure capacity within the ADAM methodology does not rely on one single method detailed above, but it combines several of them.

Based on a comprehensive observation study which is carried out in daily operations as well as in the real time simulations, tasks and activities are assessed in terms of frequency and duration. Different sector suites and traffic volumes are observed systematically. As a result, the distribution of (observable) tasks in terms of task load can be calculated. These results can be attributed to different factors like working position, traffic volume or ATM system used. It is important to state that in contrast to the ITA methodology no cognitive processes are measured within this observation study – as they are not observable.

Work load is assessed by means of the NASA TLX. A NASA TLX questionnaire is presented to each controller within daily traffic and within each real time simulation run every ten minutes, self-rating the individual work load of the past 10 minutes. This data can be attributed to different factors as well. In addition, by considering the time line this data can be correlated to the results of the observation study. As a result, indications on conform effects or discrepancies between both sources of data are identified.

Based on this data and taking into account some other factors like training aspects and assumptions on cognitive aspects, a relative capacity value comparing two different ATM systems can be calculated. This value answers the question on the initial capacity benefits in a quantitative way and states statistical evidence.

In contrast, the long term capacity benefits can not just be calculated. They depend much more on assumptions and conclusions rather than on calculations. However, discussing the long term capacity benefits gives important indications for issues to be considered to reach this benefit.

With regard to the methodology chosen, the controller's task load is composed of an observable and a non observable part. Each of these parts can be optimized to increase the capacity of the overall system, assuming that an optimization of the task load components will increase the capacity.

For the observable tasks, optimization indications can be given relatively easy. If e.g. one task has only a small amount of overall observed time, even a major optimization of this task will not contribute a lot to the overall capacity. If in contrast another task has a large amount of observed time, even a minor optimization may be relevant. One concrete example could be input of data into the ATM system: if in ATM system 1 data is written on paper and in ATM system 2 data is input into the system, it is obvious that one task substitutes the other. However, if the input process is much more time consuming than the writing process, this may indicate that the HMI to input the data may be improvable.

For the non observable tasks like monitoring or problem solving, indications can be given on the capabilities to shift those tasks from the human operator to the system component. These questions affect issues like acceptance and trust into the system rather than physical optimization. Within this discussion it is important to consider the INTEGRA ideas of not overloading one actor by shifting tasks from another one. As a consequence, clear indications can be derived to achieve the long term capacity benefit. However, in terms of concrete numbers no fixed and calculated value but only an estimation can be given.

III. EXPERIMENTAL SETUP

An ADAM investigation in principal consists of three different steps:

1. observation study in live traffic (ATM system 1)
2. real time simulation (ATM system 1)
3. real time simulation (ATM system 2)

Of course, the experimental setup tries to establish identical and comparable conditions in each of the three steps.

One major shortcoming of real time simulation investigations in ATM is the limited number of participants and experimental conditions. These limitations often do not allow for statistical analysis of the data. As this is mandatory to estimate the reliability of the data, the experimental design is set up in a way that allows for statistical analysis. It implies a significant number of participants working under different conditions.

A. *Objectives*

The single objective of an ADAM investigation is

- To assess the capacity benefits of an ATM system relatively to another ATM system

To compare both ATM systems, it is mandatory to operate both systems in identical environment, i.e. identical airspace. However, airspace structure adaptations to the new ATM system may arise to facilitate the full benefits of the new ATM system.

As a side effect, two further aspects are to be assessed to fulfill the objective: (i) proof of the feasibility of the operational concept and the operational procedures for the new ATM system and (ii) initial indications on the new ATM system acceptance.

B. Hypothesis

Hypothesis for an ADAM investigation are as follows:

1. The initial capacity benefit of the new ATM system in comparison to the old ATM system is positive
2. The long term capacity benefit of the new ATM system is higher than the initial capacity benefit for this system.

C. Experimental design

The observation study in live traffic is used to assure comparability of real time simulations with the situation in real life. In other words, it checks the realism of the simulations. Due to this fact, the observation study in live traffic is not included in the experimental design. However, this study observes identical airspaces and identical positions utilized also in the real time simulations.

For both real time simulations, the experimental design is detailed below.

1) Variables

For the investigation, several factors in both ATM systems are varied systematically with different values. These independent variables and the associated values are detailed in table 1.

TABLE I
INDEPENDENT VARIABLES

Variable	Values
Traffic volume	80% to 160% of a peak summer day 2005, varied in 6 steps
Airspace (sector)	4 sectors
Position	Executive controller, planning controller
Experience	trainee, experienced controller

Any data within the simulation system are recorded electronically. In addition, one sector suite per simulation run is used for the observation study. The observable tasks are assessed along a system of categories which is detailed in table 2. The assessment uses computer aids and logs start point and end point of each task observed via time stamps in a log file.

TABLE 2
OBSERVATION CATEGORIES

Category
Radio transmission (r/t)
Coordination within the sector team
Coordination by telephone
Verbal coordination between teams
Use of the touch input device for data input
Use of the radar screen for data input
Handling strip bay / Use of the main data window for input

From the beginning of each simulation run, each participant is presented a NASA TLX rating sheet every ten minutes. The participants have to rate the work load of past ten minutes on the six NASA TLX scales. At the end of each simulation run – lasting between 75 and 85 minutes – another NASA TLX rating sheet is presented. Now the participants have to rate the work load of the whole simulation run on the six NASA TLX scales.

2) Subjects

To achieve trustable results, at a minimum 24 air traffic controllers are needed to participate. At least 8 of them need to be experienced controllers and 8 of them need to be trainees. They are split in 3 teams with 8 controllers each. Each group in itself may either consist of experienced controllers or of trainees. The mixture of trainees and experienced controllers is needed to assess training effects with regard to operational procedures as well as familiarity with modern technologies. The experienced controller should be licensed in the sectors they staff during simulation runs.

Each participant is to attend 36 measured simulation runs and several training runs. The amount of simulation runs last at about 20-25 working days for each participant.

3) Experimental plan

Each participant has to face each of the 6 different traffic volumes three times per ATM system. This ends up in the experimental plan shown in table 3. The experimental schedule varies the order of presentation of different runs depending on the experimental group.

TABLE 3
EXPERIMENTAL PLAN

	Traffic volume [%] (100% = 2005 peak traffic)					
System	80	100	115	130	145	160
ATM system 1	9 runs	9 runs	9 runs	9 runs	9 runs	9 runs
ATM system 2	9 runs	9 runs	9 runs	9 runs	9 runs	9 runs

4) Experimental constraints

For the results, the experimental constraints with regard to the ATM tasks are of major importance. In principal, they have to be kept comparable as much as possible.

The operational concept describing how the ATM systems are operated must be available for both ATM systems. They may differ with regard to actors, functions, support tools, task distribution and team processes.

The airspace is to be comparable for both ATM systems. It may differ in terms of airspace organization like procedures or airspace structure. It may not differ in terms of lateral and vertical dimension. Letters of agreement may differ only if new support tools or support functionality is available in one of the ATM systems investigated.

The working positions used should all together cover the same ATM tasks and the same piece of airspace. However, within each ATM system they may be composed in different ways fulfilling the overall criteria.

IV. RESULTS OF THE METHODOLOGY VALIDATION EXERCISE

To proof the validity of the ADAM methodology, a validation exercise was carried out. Due to the amount of data collected, below only a few major results are detailed.

In a first step, data was refined with regard to completeness and comparability between the different groups of subjects. In addition, the first 20 minutes of each measured run were deleted because the complex traffic picture evolved slowly over time from the beginning to about 15-20 minutes. At the end of each measured run the data was deleted beyond 70 minutes, so that an evaluation duration of 50 minutes was established.

A. Results of the task analysis (observation study)

1) Comparison of the two ATM systems

The observation study assessed the start and end times of seven different categories of tasks (see table 2). Based on this data, the mean and cumulated duration in seconds for each task is calculated. In addition, results can be obtained with regard to parallel tasks. To gather results in terms of capacity, the data is to be analyzed with regard to the amount of traffic. This was done by slicing the data as well as the traffic figures into 10 minutes intervals.

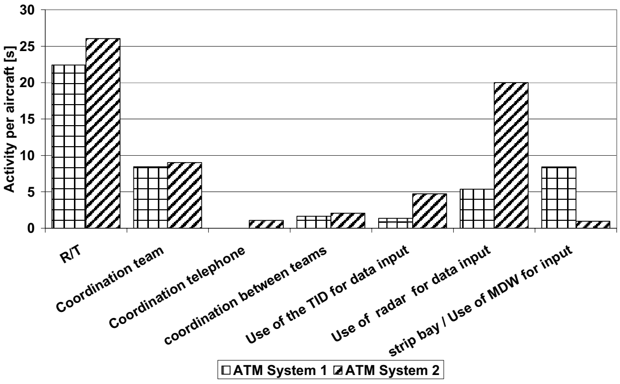


Fig. 1. Activity per aircraft per 7 task categories, executive controller (all sectors), ATM system 1 and ATM system 2

The result of this process is shown in figure 1 for the executive controller, cumulated over all sectors. It is obvious that the major part of the task is the r/t communication. A second view shows clear differences because of a rising level of automation: ATM system 2 asks for input of all data into the system, whereas ATM system 1 works with paper strips. Due to this, the mean duration of the system interaction tasks (radar, TID) rise whereas the strip marking task is more or less omitted. Statistics show a significant difference between both ATM systems ($p < 0.01$). The activity in ATM system 2 is higher than in ATM system 1.

A similar result is shown in figure 2 for the planning controller. Especially, this figure shows a clear automation effect in reduced telephone coordination activities within ATM system 2, which allows for electronic coordination mechanisms. In addition, again an increase in system interaction tasks is shown. Statistics show the identical significant difference between both ATM systems ($p < 0.01$). The activity in ATM system 2 is higher than in ATM system 1.

2) Comparison to real life

The question of realism compared to real life is essential for the validity of the investigation. For that purpose the observation study carried out during the real time simulations was reproduced in live traffic environment as well. The identical sectors were observed, but due to operational constraints other subjects than those participating in the real time simulations. The comparison of the results from this observation study in real life compared to the data available for the identical ATM system 1 in real time simulations is shown for the executive controller in figure 3 and for the planning controller in figure 4.

Both figures show that within the real time simulation on ATM system 1 more or less identical activities per aircraft

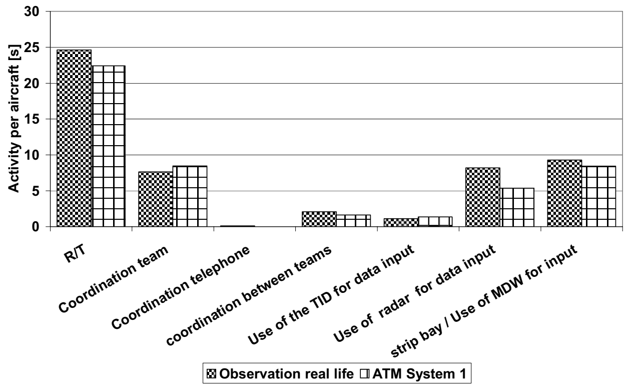


Fig. 3. Activity per aircraft per 7 task categories, executive controller (all sectors), observation real life and ATM system 1

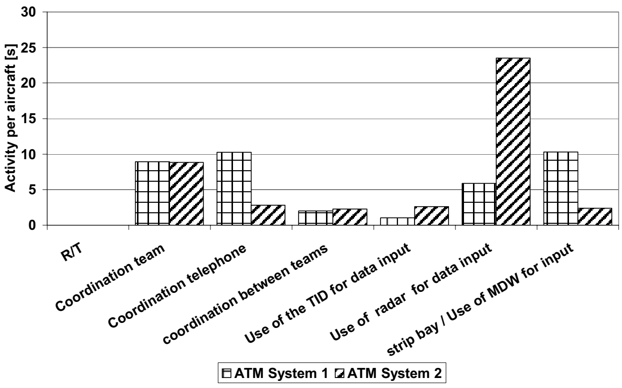


Fig. 2. Activity per aircraft per 7 task categories, planning controller (all sectors), ATM system 1 and ATM system 2

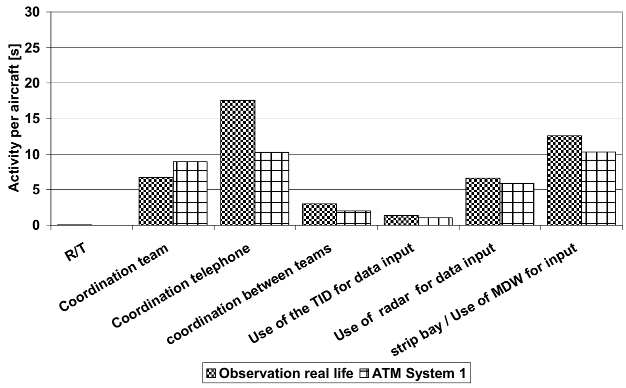


Fig. 4. Activity per aircraft per 7 task categories, planning controller (all sectors), observation real life and ATM system 1

were observed. The only major difference is shown in figure 4 for the telephone coordination at the planning controller position ($p < 0.05$). This difference is due to the fact that the coordination activities with other facilities were drastically simplified: the coordination partner was available at the phone very quickly in the real time simulations, and he accepted much more coordination requests than in real life. In summary it can be stated that obviously the real time simulations showed a high degree of realism and due to that finding results from these simulations can be transferred into the real life environment.

B. Results of the work load analysis

1) Comparison of the two ATM systems

To assess the subject's work load, NASA TLX ratings were used with all six NASA TLX scales. However, for the analysis two of the six scales were not used: whereas "physical demand" is not relevant to an air traffic controllers work (i.e. he does not handle heavy goods), the ratings on the "effort" scale indicated that the reversed scaling was misunderstood by several subjects.

The NASA TLX analysis over all subjects and conditions for the four remaining scales is shown in figure 5. It shows a significant difference ($p < 0.01$) between the ATM system 1 and the ATM system 2 on all four scales. So the results from the work load assessment clearly support the results from the task load assessment reported above.

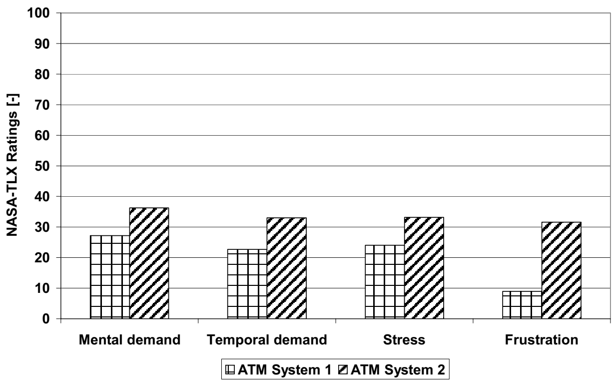


Fig. 5. NASA TLX ratings over all subjects and conditions, ATM system 1 and ATM system 2

2) Comparison to real life

The comparison of the work load data between the real time simulations and real life is not possible because of the different traffic situations in real life. In contrast to the activities captured in the observation study and which could be normalized easily, this is not possible for NASA TLX ratings. Instead, another analysis technique using a classification scheme with regard to traffic volume per 10 minute slices can be used.

C. Validation results

The initial application shows that the methodology delivers reliable results with regard to capacity benefits. In addition it allows for a detailed view on different aspects of change in terms of the technical system, the human operator and the co-operation of both. The impact of a proper operational concept is emphasized.

The workload assessment is clearly responding on different levels of traffic as well as on different ATM systems. This is true for situation awareness and trust in automation as well. However, it is to be considered that the level of maturity of the different ATM systems investigated is clearly differing. This fact may bias the measurements.

The observation study shows the shift in tasks due to the rising level of automation. Effects which were expected (e.g. "more system input" versus "paper strip interaction") could be proven clearly. In addition, a minimum time budget for interaction with each aircraft could be identified. Mechanisms were demonstrated to increase the amount of traffic that can be handled.

The adherence to the letters of agreements show the quality of service provided. Differences between the ATM systems were not found from the data. Finally, it was possible to identify and quantify clearly and reliably capacity benefits of a new ATM system.

V. TRANSFER OF RESULTS

Due to the comprehensive approach, the ADAM methodology can be used not only for capacity assessment, but also for certification issues of ATM systems. Even as a first step into training aspects, the methodology is helpful. Also, it can be easily transferred into other than DFS environments on a European level. By that, the power of this approach is clearly proven. However, the effort to gain these results has to be traded off against the advantage of early and reliable figures on capacity and efficiency.

DFS is going to refine the ADAM methodology for use in all further major implementation projects to ensure a safe, expeditious and efficient transition into the future single European sky.

REFERENCES

[1] Flynn, G., Benkouar, A. & Christien, R.: *Pessimistic sector capacity estimation*. EEC Note 21/03. Eurocontrol: Bretigny sur Orge, 2003.

[2] Eurocontrol (Ed.): *Performance Review Report 8*. Eurocontrol: Brüssel, 2005.

[3] Kallus, K. W., Barbarino, M. & Van Damme, D.: *Integrated Task and Job Analysis of Air Traffic Controllers Phase 1*. Eurocontrol: Brüssel, 1998.

[4] Kallus, K. W., Van Damme, D. & Dittmann, A.: *Integrated Task and Job Analysis of Air Traffic Controllers Phase 2: Task analysis of En-route Controllers*. Eurocontrol: Brüssel, 1999.

[5] Gingell, R., Strachan, C., Taylor, A., Kinnersly, S. & Fox, S.: *INTEGRA Metrics & Methodologies Execution Phase – Final Report*. Eurocontrol: Brüssel, 2005.

[6] Jordan, C.S.: *Experimental study of the effects of an instantaneous self assessment workload recorder on task performance*. Report No. DRA/TM(CAD5)/ 92011. Farnborough: Defence Evaluation & Research Agency, 1992.

[7] Hart, S.G. and Staveland, L.E.: Development of a NASA TLX (NASA Task Load Index): Results of empirical and theoretical research; In P.A. Hancock and N. Meshkati (Eds.), *Human Mental Workload*, Amsterdam, 1988.

A Policy Compatibility Analysis for Airport Slot Allocation

Michael A. Madas and Konstantinos G. Zografos

***Abstract*—A substantial amount of research work has been documented in the literature with the aim to review and critically assess the alternative demand management approaches aiming to allocate scarce airport slots. During the last few years in Europe, there is a policy debate on the need for a drastically revised regulatory framework and strategies for slot allocation. One of the primary policy concerns involves the policy compatibility of slot allocation strategies in different airport clusters. The ultimate objective of this paper is to develop and apply a methodological framework with the aim to assess the degree of policy compatibility of each slot allocation strategy for each airport cluster based on its profile characteristics and the embedded policy priorities. In that respect, the paper will also address: i) the slot allocation strategies, and ii) the airport typology development process.**

***Index Terms*—Airport slot allocation, airport typology, policy compatibility.**

I. Introduction

Options towards increasing airport capacity through the expansion of existing and building of new airports or alternatively through air traffic management programs address only the supply-side of the airport congestion problem. The overwhelming increase in congestion and delay externalities, in conjunction with technical constraints in providing sufficient capacity, stress the need for a thorough examination of demand management instruments as a potential vehicle for bridging the gap between supply and demand for scarce airport slots[1].

This increasing imbalance between capacity and traffic has resulted in congestion and delay figures that have drawn the attention of aviation policy makers and regulatory bodies investigating alternative means of coping with the aviation capacity gridlock. During the last few years, the European Commission pursues a radical revision of the existing slot allocation regime towards the adoption of market-driven allocation mechanisms. In response to this, a substantial amount of research work has been documented in the literature [1]-[6]. In this effort, one of the primary policy concerns lies on the appropriateness and policy compatibility of alternative slot allocation strategies in different airport settings / profiles, as well as the selection and implementation of a slot allocation strategy that will fulfill certain policy criteria.

The major objective of this paper is to develop and apply a methodological framework for the systematic assessment and selection of the most appropriate slot allocation strategy for different airport profiles on the basis of certain policy criteria and priorities (i.e., policy compatibility analysis). To that end, the following intermediate goals have been tackled: i) to identify and briefly discuss a series of airport slot allocation strategies based on the integration of administrative measures, rules, and market-based instruments, and ii) to present a systematic airport typology development process (i.e., airport cluster analysis) that will describe and classify the various airport profiles according to specific clustering variables.

II. Approach

The policy compatibility analysis aims to formulate

Manuscript received January 14, 2006.

M. A. Madas is a Ph.D candidate of the Department of Management Science and Technology, Athens University of Economics and Business, Evelpidon 47A and Lefkados 33, 113 62, Athens, Greece, Tel.: +30 210 8203673-5, Fax: +30 210 8203684, e-mail: mmadas@aueb.gr.

K. G. Zografos is a Professor of the Department of Management Science and Technology, Athens University of Economics and Business, Evelpidon 47A and Lefkados 33, 113 62, Athens, Greece, Tel.: +30 210 8203673-5, Fax: +30 210 8203684, e-mail: kostas.zografos@aueb.gr (corresponding author).

[1] Slots were introduced as an expression of capacity in late 1960's to deal with congestion and delay problems and represents (according to the European Commission Regulation 95/93 and its recent amendments) the permission given to a carrier to use the full range of airport infrastructure necessary to operate an air service at a slot-controlled airport on a specific date and time for the purpose of landing or take-off.

an evaluation framework that will provide guidance for the selection of the most appropriate and compatible (with respect to specific policy criteria) slot allocation strategy for various types / profiles of airports. In order to perform a policy compatibility analysis, three basic methodological steps should be followed (Fig. 1): i) identification of the alternative slot allocation strategies and options (Step 1, i.e., "what" should be evaluated), ii) development of a typology of airports (i.e., airport cluster analysis) (Step 2, i.e., "where / in which airport context" the identified strategies should be evaluated), and iii) assessment of the policy compatibility of the identified strategies with the elicited airport profiles / clusters (Step 3, i.e., "how" the identified strategies should be evaluated for each airport cluster with respect to their conformance / compatibility with policy criteria).

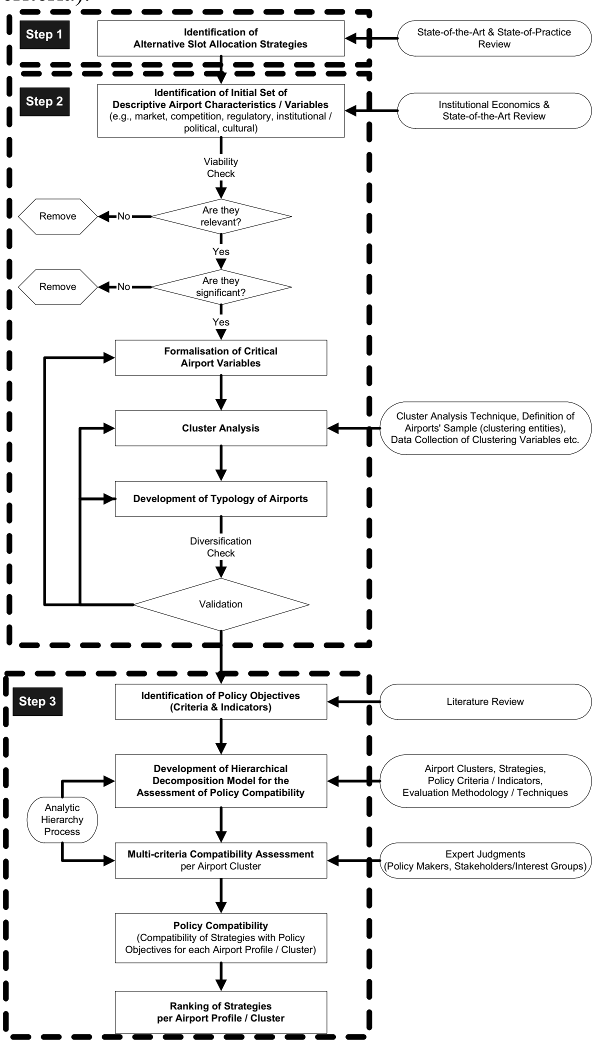


Fig. 1. Methodological approach

As illustrated in Fig. 1, the policy compatibility study was initiated with the identification of the alternative slot allocation strategies. Based on a state-of-the-art and state-of-practice review of slot allocation instruments, a number of distinct and integrated airport slot allocation strategies were defined [6]. The identification of strategies was followed by the determination of the different airport clusters through the development of a typology of airports. The development of the airport typology was performed through cluster analysis. The selection of the candidate variables for the cluster analysis was based on the four-level framework introduced by Williamson in the theory of institutional economics [7] and the application of this framework for seaports with some parallels drawn for airports [8].

According to this four-dimensional scheme, the following categories of characteristics and attributes for the development of the airport typology were identified: i) cultural / behavioral characteristics, ii) regulatory / legal characteristics, iii) institutional / political characteristics, iv) market characteristics distinguished between traffic, market access, and competition attributes, and v) capacity (supply-side) attributes, which represent a new category / level in the four-layer scheme introduced by Williamson [7]. This latter category (i.e., capacity) has been considered and examined as an additional, separate level of analysis on the grounds that capacity is a major determinant of airports (and transport systems in general). Besides, capacity (i.e., supply) in conjunction with traffic constitute the primary root of the congestion problem that eventually dictates the need for a more efficient slot allocation. As a result, the following types of variables were selected for airport clustering purposes in the context of slot allocation: i) capacity metrics (i.e., declared capacity), ii) traffic / congestion metrics (i.e., aircraft and passenger movements, average delay per movement), and iii) measures of effectiveness of slot allocation and utilization including competition / entry barriers (i.e., misused slots, unsatisfied demand, slot mobility). Finally, one additional variable (i.e., grandfathered over total number of slots) has been included in order to provide a measure of the competition and the ease of market entry for newcomers.

The selection of the clustering variables was thereafter followed by the determination of the most appropriate cluster analysis technique and the definition of the sample of airports (i.e., clustering entities / objects) participating in the cluster analysis. Concerning the cluster analysis technique, several techniques can be found in the literature with the most popular being the hierarchical agglomerative, the iterative partitioning, and the factor analytic. The hierarchical agglomerative clustering was applied for the purposes of the cluster analysis under consideration due to a number of characteristics, advantages, and properties exhibited by these methods [9]: i) they are the most frequently used in such problems and conceptually simple to implement, ii) they produce tight and non-overlapping clusters of cases, iii) they can be efficiently applied for classifying a reasonable number (e.g., less than 400) of cases (i.e., airports) into groups that are homogeneous within themselves and heterogeneous between each other, and iv) they do not necessitate the a priori determination of the number of resulted clusters (as in the case of iterative partitioning methods. The development of the airport typology through the cluster analysis was concluded by the diversification check of the resulted airport clusters.

The final step of this analysis was signified by the assessment of the policy compatibility. The objective of the policy compatibility analysis was to assess the compatibility of the proposed slot allocation strategies in the various airport clusters with respect to policy criteria (and their relative importance weights in each airport cluster). The policy criteria involved in the selection of the strategy in each airport cluster have been elicited from the literature and other relevant policy studies [3], [5]. In order to accommodate the multi-criteria evaluation problem requirements, a multi-criteria assessment technique should be considered. Among the available techniques, the most appropriate technique for the problem at hand was found [10] to be the Analytical Hierarchy Process (AHP) developed by Saaty [11].

III. Overview Of Strategies

Based on the state-of-the-art and state-of-practice review of slot allocation instruments [1]-[6], a number of distinct slot allocation strategies were identified. Table I sketches the identified strategies with their key features, rules, and instruments deployed for the operationalisation of each strategy.

In particular, the “Enhanced Status Quo” strategy involves the minimum departure from the existing system on the grounds that it fully maintains the overriding principle of historic slot holdings based on grandfather rights. It retains the rationale of administrative coordination of slot allocation in conjunction with primary slot trading (coordinated trading). Coordinated trading stands for the establishment of a trading environment that will be centrally administered by a mutually trusted allocation entity. In coordinated trading, the slot pool will be subject to a slot trading process in the form of an informal, one-by-one auction with direct negotiations of the airlines with the central allocation entity.

TABLE I
Alternative Slot Allocation Strategies and their Integrated Instruments and Rules

	STRATEGIES					
	Strategies & Components	Enhanced Status Quo	Gradual	Controlled Trading	Congestion Pricing	Big Bang
INSTRUMENTS & RULES	*Grandfather rights*	✓	✓	✓	Removed	Removed
	Centralized trading with policy criteria	✓				
	Primary trading			✓		
	Secondary trading		✓	✓		✓
	Auctions		✓			✓
	Congestion Fees				✓	
	Recycling	✓	✓	✓		✓
	Use it or lose it rule	✓	✓			
	Policy-designated slots	✓	✓	✓	✓	✓
	All slots?	(Only pool)	(Only pool)	✓	✓	✓

The “Gradual” strategy also involves a conservative approach albeit with a more clear orientation to market mechanisms and a slightly more drastic revision of the status quo especially with regards to secondary allocation. In principle, it also retains grandfather rights, but it attempts an application of market-oriented mechanisms in two parallel directions. Besides grandfather rights and policy slots, the remaining slots will be auctioned at the airport level with monetary trading between airlines being also allowed on a secondary level. Grandfather rights will be still valid and recognized as historic slot holdings but they will not be tradable on a monetary basis. The non-grandfathered and not policy-designated slots will be subject to an auction-based allocation accompanied by secondary trading. At the primary allocation level, the pool slots will be allocated by means of decentralized auctions organized on an airport-specific basis by a mutually trusted entity. At the secondary level, monetary trading of auctioned slots will be allowed through bilateral airlines’ negotiations.

The “Controlled Trading” strategy essentially combines conservative and innovative elements in one strategy. In particular, it retains with slight modifications / adaptations the principle of grandfather rights, but simultaneously allows full (primary and secondary) monetary trading based on bilateral negotiations either between the airport and airlines (primary trading) or between airlines (secondary trading). The characterization “controlled” trading stems from the principle that although full (i.e., primary and secondary) trading is allowed, primary allocation is self-controlled by the historic slot holdings, which could be also subject to monetary trading. Grandfather rights will be temporarily revoked so that a temporary pool of slots will be built with: i) slots that will be (eventually) retained by their owners during the previous scheduling season (x%), ii) slots that will be redistributed on the basis of performance criteria among existing carriers operating in this airport (y%), and iii) slots that will be recycled and returned to the pool (z%). Slots that are not grandfathered and not allocated based on policy criteria will be subject to primary and secondary trading. At the primary trading level, the pool slots will be owned by the airport and will be traded to airlines through a centralized trading environment based on one-to-one negotiations with airlines. At a secondary level, airlines having acquired some particular slots through primary trading will have the option to secondarily trade these slots.

The “Congestion Pricing” strategy represents the most direct pricing method for addressing the real causes of the mismatch between capacity and demand for airport operations. Under the congestion pricing strategy, grandfather rights will be abandoned and a congestion-based scheme with fees varying with congestion throughout the day will be set by an administrative authority. In particular, each carrier could operate at any time or slot by paying the corresponding scarcity rent (i.e., congestion fee). The congestion pricing strategy can be structured to include different components / types of fees addressing various policy or operational objectives. Some of the most widely discussed [3] pricing components will be integrated to produce a three-part tariff that will combine: i) the traditional weight-based fees and passenger surcharges, ii) a flat reservation fee applied per movement in the form of membership dues or “no-show” penalties, and iii) a congestion-based fee. The congestion-based fee will be the newly introduced scarcity surcharge applied during congested period(s) per movement and will vary within the day based on congestion levels and patterns. More specifically, it is proposed to apply a flat fee (only during congested periods) that will be entirely independent of the aircraft weight, but will vary with congestion throughout the day.

The “Big Bang” strategy represents the opposite extreme vis-à-vis the “Enhanced Status Quo” and the “Gradual” strategy on the continuum of the proposed strategies. Grandfather rights will be abandoned with the entire slot pool being allocated by means of market-based instruments (i.e., decentralized auctions accompanied by secondary trading). More specifically, the size of the pool will be administratively declared by the slot coordinator. Grandfather rights will be removed so that the entire slot portfolio will be directly available for (re)allocation. This removal will be thorough (abolition of all grandfather rights) and immediate (not gradual) [3]. A portion of the pool slots will be set aside for policy-driven slot allocation purposes. The remaining (not policy-designated) slots will be thereafter subject to decentralized airport auctions. These auctions will be conducted with a standard frequency (every 3-5 years) and will be coordinated by a mutually trusted authority. Auctioned slots will be exempted from landing fees and their duration will depend on the standard auction frequency (3 or 5 years). In a secondary level, auctioned slots will be tradable between airlines on a monetary basis.

A more elaborated description of the strategies and their functional principles and integrated instruments can be found in [6].

IV. AIRPORT CLUSTERS

This section presents the results of the cluster analysis and the elicited airport typologies. The sample of airports acting as clustering entities included European Union airports designated (on a year-round basis) both as "Level 3 (Schedule Coordination Request – SCR)" according to IATA designation and "Fully Coordinated" as envisaged by European Commission designation guidelines in 2002-2003. The selection of the particular sample of airports stems from the fact that these are the largest and practically the busiest airports playing a strategic role in European air transport, thus attracting the major attention and policy acts from aviation policy makers. Therefore, it can be safely claimed that these "Level 3" and "Fully Coordinated" European Union airports (mainly Category 1) constitute the appropriate sample of airports where some form of demand management through slot allocation is or will be soon necessary as a result of the experienced or anticipated congestion problems. In order to develop the airport typology, the following methodological aspects and cluster analysis properties have been established (Table II).

TABLE II
CLUSTER ANALYSIS PROPERTIES

Cluster Analysis Properties & Methodological Decisions	
Cluster analysis method (Hierarchical Vs. K-means)	Hierarchical method (i.e., the resulted classification has an increasing number of nested clusters and the result resembles a phylogenetic classification). It is the most widely used technique, fits better to the specific conceptual problem, it is simple, fast, and reliable, and does not necessitate the *a priori* determination of the number of clusters (required by K-means clustering technique).
Agglomerative Vs. Divisive	Agglomerative method: all cases / airports are initially considered as single-member clusters which are gradually fused to form larger, but still homogeneous clusters.
Monothetic Vs. Polythetic	Polythetic agglomerative (i.e., based on multiple clustering variables).
Type of clustering variables / data	Interval data (number of passengers, number of aircraft movements, declared capacity in hourly aircraft movements, % of slots utilized or grandfathered).
Clustering objects (cases Vs. variables)	Cases (i.e., airports)
Clustering sample	"Fully Coordinated" and "Schedule Coordination Request – (SCR)" EU airports
Distance Vs. Similarity Measures	Distance measures (due to the type of selected variables, i.e., interval).
Type of distance measures	Squared Euclidean distance
Clustering algorithm	Complete Linkage ("Maximum or Furthest Neighbor Method") algorithm: it is clearly appropriate for Euclidean distance metrics and normally produces very tight clusters of similar cases, which conceptually fits with the problem at hand.
Software used	The (Hierarchical) Cluster Analysis module of SPSS (version 12.0).

Based on the aforementioned properties, four clusters (highlighted in different colors) emerged from the clustering sample of airports on the basis of the selected clustering variables (Table III).

TABLE III
AIRPORT CLUSTERS

Airport	IATA Code	Cluster	Airport	IATA Code	Cluster
Helsinki / Vantaa	HEL	4	Brussels	BRU	2
Berlin / Schonefeld	SXF	4	Copenhagen	CPH	2
Berlin / Tempelhof	THF	4	Dusseldorf	DUS	2
Berlin / Tegel	TXL	4	Munich	MUC	2
Dublin	DUB	4	Stuttgart	STR	2
Turin	TRN	4	Milan / Malpensa	MXP	2
Milan / Linate	LIN	4	Rome / Fiumicino	FCO	2
Milan / Bergamo	BGY	4	Barcelona	BCN	2
Venice	VCE	4	Madrid / Barajas	MAD	2
Florence	FLR	4	Arlanda	ARN	2
Naples	NAP	4	Paris / Charles de Gaulle	CDG	1
Rome / Ciampino	CIA	4	Paris / Orly	ORY	1
Palermo	PMO	4	Frankfurt	FRA	1
London / Stansted	STN	4	Amsterdam / Schiphol	AMS	1
Manchester	MAN	4	London / Heathrow	LHR	1
Vienna	VIE	3	London / Gatwick	LGW	1
Athens	ATH	3	Lyon	LYS	Unclassified
Thessaloniki	SKG	3	Catania	CTA	Unclassified
Lisbon	LIS	3	Faro	FAO	Unclassified
Porto	OPO	3	Funchal	FNC	Unclassified
Alicante	ALC	3	Bilbao	BIO	Unclassified
Fuerteventura	FUE	3	Tenerife Norte	TFN	Unclassified
Gran Canaria	LPA	3	Bromma	BMA	Unclassified
Lanzarote	ACE	3	Eindhoven	EIN	Unclassified
Malaga	AGP	3	Rotterdam	RTM	Unclassified
Palma de Mallorca	PMI	3			
Tenerife Sur	TFS	3			

Cluster 4-"Small National Spokes"

The airports included in this cluster are mainly small, satellite or regional airports acting as the spokes of their national airport network, respectively. Typical examples of airports included in Cluster 4 are Berlin Tempelhof, Berlin Tegel, Dublin, Turin, Milan Bergamo, Milan Linate, Venice, London Stansted, and Manchester.

Cluster 3-"Large National Spokes & Small National Hubs"

This cluster contains small and medium-sized airports acting mostly as larger (as compared to Cluster 4) spokes of the national airport network or small national hubs channeling traffic from the national spokes to international hubs and vice versa. Typical examples of airports included in Cluster 3 are Malaga, Thessaloniki, Palma de Mallorca, and Porto ("large national spokes"), as well as Vienna, Athens, and Lisbon ("small national hubs").

Cluster 2-"Large International Hubs"

Cluster 2 contains major, metropolitan airports of the European airport network acting mostly as large international hubs (at least for certain national carriers) with focus on intra-European routes and a growth potential to establish one of the major European hubs included in Cluster 1. Practically, the airports included in Cluster 2 are primary and secondary large hubs of some major European airlines, which operate these airports as servers of traffic both between international destinations, as well as between domestic and international destinations. Typical examples of airports included in Cluster 2 are Munich, Brussels, Malpensa, Madrid, and Barcelona.

Cluster 1-"Super Hubs"

Cluster 1 airports represent the largest, busiest, and most congested European airports with a worldwide presence and a strategic role in the European airport network. The airports included in Cluster 1 are the primary hubs of the major European airlines (i.e., Lufthansa, Air France / KLM, British Airways) operating these airports as the major intra and extra-EU hubs by accommodating traffic mostly between international airport destinations. Paris / Charles de Gaulle, Frankfurt, Amsterdam / Schiphol, and London Heathrow are typical examples of airports included in Cluster 1.

Based on the technical validation of the cluster analysis results, the following observations can be made:

- The resulted clusters are very compact, distinct, non-overlapping, and of approximately equal number of homogeneous cases (i.e., airports). These are quite dissimilar to each other with the exception of Clusters 3 and 4 representing small national hubs and secondary airports, respectively.
- Two additional small clusters have been found; one containing BIO, TFN, and LYS, which assimilates the membership of Cluster 2, with another containing FNC, BMA, CTA, EIN, FAO, and RTM exhibiting similar characteristics with the membership of Clusters 3 and 4. However, both clusters are rather loose, while their inclusion in one of the tighter clusters would reduce the homogeneity and internal similarity of the other clusters rather than contributing to the interpretation of their behavior. As a result, these two loose clusters have not been considered as autonomous clusters, but rather as single-member clusters for which further analysis is necessary (with the use of additional variables).

V. POLICY COMPATIBILITY ANALYSIS

The multi-criteria evaluation problem at hand has been confronted with the use of the Analytical Hierarchy Process (AHP) [11]. AHP has some notable advantages as a multi-criteria assessment method and fulfils the following methodological properties that are absolutely in alignment with the characteristics and methodological requirements of the evaluation problem under consideration [10], [11]. In particular, the selected evaluation technique:

- Considers multiple (even conflicting) criteria, priorities, and trade-off's (e.g., different priorities of policy objectives in the different airport clusters).
- Expresses and quantifies the relative importance of the various criteria and indicators in non-homogeneous panels of experts / judges.
- Enables multiple judgments and group decision making but does not insist on consensus; instead, it synthesizes the outcome of diverse judgments.
- Compiles the assessments and expert judgments of various decision makers / experts and identifies "compromising" solutions. Eventually, it leads to an overall estimate of the desirability - with respect to performance, compatibility, importance etc. – of each alternative (e.g., the most compatible strategy for each

airport typology based on the different priorities of the policy criteria and the relative compatibility of strategies in each airport typology).

- Tracks the logical consistency of judgments used in determining priorities and provides an easy quantitative way in order to improve logical consistency.
- Enables sensitivity analysis.
- Enables the use of intangible, qualitative criteria by providing a scale for measuring and establishing their priorities.
- Provides a hierarchical structuring of the evaluation problem (Fig. 2) by decomposing it from the “Ultimate Evaluation Goal” (i.e., compatibility assessment) to the “Evaluation Criteria” (i.e., policy criteria), “Evaluation Indicators” (i.e., policy indicators), and “Evaluation Alternatives” (i.e., slot allocation strategies) that are applicable to each airport cluster (i.e., typology) in the context of the herewith examined problem.

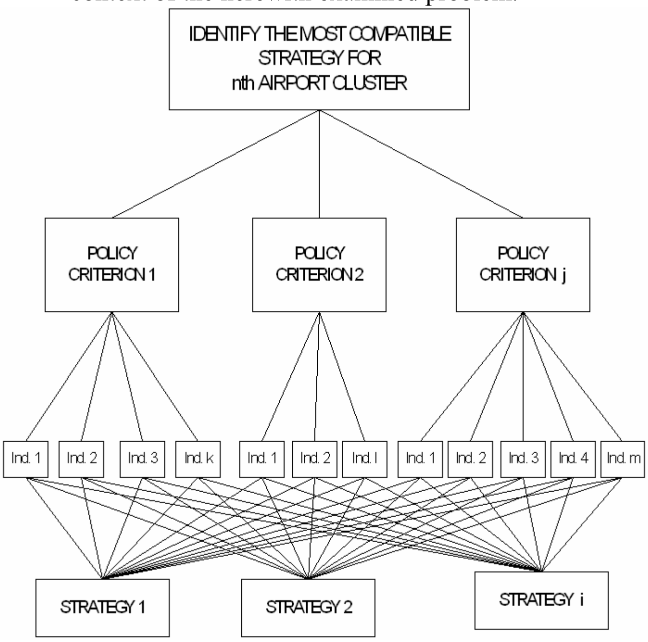


Fig. 2. Hierarchical structure of the policy compatibility assessment.

Based on the above presented AHP model, the assessment levels are structured as follows:

- Level 0 (“Ultimate Evaluation Goal”): it represents the assessment goal, that is, the policy compatibility assessment.
- Level 1 (“Evaluation Criteria”): it assesses the priorities / weights of the policy criteria. The policy criteria (Table IV) are further broken down to policy indicators.
- Level 2 (“Evaluation Indicators”): it assesses the priorities / weights of the policy indicators (Table IV). This level provides the indicators based on which the strategies are comparatively assessed (Level 3).
- Level 3 (“Evaluation Alternatives”): it represents the lowest level of assessment in which the strategies are comparatively assessed. This level’s assessments eventually determine the overall ranking of strategies in each airport cluster.

TABLE IV
POLICY COMPATIBILITY CRITERIA & INDICATORS

Criterion & Indicators	Definition
Efficiency: *Allocative Efficiency*	It measures the capability of the strategy to allocate slots to those with the greatest willingness to pay.
Efficiency: *Competitive Efficiency*	It measures the strategy’s capability of promoting competition through the elimination of entry barriers to newcomers and discriminatory practices in favor of established carriers.
Efficiency: *Infrastructural Efficiency*	It measures the extent to which the slot allocation proceeds are efficiently distributed to those having the means at their disposal to eliminate the scarcity.
Cost: *Cost-relatedness*	It measures the extent to which airport charges are representative of actual costs.
Cost: *Implementation Cost*	It measures the costs (e.g., transaction, organization / coordination, technology) envisaged for the implementation of the strategies.
Implementation: *Complexity*	It measures the degree of implementation complexity of a strategy in the form of administration, preparatory time required, as well as the necessary organizational arrangements and coordination efforts.
Implementation: *Flexibility*	It measures the degree of flexibility of strategy’s implementation in terms of temporal or spatial flexibility.
Implementation: *Fiducial Protection*	It measures the implementation phasing / graduality of a strategy.
Acceptability: *Stakeholders’ Inertia*	It measures the degree of expected resistance of the affected stakeholders (e.g., carriers, airports) to the changes introduced by the various strategies.

VI. CONCLUSIONS

This paper presents a systematic methodological framework for developing and evaluating strategies in the airport slot allocation context. The ultimate objective of the presented work was to perform a policy compatibility analysis aiming to assess in a structured and scientifically solid manner the most appropriate slot allocation strategy for each airport cluster based on its profile characteristics and the embedded decision making and policy priorities.

Currently, a targeted panel of experts and key stakeholders has been selected and is surveyed with the use of a detailed survey instrument capturing the methodological aspects and requirements of the multi-criteria assessment problem and the associated AHP operational properties. This expert panel includes field experts mainly from policy makers, collective industry bodies, airport operators, airlines, and academic / research experts in slot allocation. The policy compatibility analysis is expected to provide answers on the following: i) what is the most appropriate slot allocation strategy in each airport cluster, ii) how the appropriateness and compatibility of each strategy changes with different importance weights assigned to the various policy criteria and their associated indicators (sensitivity analysis), and iii) which are the different viewpoints and assessments of the various interest groups involved in the survey and why (if any) do they vary with their industry roles and business goals.

REFERENCES

[1] DotEcon Ltd., "Auctioning airport slots", Report for HM Treasury and the Department of the Environment, Transport, and the Regions, DotEcon Ltd. London, U.K., 2001.

[2] Federal Aviation Administration (FAA), "Notice of alternative policy options for managing capacity at LaGuardia airport and proposed extension of the lottery allocation", Federal Register, 66 (113), June, 31731-31748; Docket FAA-2001-9852, 2001.

[3] Technology University of Berlin (TUB), "Possibilities for the better use of airport slots in Germany and the E.U.", Technical report prepared by the Technical University of Berlin (TUB), Workgroup for Infrastructure Policy, Berlin, Germany, 2001.

[4] K. G. Zografos and M. A. Madas, "A critical assessment of airport demand management strategies in Europe and U.S.: a comparative assessment", *Transportation Research Record*, No. 1850, pp. 41-48, Transportation Research Board, Washington, D.C., 2003.

[5] National Economic Research Associates (NERA), "Study to assess the effects of different slot allocation schemes", Technical Report prepared for the European Commission, London, U.K., 2004.

[6] M. A. Madas and K. G. Zografos, "airport slot allocation: from instruments to strategies", *Journal of Air Transport Management*, Vol. 12, Issue 2, pp. 53-62, 2006.

[7] O. E. Williamson, "The new institutional economics: taking stock, looking ahead", *Journal of Economic Literature*, Vol. 38, No. 3, pp. 595-613, 2000.

[8] CONNEKT, "Institutional arrangements in seaports: an institutional economic analysis of seaports", CONNEKT Research Paper, CONNEKT, EIM @ YourService BV, Delft, The Netherlands, 2003.

[9] M. S. Aldenderfer and R. K. Blashfield, "Cluster analysis", Series: Quantitative Applications in the Social Sciences, Series / Number 07-044, SAGE Publications, Beverly Hills, California, 1984.

[10] K. G. Zografos and I. M. Giannouli, "Development and application of a methodological framework for evaluating the cost-effectiveness of an air traffic management system for Europe", *Transportation Research Record*, No 1744, pp. 52-64, Transportation Research Board, Washington, D.C., 2001.

[11] T. L. Saaty, "Decision making for leaders: the Analytic Hierarchy Process for decisions in a complex world", RWS Publications, Pittsburgh, PA., 1990.

Michael A. Madas holds a bachelor degree in Marketing and Operational Research of the Athens University of Economics and Business, and a M.Sc. Degree in Decision Sciences (Logistics and Transport Specialization) of the Athens University of Economics and Business. He is currently a Ph.D. candidate in the Department of Management Science and Technology of the Athens University of Economics and Business in the area of "Management of Airport / Air Transport Systems". He has substantial expertise in airport terminal operations, airport demand management and slot allocation, aviation demand forecasting, and transport pricing.

Konstantinos G. Zografos is a Professor at the Department of Management Science and Technology of the Athens University of Economics and Business (AUEB), Director of the Transportation Systems and Logistics Laboratory (TRANSLOG), and Coordinator of the Transportation and Logistics Specialization of the Graduate Programme. Professor Zografos has been on the faculty of the University of Miami, Coral Gables, Florida, U.S.A. His professional expertise, research and teaching interests include: Applications of Operations Research and Information Systems in Transportation Systems and Logistics Management with particular emphasis on: Air Traffic Management, Airport Planning and Operations, Fleet Management, Vehicle Routing and Scheduling, Incident Management and Emergency Response Logistics, Transportation of Hazardous Materials, and Risk Assessment.

He has published more than 90 papers and he is a member of the TRB Committee on Airfield and Airspace Capacity and Delay (AV060), member of TRB Committee on Freight Transportation Planning and Logistics (A1B02), and member of TRB Committee on Transportation Network Modeling (A1C05).

Professor Zografos has been awarded for his doctoral studies the ENO foundation for transportation fellowship. He has been one of the six finalists of the Competition for the 2005 Franz Edelman Award for Achievement in Management Science and Operations Research and he received the 2005 President's Medal I of the British Operations Research Society.

Adopting the Agile Unified Process for Developing a DSS for Airport Strategic Planning

Roland Wijnen
Faculty of Aerospace Engineering
Delft University of Technology
2629 HS, Delft, The Netherlands
Email: r.a.a.wijnen@tudelft.nl

***Abstract*— For efficient business, an airport must develop strategies for its development. The increased complexity of an airport system and an uncertain future make this increasingly difficult. A decision support system might be able to help. However, developing such a system itself is not easy. Successfully developing a DSS, requires an agile development process and knowledge from both the aeronautical as well as the policy analysis field. Our research has shown that the policy analysis approach is an effective framework for the design of the DSS. The policy analysis approach defines the scope for the analytical activities from the aeronautical field. Together with a domain driven design, a system can be designed and implemented that captures the best of both worlds to support airport decision makers in strategic planning.**

I. Introduction

The world in which airports operate is undergoing rapid change. Traditional approaches to planning (e.g., Master Planning) are becoming less appropriate. The stakeholders associated with airports are changing as well. The main new stakeholder, especially for privatized airports, is the financial market, which sets new requirements on financial performance. Furthermore, trends of increasing privatization and liberalization put time-pressures on strategic decision making: opportunities have to be seized and threats dealt with quicker than ever before.

This increasingly complex and dynamic set of circumstances motivates the need for a Decision Support System (DSS) that offers systematic problem analysis and supports multiple stakeholders. Traditionally, DSS's have been designed to support specific decisions or portions of the decision making process. They have rarely been designed around a complete, comprehensive analytical framework for decision making. Here, a DSS for airport strategic planning is presented that is based on the policy analysis approach to problem solving. It is a systematic approach that is generic enough for addressing the wide range of airport planning problems and is widely accepted for analyzing a diversity of problems.

The increased efficiency provided by this DSS will free the advisers of decision makers from time spent on configuring the process (hiring experts, selecting tools, getting information & data) every time a new strategic plan needs to be created or an existing one needs to be updated. Their valuable time can be used, instead, to be more creative when searching for policies that address the problems.

II. Current Strategic Planning

With respect to the current strategic planning process, only the roles of the people involved are described in some detail. The information and process, although very important for understanding what the DSS should provide in terms of automation and functionality, are only touched upon. A discussion of the problems related to the current approach concludes this section.

Three major roles (performed by one or multiple persons) that can be distinguished are the Decision Maker, Decision Adviser and Domain Expert. The decision maker is the person that makes the actual decisions in order to meet the business objectives that are associated with a vision for the airport's development. The decision adviser is a person that advises the people that make the actual decisions. The decision advisers explore the policies that could be implemented for meeting the objectives set by the decision makers. In order to accomplish this task, they hire external experts, use computer tools, and consult in-house domain experts. The domain expert is a person that has specific knowledge of the airport system and its operation. The domain expert uses various tools to provide quantitative information about a specific aspect (e.g. capacity and delay, environmental impacts) of the (future) airport performance.

Part of the strategic planning process is mobilizing resources such as internal staff, consultants, and experts representing other stakeholders for help with the problem analysis. In addition, a lot of information and data need to be collected or generated as input for the analysis. Conducting the quantitative part of the analysis involves the selection of the appropriate tools and processing their results. The inputs and the results need to be consistent and valid, which usually requires a significant effort. If the decision maker is satisfied with the results, a concrete strategic plan can be created. The plan is communicated to the other stakeholders, which usually initiates a discussion about the plan itself and the underlying analysis process, possibly requiring additional analysis. After reaching agreement among the stakeholders, the final strategic plan can be issued. And these activities will have to be repeated when a new problem has to be analyzed at some time in the future.

A number of problems is associated with the current process. One problem is the use of a limited number of scenarios for modeling possible future situations. Two other problems are the design of non-optimal policies (it is not clear whether the policies of the strategic plan are truly effective in addressing specific issues of the decision-making problem) and the consideration of a non-exhaustive list of policies (policies are arbitrarily left out of the analysis instead of using a quantitative screening). Another problem is the inefficient use of resources because a number of parties is involved in conducting the analysis. These parties might use different assumptions and data, which makes it difficult to combine results afterward. An integral view on the future performance of the airport system can only be provided by manually collecting, combining and post-processing the results. Stakeholders might also argue about results, assumptions, and the methodologies that were used. Finally, there are difficulties in revising the plan. While executing the previously described process, new issues might have come to the surface which makes the proposed plan out of date already. Because of that, the plan should be updated accordingly but usually there are not much resources left to do so.

The Solution

The strategic planning process could be improved by partially automating the activities conducted within the current process. Only automation however is not sufficient. The strategic planning process can only be improved if all the people, information and tools are connected appropriately through a Decision Support System (DSS).

Developing such a software system is a complex task and therefore the adoption of a well-defined software development process is required. For this project, the Agile Unified Process (AUP) is adopted. The Agile Unified Process[1] (AUP) is a simplified version of the Rational Unified Process[2] (RUP). The reason for using the AUP is because it is an agile and iterative approach to software development. An adaptive evolutionary development approach has been recognized as crucial by the DSS development community since the early 1980's [1]. The disciplines (like modeling, implementation, tetsing) are performed in an iterative manner, defining the activities which development team members perform to build, validate, and deliver working software which meets the needs of their stakeholders. During an iterative process incremental releases are delivered over time, instead of the 'big bang' approach where software is delivered all at once. More information about the AUP and its philosophies can be found in [2]. The remaining sections of this paper describe the results of the current development process of this DSS, which mainly involves modeling and implementing focusing on the architecture of the system.

[1]This is a summary from `http://www.ambysoft.com/unifiedprocess/agileUP.html`.

[2]Rational Unified Process is a trademark or registered trademark of Rational Software Corporation in the United States and in other countries.

III. The Vision

The vision is an (AUP) artifact that states the problem to be addressed, identifies who the stakeholders and users are and what high-level functionality needs to be provided.

A. Problem Statement

Section II described the current strategic planning process in general. The problems associated with the current approach were also identified. Summarizing, the current process consumes a significant amount of resources for collection of information and data, selecting and running the appropriate tools for conducting the quantitative part of the analysis, and processing the results. Usually, a significant effort is required to make sure that the inputs and the results are consistent and valid. Finally, the results have to be translated into a concrete strategic plan, which needs to be presented to and discussed with the other stakeholders so that a final plan can be issued. All of these activities have to be repeated when a new problem has to be analyzed at some time in the future.

B. Positioning

There is a need for a DSS that offers systematic problem analysis and that supports multiple stakeholders. Traditionally, DSS have been designed to support specific decisions or portions of the decision making process. They have rarely been designed around a complete, comprehensive analytical framework for decision making. A DSS for airport strategic planning is envisioned that is based on the policy analysis approach to problem solving.

The increased efficiency provided by this DSS will free time spent on configuring the process (hiring experts, selecting tools, getting information & data) every time a new strategic plan needs to be created or an existing one needs to be updated. Valuable time can be used, instead, to be more creative when searching for policies that address the problems. The system is called HARMOS, which is an an acronym for 'Holistic Airport Resource Management and Optimization System'.

C. Stakeholder Descriptions

One of the first activities in the modeling discipline is identification of the stakeholders and users of the system. This section provides a list of the stakeholders and users, their goals, and the key problems that they perceive. It does not describe their specific requirements, as these are captured by the so-called Use Case Model (discussed later in V-A). The stakeholders identified are the airport operator, airlines, ATC authority, the general public, the aviation authorities, governmental and political organizations, banks, and businesses associated with the airport. Note that in the project's Vision artifact the stakeholders is described in much more detail. Every stakeholder could actually also be a user of the system since HARMOS is supposed to be a system that is commonly used. However, only certain types of people from the stakeholder community will actually work with the system. These people are identified in the next section.

1) User Summary : In total, eight actors have been identified, i.e. three primary, three secondary and two off-stage actors.

The primary actors are those actors that have user goals fulfilled through using services of the system. They need to identified because through them the user goals, which drive the use cases (section V-A), are found. The domain expert is a person that has specific knowledge of (a part of) the airport system and/or operational aspects. This actor will interact heavily with the system since experts are needed for populating the system with data, interpreting and validating the various results generated by the system. The decision adviser is a person that advises the people that make the actual decisions. They need to be able to test different policies and strategies (i.e. combinations of policies) for developing the airport for different futures. They also need to be able to clearly present them to the decision makers so that the decision maker can make a good comparison and create a robust strategic plan. The Application Manager is an IT expert with specific installation and maintenance experience (including database management) of the system.

A supporting actor provides a service to the system. This is often a computer system but could also be an organization. The supporting actors identified for this system are the development team, consultancy firms, and third party tools. The latter are the software applications used for assessing the performance of a specific part of the airport system.

An offstage actor has an interest in the behavior of the system but is not primary or supporting. They need to be identified because it needs to be ensured that all necessary interests are identified and satisfied. The decision maker is the person that makes the actual decisions. For the decision maker it is important that the system can be used efficiently by his advisers and experts such that it provides him with effective solutions for decision making problems. Another actor are the organizations that are affected by the airport development and planning, since these organizations would like their objectives to be taken into account during the policy analysis to ensure the resulting strategic plan is acceptable for them too.

The previous list also illustrates how the roles of the people involved in the strategic planning process (section II) map to the actors for the software system.

2) High-level Goals and Problems: The key goals of the stakeholders that were identified are:

- A more transparent analysis and planning process: Increased transparency of the analysis process is important to enable efficient communication with other stakeholders. A transparent plan is important because the decision maker will have to communicate it to the other stakeholders, including the general public and politicians.
- More efficient analysis: The system should enable an analysis process that can be deployed efficiently within organizations, using its resources to their full potential;
- Robust policies: A strategic plan is composed of a mix of policies to be implemented at various times during the planning period that is considered. This plan should be composed of policies that perform well over a range of external factors (i.e. are robust);
- Faster consensus or compromise: The world is changing fast and so is the context for a particular decision making problem. Is it therefore crucial that the different stakeholders reach consensus early or find a compromise solution within a reasonable time;
- Support for adaptive planning: Even if early consensus or compromise is accomplished, the initial plan could run out of date because of changing external factors not thought of at the time the analysis started. The system should therefore support revision of the initial plan easily, without requiring the need to redo everything again every time.

The key problems perceived by the stakeholders have already been discussed in section II.

3) User-Level Goals: The users need a system to fulfill the goals presented in Table I. The goals are directly associated with a use case (abbreviated with UC). A use case is a textual description of the interaction between a system and its users serving as a contract for system behavior.

TABLE I
ACTOR-GOAL LIST

Actor	**Goal**
Decision Adviser	Define Decision Making Context (UC3)
	Create Scenario (UC4)
	Explore Business as Usual (UC5)
	Design Policy (UC6)
	Explore Strategies (UC7)
	Evaluate Results (UC8)
	Compose Strategic Plan (UC10)
	Disseminate Strategic Plan (UC11)
	Update Strategic Plan (UC12)
Domain Expert	Validate the System (UC1)
	Specify System Model Characteristics (UC2)
	Create Scenario (UC4)
	Explore Business as Usual (UC5)
	Design Policies (UC6)
	Evaluate Results (UC8)
	Execute Performance Analysis (UC15)

D. Key Principles of the DSS

The previous sections showed there is a need for a new approach toward decision support for strategic planning. A DSS for strategic planning should therefore adequately address the problems and support the high-level goals of the stakeholders. The key principles used to realize this system are:

- Policy Analysis Approach: The current process used to develop strategic plans is mainly based on an ad-hoc approach, which is not very efficient in terms of the deployment of resources and consistency of information. A lot of coordination and organization is required, leaving less time available for being creative in exploring effective solutions for the decision-making problem. These issues can be addressed by adopting a systematic and well-defined approach for analyzing decision-making problems. The policy analysis approach is such an approach;

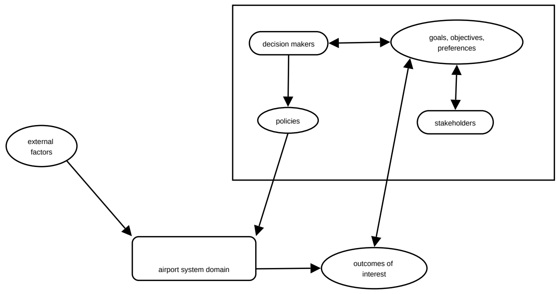


Fig. 1. Framework for policy analysis

- Integrative Design: The entire design of the DSS is integrative. The main aspects that are directed toward integration are:
  - People: The whole idea of HARMOS is to get all stakeholders actually involved in the strategic planning process from the beginning, working together on the solutions for the decision-making problems, trying to understand each other's perspective. Besides this, the people involved in strategic planning need to be properly serviced by the system, which is the reason why the Decision Maker, Decision Adviser and Domain Expert were all identified as an actor of the system;
  - Information: Consistent and valid data is crucial for providing quantitative information about the airport performance. Also the assumptions and input that are used to provide the information is to be available at all times;
  - Tools: The tools used to perform quantitative analysis for the airport's performance are controlled (input, execution, and output) by the system to provide those results that are relevant in the decision making context.
- Optimization: Mathematical optimization techniques can be used for designing policies (UC6). If a policy has a quantifiable objective, optimization is perfectly suited for designing the policy itself.

The policy analysis approach is discussed in more detail in the next section because it is the principle that defines the scope for the system's design. A more in-depth discussion of the policy analysis approach can be found in [3]. The other key principles are not discussed separately.

IV. Policy Analysis Approach

Policy analysis is aimed at facilitating the policymaking process by producing policy relevant information. The policy analysis approach has been described extensively already by [4] among many others. Here, a brief summary of the policy analysis approach is given. The framework that is used for policy analysis is presented in Figure 1.

The figure shows the main concepts and what their relationships are to the system under analysis and to the other concepts. The system domain is the (mathematical) representation of the system (here the airport, its environment and the relevant systems it interacts with) that is subject to policy analysis. The external factors are the forces that influence the system but that cannot be influenced by the decision makers. Policies are the forces used by the decision makers to control the system. Stakeholders are the entities that have an interest in the operation of the system and its effects. Policy or decision makers are the people that define and implement the actual policies. Objectives, Goals and Preferences are the decision maker's 'standards' and attitude toward meeting those standards. The outcomes of interest are the system outcomes related to the policy objectives.

For sake of brevity, the analytical disciplines within the policy analysis can not be be presented here extensively. Only the analytical disciplines that play a role in the use cases (Table I) that have been realized during the recent development effort are discussed.

Like shown in Figure 1, at the heart of the system description is a system model that represents the policy domain. The 'descriptive modeling and analysis' discipline is about providing an integral system description of a policy field. Analysis of decision making problems in the context of strategic planning requires a model representing the airport itself, its environment and other systems it interacts with, at the appropriate level (macro- and/or mesoscopic) of detail. Functionality for specifying the characteristics of this system model is obviously to be provided by HARMOS (UC2: Specify System Model Characteristics).

The 'impact or performance analysis' discipline is about determining the outcomes produced by the system, either with or without policies in effect. Based on the system model and the forces applied to it (external factors and policies), the appropriate tools should be selected for conducting impact analysis (UC15: Execute Performance Analysis). For an airport, the tools required are related to capacity and delay, noise, emissions, third party risk, and finance (including cost-benefit analysis).

Other disciplines that are defined are 'forecast analysis' (UC4), 'design' (UC6), 'evaluation' (UC8) and 'implementation' (UC12). These disciplines are not further described here because the use cases they relate to (indicated between brackets) are not yet considered at the moment.

The policy analysis process generally involves performing the same set of logical steps. The steps are not always performed in the same order, and there is usually feedback among the steps. The first three steps are (1) identification of the problem, (2) identification of the objectives of the new policy, and (3) deciding on criteria (measures of performance and cost) with which to evaluate alternative policies. The functionality covering these three steps is included in UC3: Define Decision Making Context. The next steps are (4) select the alternative policies to be evaluated (also UC3), (5) analyze each alternative policy (UC6: Design Policy), (6) compare the alternatives in terms of projected costs and effects (UC8: Evaluate Results). If none of the alternatives examined so far is good enough to be implemented, one should return to step (4). Step (7) concerns implementation of the chosen alternative(s) (UC11: Disseminate Strategic Plan) and step (8) is about monitoring and evaluating the results, returning to one of the previous steps if necessary (UC12: Update Strategic Plan).

As described, this process exhibits an iterative character. By using optimization (one of the key principles of the DSS) the efficiency of this process can be improved, because optimization techniques can be used to design policies that directly meet the objectives. The application of optimization techniques has already been tested in a HARMOS prototype which was developed during the so-called Inception phase. An optimization module that finds the best policy for allocation of flights to runways and tracks based on multiple objectives (delay, noise and third party risk) is described in [5]. An optimization module that finds the optimal arrival procedure with respect to the number of awakenings is described in [6].

V. HARMOS

The current work on HARMOS requirements, analysis and design (part of the Modeling discipline of the AUP) is described, showing how the key principles presented in section III-D are applied. First, the requirements are described in more detail by briefly discussing a few use cases from the Use Case Model. These use cases only specify what the system is supposed to do; not how it is being done.

Analysis of the use cases results in so-called use case realizations that specify design elements which are implemented so that the functionality described by the use cases is actually delivered. Although a large number of use cases have already been identified (see Table I), not every single one of them is to be realized at the same time. An iterative approach focuses on only a few of them during each development iteration. The current development focus is on the architecturally significant use cases, so that a stable architecture is realized. Section V-A discusses some of those use cases followed by the presentation of the logical view (section V-B) on the architecture.

A. Use Case Model

As mentioned already, a use case model describes the interaction between a system and its users. Writing use cases [7] is an effective way of identifying and specifying the functional requirements of a system. The next subsections present the use cases in a very brief format, only describing their essence; within the project use cases in fully dressed format (2 to 3 pages of text) are used.

1) Validate the System (UC1): Before one can start using the system, it needs to be validated. The development team together with the domain experts specify the system model characteristics (UC2) and execute a performance analysis (UC15). The results from this analysis are compared (UC8) with information about the known performance.

2) Specify System Model Characteristics (UC2): Without a model of the system that is subject to policy analysis, the quantification of airport performance and how it is affected by policies and external factors (the forces on the system) is impossible. The domain experts specify all the characteristics of the airport (like the runway system, terminals, flight operations), its environment (atmospheric conditions, population density, etc.) and other systems (e.g. ATM separation standards). This use case drives the design of the Data Model, which is another artifact.

3) Execute Performance Analysis (UC15): This is a use case at the subfunction level since it does not directly satisfy a user goal. Note that simply executing a particular performance analysis (e.g. a noise impact study) does not satisfy a goal in itself. A goal is only satisfied when such an analysis is done within the decision making's context. The use case is described anyway because it involves executing, and controlling a number of tools. The execution and control of the tools should be independent of the tools actually used (as identified in the vision). The realization of such a feature is pretty complex and of high importance for the business value so it should be addressed during the architectural design.

B. Logical View of the Architecture

The architecture is based on the LAYERS PATTERN [8], having four layers, namely the user interface layer (UI), application layer (containing objects for controlling the flow of the application), the domain layer containing an object-oriented model of the domain and the technical services layer (including lower level services used by the higher layers for getting their job done). Each of the layers is further partitioned in packages, which are logically grouped objects based on a specific domain aspect. The domain and technical services layer and their packages are described in the following sections.

1) Domain: The domain model, an object-oriented model of the domain (i.e. airport strategic planning), is contained in this layer. A domain model is an official RUP artifact that conceptually describes the domain. In this project a domain model, which is an example of a domain logic pattern [9], is used as the central focus of development throughout the entire project. Such an approach is known as *Domain-Driven Design* [10].

The objects in this layer are organized in packages. The definition of those packages should not be made from the developer's point of view (possibly resulting in a structure that nobody else can understand) but should be inspired by the domain. In order to identify these packages, the concepts from the domain need to be discovered and defined consistently (a glossary is therefore another important artifact). The package names (like scenario, policy, system model, outcomes) and responsibilities chosen therefore reflect the concepts of the policy analysis framework (Figure 1). Other packages (like analysis and optimization) have been based on concepts related to airport analysis in general. Only the core, system model and analysis package are described in the next paragraphs.

a) Core: The core package includes basic functionality required for running the application. For now, three classes, i.e. study, context, and case, are defined in this package. First, the responsibility of the case class is explained. Besides the categorization (external factors, policies, system, outcomes of interest, stakeholder objectives) according to the policy analysis framework, the entire set of input data and results generated during the process needs to be organized in a way such that each user can work with the DSS efficiently. Organizing this requires a case management mechanism. A case is a consistent collection of information incorporating input data, assumptions, and results. Four cases are defined: validation, reference, policy and strategic case. The development efforts described here focus on the validation case and therefore only this case is described in more detail. A validation case captures the current airport characteristics and its estimated performance. This information is then compared to e.g. empirical data on real-world performance in order to validate the DSS (UC1). The goal of the development effort described in this paper is to provide all the functionality for validating the DSS. Since it is difficult to describe the resulting software here, a demo of this functionality is planned during the presentation of this paper during the conference. A study is related to a single decision making problem and servers as a container for cases. The context class has a one-to-one relationship and is responsible for keeping the information related to the decision making context (UC3).

b) System Model: This package contains the model of the system that is the subject of the policy analysis. Since HARMOS is a DSS for airport strategic planning, the model incorporates all the elements of an airport, its environment and systems it interacts with, that are important within the long-term planning context (UC2).

c) Analysis: The responsibility of this package is setting up a particular type of airport performance analysis (UC15). It does not have the the responsibility for putting the analysis in the context of the decision making problem. This is a responsibility of the case class from the core package.

2) Technical Services: The functionality in this layer is rather generic and common, like mechanisms for accessing data, persisting objects, and interacting with external tools. At the moment, three packages have been identified, namely persistency (not described any further), data access, and tool wrappers.

a) Data Access: A significant amount of data is required for airport modeling and planning. Some of this information is to be taken from existing sources, like OAG schedules, ICAO/IATA code listing of airlines, aircraft and (destination) airports, engine emission and noise data, etc. This package is responsible for disclosing such existing information.

b) Tool Adapters: Like discussed in the vision, the tools are supporting actors in the form of external systems. Therefore, a mechanism (or tool adapter) needs to be designed that allows HARMOS to interact with these external systems. One of the key features of the DSS is that it should be independent of tools. An airport organization has to be able to select their preferred tools for collaborating with the HARMOS system. For each of the tools that is to be interacting with HARMOS an adapter needs to be provided. By default, adapters for the FAA airfield capacity model, the Integrated Noise Model (INM), and the EPA AERMOD dispersion tool are provided.

For a number of airport performance (emissions, third party risk, finance) aspects, no suitable existing tools have been found. For these aspects, tools will be developed in-house, like a tool for computing airside and landside delays (based on our experience with implementing a delay model for the Airport Business Suite [11]).

VI. Conclusion

This paper identified the need for a DSS for airport strategic planning. It described how such a software system is being developed. Only a limited overview of the development process and associated artifacts was provided. Based on that, the following can be concluded:

- It is very important to have a Vision that directs the developments efforts;
- An agile development process is critical for being able to deal with and learn from changes during the project;
- Both the field of aeronautical engineering and policy analysis are required as a knowledge base for building a DSS for airport decision making. This project again confirms that a DSS cannot be designed successfully if its design is exclusively driven by a technical perspective [12].
- The policy analysis approach can effectively be used as the basis for a DSS;
- A domain driven design helps defining a software system that truly captures the knowledge of the domain exploiting object technology to its full extent.

References

[1] Sprague, R. H. Jr., Carlson, E. D., *Building Effective Decision Support Systems*. Prentice-Hall, London, 1982.

[2] S. W. Ambler, “The Agile Unified Process,” http://www.ambysoft.com/unifiedprocess/agileUP.html.

[3] Walker, W.E. and Wijnen, R.A.A., “Using the Policy Analysis Framework to Design a DSS for Airport Strategic Planning,” in *To be presented at the International Conference on Creativity and Innovation in Decision Making and Decision Support*, June 2006.

[4] W. E. Walker, “Policy Analysis: A Systematic Approach to Supporting Policymaking in the Public Sector,” *Journal of Multicriteria Decision Analysis*, vol. 9, no. 1-3, pp. 11–27, 2000.

[5] Hebly, S.J. and Wijnen, R.A.A., “Development of a Runway Allocation Optimisation Model for Airport Strategic Planning,” in *Proceedings of the 11^{th} Eleventh Australian Aeronautical Conference*, 2005.

[6] van Loo, R.P. and Wijnen, R.A.A., “Integration of NOISHHH into a decision support system for strategic airport planning,” in *Proceedings of the 11^{th} Eleventh Australian Aeronautical Conference*, 2005.

[7] A. Cockburn, *Writing Effective Use Cases*. Addison-Wesley Professional, 2001.

[8] C. Larman, *Applying UML and Patterns*. Prentice Hall, 2005.

[9] M. Fowler, *Patterns of Enterprise Application Architecture*. Addison-Wesley, 2003.

[10] E. Evans, *Domain Driven Design*. Addison-Wesley, 2004.

[11] Warren E. Walker et.al., “The Airport Business Suite: A Decision Support System for Airport Strategic Exploration,” in *Proceedings of the 3^{rd} AIAA Aviation Technology, Integration and Operations Symposium*, Denver, Colorado, U.S.A., Nov. 2003.

[12] Lucas H. C. Jr., *Why Information Systems Fail*. Columbia University press, New York, 1975.

Session 10

Security

Increased Detection Performance in Airport Security Screening Using the X-Ray ORT as Pre-Employment Assessment Tool

Diana Hardmeier, Franziska Hofer, Adrian Schwaninger

***Abstract*—Detecting prohibited items in x-ray images of passenger bags is one of the most important tasks in aviation security. This screening process includes both, knowledge-based and image-based factors. That is, the knowledge about which items are prohibited and what they look like in x-ray images (knowledge-based factors) and the ability to cope with bag complexity, superposition and rotation of the threat item (image-based factors). The X-Ray ORT was developed to measure how well screeners and novices can cope with image-based factors. Schwaninger, Hardmeier, and Hofer (2004) could show that image-based factors are rather independent of knowledge and therefore can only be partly enhanced through training. As these image-based factors are very important in all x-ray screening tasks, using the X-Ray ORT as pre-employment assessment tool should result in a remarkable increase in detection performance of screeners in the future. To test whether the X-Ray ORT is a useful tool to select job applicants, detection performance of screeners selected with and without the X-Ray ORT was compared in the Prohibited Items Test (PIT), which mainly measures knowledge-based factors. This means that one group of job applicants (all novices) was hired using the X-Ray ORT, whereas the other group was hired without the X-Ray ORT. Both groups of screeners had undergone initial classroom training and a minimum of one year working experience in screening carry-on baggage when they took the PIT. Results evidence that in fact detection performance in the PIT is significantly higher for the group selected with the X-Ray ORT than detection performance of screeners selected without the X-ray ORT.**

Furthermore, results reveal reliable and valid measurement of detection performance in both tests, the ORT and the PIT.

***Index Terms*—Aviation Security, Object Recognition, Pre-Employment Assessment, Reliability, Validity.**

I. Introduction

NOWADAYS, civil aviation has become more important and passenger flow still increases yearly. As a result, work load in aviation security increases enormously. To ensure effective and efficient work, it is very important to select and train people accurately. One of the most important tasks of aviation security screeners is detecting prohibited items such as guns, knives, improvised explosive devices (IEDs) and other prohibited items in passenger bags. During rush hours at checkpoints the decision whether a bag is OK (i.e. contains no prohibited item) or NOT OK (contains a prohibited item) has to be made within four seconds. This short time requires both, the profound knowledge about prohibited items and their appearance in x-ray images, as well as the ability to cope with image-based factors such as bag complexity, superposition and rotation of the threat item.

Referring to a general visual cognition model, recognition is defined as a successful matching of the stimulus representation with the visual memory representation. Based on this model [1] revealed two main factors in detecting threat items in x-ray screening, knowledge-based and image-based factors. First, screeners have to know which objects are prohibited and what they look like in x-ray images in order to recognize them (knowledge-based factors). As the appearance of prohibited items in x-ray images can differ remarkably from real life, training is very important in order to recognize them. In addition, it could be shown that an effective training system like X-Ray Tutor can significantly increase detection performance by reducing the false alarm rate. That is, through training screeners learn to distinguish reliably similar looking threat and non-threat items. Second, image-based factors influence detection performance in X-ray images enormously. [1] have shown three different types of image-based factors, namely bag complexity, superposition and rotation of the threat item. A threat item is more difficult to detect if it is shown in a close-packed bag as other objects can distract attention (effect of bag complexity). In addition, the more the threat item is superimposed by other objects in the bag, the harder it becomes to detect it (effect of superposition). Furthermore, a rotated threat item is more difficult to detect than a threat item shown in the frontal view (effect of viewpoint). These image-based factors are relatively independent of training and therefore rather referred to visual abilities. The ability to cope with image-based factors can be measured using the X-Ray ORT. This test consists of 256 x-ray images, half of them including either a gun or a knife.

Manuscript received February 25, 2006.

D. Hardmeier is with the University of Zurich, Department of Psychology, 8032 Zurich, Switzerland (phone: +41-44-3852; fax: +41-44-3856; e-mail: d.hardmeier@psychologie.unizh.ch).

F. Hofer is with the University of Zurich, Department of Psychology, 8032 Zurich, Switzerland (e-mail: f.hofer@psychologie.unizh.ch).

A. Schwaninger is with the University of Zurich, Department of Psychology, 8032 Zurich, Switzerland and Max Planck Institute for Biological Cybernetics, 72076 Tübingen, Germany (e-mail: a.schwaninger@psychologie.unizh.ch).

These threat items are shown in the frontal and rotated view, more or less superimposed by other objects in the bag, in a close-packed or rather empty bag.

The above described image-based factors are supposed to play a key function in the x-ray screening process. Coping with bag complexity, superposition and viewpoint of a threat item can only be partly enhanced through training and is therefore rather dependent on the ability of each screener (see also [1]). Because these abilities play an important role in all x-ray image interpretation processes, screeners who have the relevant visual abilities should not only have a much better detection performance when untrained, but also after training and some working experience, compared to screeners who are less endued with image-based factors. To test this assumption, we compared detection performance of screeners, who were hired one year before using the X-Ray ORT with detection performance of screeners, who were selected using not fully standardized selection procedures. To compare detection performance of the two groups, we used the Prohibited Items Test (PIT). The PIT is a test including all kinds of prohibited items in x-ray images and therefore allows measuring knowledge-based factors in x-ray screening. This test provides a good possibility to measure the screener's detection performance of prohibited items independently of the selection process. Furthermore, reliability and validity of both tests, the ORT and PIT were evaluated.

II. METHOD

A. Participants

Two groups of aviation security screeners participated in this study. The experimental group consisted of 101 participants (71 male and 30 female) between 19 and 55 years (M = 35.25 years, SD = 9.79 years), who were all hired as security screeners based on the results of the X-ray ORT, which was used as part of the pre-employment assessment procedure. When taking the X-ray ORT, these job applicants had no x-ray image interpretation experience at all. Besides the X-Ray ORT, this group had to pass the color blindness test, an English test and a job interview as well in order to get employed. These screeners had about one year working experience in x-ray screening when this study was conducted (i.e. when taking the PIT).

The control group consisted of 453 screeners (141 male and 312 female) between 24 and 65 years (M = 48.94 years, SD = 9.09 years), who were hired without the X-ray ORT, but using an old selection procedure, which consisted of a color blindness test, an oral English test and a job interview. Working experience of these aviation security screeners varied from two years to 26 years (M = 9.71 years, SD = 5.50 years) when conducting this study (i.e. when detection performance in the PIT was compared to the experimental group).

B. Material

1) The X-Ray Object Recognition Test (X-Ray ORT)

The X-Ray ORT consists of 256 x-ray images and measures mainly image-based factors in x-ray screening. Therefore, only guns and knives are used as these threat objects are known by most people independent of visual experience or training and therefore are also well known by novices. Furthermore, all images are shown in black and white to eliminate color-diagnostic information for experts. To measure how good test candidates can cope with image-based factors, the image-based factors bag complexity, superposition and viewpoint are varied systematically with each other. That is, eight guns and eight knives were each combined with two bags with low complexity levels and two bags with high complexity levels, but once only little and once more superimposed by other objects in the bag. Furthermore, each bag is shown once with and once without a threat item. That is, half of the trials in the X-Ray ORT are completely harmless bags and contain neither a gun nor a knife. In the test, each image is shown for four seconds on the computer screen. Then, the test candidate has to decide whether the bag is OK (contains no gun and no knife) or NOT OK (contains a gun or a knife) by clicking the respective button on the screen. Additionally, test candidates are asked to indicate how sure they are in their decision clicking on a 50 point rating scale on the screen. For a closer description of the test design refer to [2].

2) Prohibited Items Test (PIT)

The Prohibited Items Test (PIT) was developed to measure how well aviation security screeners know what prohibited items look like in x-ray images. The PIT contains all kinds of prohibited items and thus measures mainly knowledge-based factors in x-ray screening. All prohibited items in the PIT can be classified into seven categories by ECAC, ICAO and EU prohibited items lists. A total of 19 guns, 27 sharp objects, 14 hunt and blunt instruments, 5 highly inflammable substances, 17 explosives, 3 chemicals and 13 other prohibited items (such as ivory, crocodile) are shown. In total the PIT includes 160 trials, half of them including prohibited items and half of them containing no prohibited items at all. 68 of the trials containing a prohibited item included exactly one prohibited item, whereas the other twelve trials included two or three prohibited items at once[1]. As this test was developed to measure mainly knowledge-based factors in x-ray images, all threat items were shown in an easy view, combined with bags of medium complexity level and medium superposition. Thus, all three image-based factors are kept relatively constant in the PIT. Furthermore, all images were shown in color to provide a realistic test environment.

Test taking procedure in the PIT was similar to the X-Ray ORT. First, a self-explanatory instruction was shown

[1] This was done to assure face validity. In reality more than one prohibited item can be in a passenger bag. Note that only bags including one prohibited item were used for analysis.

explaining the task followed by some exercise trials to familiarize the participants with the test taking procedure. After each of the six exercise trials a visual feedback was given whether the bag was OK (contains no prohibited item) or NOT OK (contains at least one prohibited item). In the test itself no more feedback was given to the test candidates. In the PIT, all images are displayed for a maximum of ten seconds on the screen. Test candidates have to decide whether the bag contains one or more prohibited items by clicking the OK or NOT OK button on the screen. If the bag is judged as NOT OK, screeners have to indicate to which of the seven categories the prohibited item(s) belongs to by clicking on the respective button(s)[2]. Besides giving the answer OK or NOT OK, test candidates have to indicate how sure they are in their decision by clicking on a 50 point rating bar on the screen. Pressing the space bar, the next image is shown. There are four blocks of trials, after which test candidates could take an individual short break if wanted. The order of blocks is counterbalanced across four groups of participants. Within each block the order of trials is random.

C. *Procedure*

To test whether the X-Ray ORT is a useful pre-employment assessment tool, detection performance of screeners selected without the X-Ray ORT[3] and screeners who were hired with the X-Ray ORT was compared using the test results in the PIT. All screeners who were hired with the X-Ray ORT had completed a classroom training and about one year of working experience when taking the PIT. Experience of screeners selected without the X-Ray ORT varied between two and 26 years when taking the PIT (for more details on detection performance and working experience see [3]).

III. RESULTS

All test results were calculated using the "nonparametric" detection performance measure A' (see [4], [5]). A' takes into account the hit rate (i.e. bags containing a prohibited item judged as NOT OK) as well as the false alarm rate (i.e. harmless bags judged as NOT OK). This is especially important considering the task of an aviation security screener. A screener, who judges nearly all bags as NOT OK, would for sure have a high hit rate, but at the same time a very high false alarm rate and thus be very inefficient in his job. A good screener is expected to recognize most forbidden objects without being mistaken. For further information on detection performance measures, calculation and assumptions about A' see [6], [7] or [8].

A. *Reliability and Validity of the X-Ray ORT*

Reliability of the X-Ray ORT is very high for trained aviation security screeners and novices. Cronbach Alpha values range from .887 to .966 for screeners and from .907 to .970 for novices. As well split-half reliabilities (> .781 for screeners and > .778 for novices) support reliable measurement of detection performance using the X-Ray ORT. For more details about reliability of the X-Ray ORT see [2].

Different validity measures of the X-Ray ORT were evaluated by [2] in order to determine whether the test measures what it is supposed to measure and whether it can be used in making accurate decisions. Internal, convergent and discriminant validity measures evidence the former, whereas criterion-related validity refers to the correctness of decisions. Large effects of bag complexity, superposition and viewpoint could be shown for aviation security screeners and novices and support high internal validity. Furthermore, convergent and discriminant validity could be shown based on all 453 screeners selected with the old selection procedure correlating results in the X-Ray ORT with results in the PIT ($r = .61$, $p < .001$) and results in the computer-based questionnaire (CBQ) ($r = .27$, $p < .001$), respectively. The CBQ is a multiple choice questionnaire including airport specific questions about safety and security regulations at airports. Therefore, neither the ORT nor the PIT should show a high correlation with the CBQ. Criterion-related validity was examined by correlating detection performance in the X-Ray ORT with on-the-job performance measured by Threat Image Projection (TIP) data ($r = .41$, $p < .001$). TIP systems project fictional threat images into real passenger bags during work. Therefore, TIP allows measuring on-the-job detection performance. After each TIP image screeners receive a feedback message that a fictional threat item was present. TIP data were aggregated over a period of 17 months of 86 aviation security screeners. Detection performance was calculated using A' scores, i.e. hit and false alarm rates. The correlation between the X-Ray ORT and TIP data evidences that abilities measured with the X-Ray ORT are indeed important determinants of detection performance on-the-job. For more details about calculation of these validity measures see also [2].

B. *Reliability and Validity of the PIT*

As for the X-Ray ORT, Cronbach Alpha and split-half reliabilities were calculated with 453 aviation security screeners for the PIT. All reliability measures are based on percentage corrects (PC), i.e. hits and correct rejections, as well as on confidence ratings (CR), i.e. how sure screeners were in their decision. Based on signal detection theory, reliabilities were calculated for N trials (bags without a prohibited item) and SN trials (bags with prohibited items) separately. All reliability measures are listed in Table 1 for the two groups of screeners separately. All values are very similar for both groups and support reliable measurement of detecting threat items in x-ray images. Cronbach Alpha values are ranging from .870 to .943 and split-half reliabilities from .864 to .944.

[2] The answer to which of the seven categories the prohibited item(s) belonged to, was not used for the data analysis.

[3] Screeners selected without the X-Ray ORT had to take a color blindness test, as well as a common job interview.

TABLE I

RELIABILITY ANALYSES (PIT)

Reliability Coefficients		PC SN	PC N	CR SN	CR N
Screeners (Control Group N=453)	Cronbach Alpha	.874	.901	.910	.928
	Split-half (Guttman)	.871	.914	.900	.936
Screeners (Experimental Group N=101)	Cronbach Alpha	.908	.943	.870	.883
	Split-half (Guttman)	.878	.944	.877	.864

Cronbach Alpha values and split-half reliabilities (Guttman) of the PIT calculated for screeners selected without the X-Ray ORT (N=453) and screeners selected with the X-Ray ORT (N=101): PC = percentage correct, CR = confidence ratings, SN = signal plus noise trials, N = noise trials.

Validity of the PIT can be examined calculating convergent, discriminant and criterion-related validity. These measures were calculated based on all 453 aviation security screeners who were selected without using the X-Ray ORT as pre-employment assessment tool. Convergent validity was tested correlating test scores in the PIT with test scores in the X-Ray ORT. A' scores in the PIT correlated significantly with A' scores in the X-Ray ORT (r = .61, p < .001) indicating convergent validity. This rather high correlation makes sense because both tests investigate x-ray image interpretation and obviously also in the PIT image-based factors are relevant. Furthermore, correlation between A' scores in the PIT with percentage correct answers in the computer-based questionnaire (CBQ) indicates discriminant validity (r = .26, p < .001). As for the X-Ray ORT, criterion-related validity was estimated using threat image projection (TIP) data of the same TIP-library used for the validation of the X-Ray ORT (for more details about this library please see [2]). Correlation between test results in the PIT and on-the-job detection performance (TIP data) was r = .54 (p < .001). Thus, test results in the PIT can be used to predict on-the-job performance of screeners to a certain degree.

C. *Evaluation of the X-Ray ORT as pre-employment assessment tool*

In order to investigate whether the X-Ray ORT is a valuable tool for pre-employment assessment, the mean detection performance of both groups in the PIT was compared (see Figure 1). A significant difference in detection performance of prohibited items between screeners selected without the X-Ray ORT and the group hired with the X-Ray ORT can be shown. The job applicants who were selected with the X-Ray ORT are significantly better in detecting prohibited items in x-ray images, $t(552) = 14.51$, $p < .001$ one year after employment. To test whether the difference in detection performance is influenced by the age of screeners or working experience (see [3] for the influence of these factors on x-ray detection performance) an analysis of covariance (ANCOVA) with selection procedure as between-participants factor and age and working experience as covariates was conducted. Results show that even if these two covariates are considered, detection performance of the screeners selected with the X-Ray ORT is significantly higher compared to the other screeners, with an effect size of η^2 = .07, $F(1, 548)$ = 38.82, MSE = .004, p < .001.

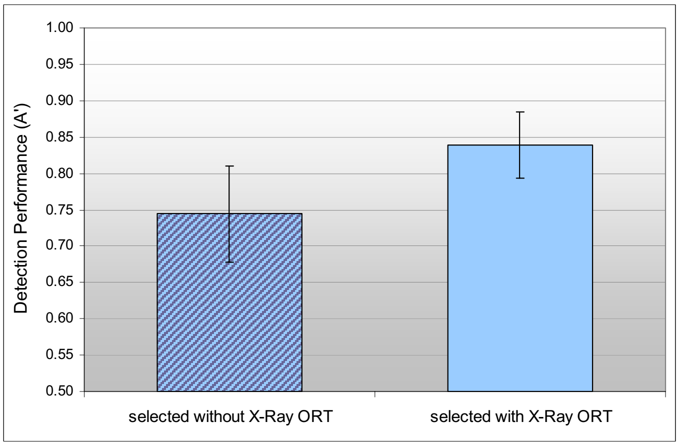


Fig. 1. Detection performance with standard deviations in the PIT for screeners selected without the X-Ray ORT (Control Group) left and screeners selected with the X-Ray ORT (Experimental Group) right.

IV. DISCUSSION

Overall, the results of the present study show that employing job applicants based on their results in the X-Ray ORT results in an increased detection performance in x-ray screening one year after employment, compared to detection performance of screeners selected without the X-Ray ORT. In this study, detection performance of two groups of screeners, each with a working experience of at least one year, was compared using the PIT, a computer-based x-ray screening test which measures rather knowledge-based factors in x-ray image interpretation. Compared to the screeners of the control group, who were not selected using the X-Ray ORT, job applicants who were hired based on the results in the X-ray ORT as pre-employment assessment tool performed significantly better in the PIT one year after employment. The effect size with η^2 = .07 ([9]) is still eminent, even when possible influences of the factors age and working experience are considered as covariates. Therefore, the ability how well someone can cope with the image-based factors, i.e. bag complexity, superposition and viewpoint, predicts detection performance in x-ray screening to a certain degree at a later date.

In addition, statistical analyses show for both x-ray performance measurement instruments high reliability and validity. The X-Ray ORT is not only a highly reliable and valid tool for measuring how well *novices* can cope with image-based factors in x-ray images, but also how well aviation security screeners with several years of working experience can handle these factors [1]. Furthermore, in this study we show that the PIT is a reliable instrument for measuring visual knowledge in the x-ray image interpretation task. Cronbach Alpha values were all > .87 and Guttman split half reliabilities > .86. Validity of the PIT was examined calculating convergent, discriminant and criterion-related validity. The large correlation between the PIT and the X-Ray ORT (r =.61) supports convergent validity. A rather low correlation of r = .26 between the PIT and the CBQ (a test measuring general knowledge about security issues at airports) evidence discriminant validity. Furthermore, criterion-related validity of the PIT is also quite high (r = .54).

To further investigate whether the X-Ray ORT used as pre-employment assessment tool can also predict on-the-job detection performance, Threat Image Projection (TIP) data could be measured. Currently, this is examined in a recently started study, in which the two screener groups will be compared with regard to their TIP performance. Because TIP data are only reliable when a large TIP-library with realistic images is used and data are aggregated over several months [10], results are not yet available. As both tests, the X-Ray ORT and the PIT show high criterion-related validity, it can be assumed that the X-Ray ORT can effectively predict on-the-job detection performance.

Besides the importance of a valid and reliable pre-employment assessment procedure, intensive individual adaptive computer-based training (CBT) is also very important to improve detection performance of security screeners during work (see for example [11] for an evaluation study of CBT). In this context, it would be interesting if screeners with high values in the X-Ray ORT show a larger training effect than screeners with low performance in the X-Ray ORT. It could be assumed that screeners who are good in coping with image-based factors profit more from training than screeners who have problems with image-based factors. This is currently also under investigation.

ACKNOWLEDGMENT

This research was financially supported by Zurich Airport Unique, Switzerland. We are thankful to Zurich State Police, Airport Division for their help in creating the stimuli and the good collaboration for conducting the study.

REFERENCES

[1] A. Schwaninger, D. Hardmeier, and F. Hofer, “Measuring visual abilities and visual knowledge of aviation security screeners,” *IEEE ICCST Proceedings*, vol. 38, pp. 258–264, 2004.

[2] D. Hardmeier, F. Hofer, and A. Schwaninger, “The X-Ray Object Recognition Test – A reliable and valid instrument for measuring visual abilities needed in x-ray screening,” *IEEE ICCST Proceedings*, vol. 39, pp. 189–192, 2005.

[3] J. Riegelnig and A. Schwaninger, “The Influence of Age and Gender on Detection Performance and the Criterion in X-Ray Screening,” *ICRAT Proceedings*, submitted for publication.

[4] J.B. Grier, “Nonparametric indexes for sensitivity and bias: Computing formulas,” *Psychological Bulletin,* vol. 75, 424-429, 1971.

[5] R. E. Pastore, E. J. Crawley, M.S. Berens, and M. A. Skelly, “”Nonparametric” A’ and other modern misconceptions about signal detection theory,” *Psychonomic Bulletin & Review*, vol. 10, no. 3, pp. 556-569, 2003.

[6] D. M. Green and J. A. Sweets, *Signal Detection Theory and Psychophysics.* New York: Wiley, 1966.

[7] H. Stanislaw and N. Todorov, “Calculation of signal detection theory measures,” *Behavior Research, Instruments, & Computers*, vol. 31, no. 1, pp. 137–149, 1999.

[8] N. A. MacMillan and C. D. Creelman, *Detection theory: A user’s guide.* Cambridge: University Press, 1991.

[9] J. Cohen, *Statistical power analysis for the behavioural sciences.* New York: Hillsdale, 1988.

[10] F. Hofer and A. Schwaninger, “Using threat image projection data for assessing individual screener performance. In: C.A.Brebbia, T. Bucciarelli, F. Garzia, and M.Guarascio, *Transactions on the Built Environment (82), Safety and Security Engineering* (pp. 417-426). Wessex: WIT Press, 2005.

[11] A. Schwaninger and F. Hofer, Evaluation of CBT for increasing threat detection performance in X-ray screening. In: K. Morgan and M. J. Spector, *The Internet Society 2004, Advances in Learning, Commerce and Security* (pp. 147-156). Wessex: WIT Press, 2004.

Assessing X-Ray Image Interpretation Competency of Airport Security Screeners

S. Koller, and A. Schwaninger

Abstract—**Baggage screening using x-ray equipment is an essential element of airport security. In order to measure threat detection performance of airport security screeners in a reliable, valid, and standardized way, the X-Ray Competency Assessment Test (X-Ray CAT) has been developed. Based on findings of object recognition studies, X-Ray CAT was designed considering the image-based factors bag complexity, superposition, and viewpoint. Furthermore, it consists of two sets of similar looking threat objects, which allows measuring transfer effects resulting from computer-based training. This study confirmed results obtained in earlier object recognition studies indicating different detection scores for different categories of threat objects and decreasing detection performance with increasing rotation of the threat object. Reliability analyses showed that the X-Ray CAT features high Cronbach Alpha and split half reliability values. Therefore, this test is a useful instrument for initial and recurrent certification and competency assessment of x-ray operators.**

Index Terms—**Airport security, x-ray screening, screener performance measurement, certification, competency assessment, quality control**

I. Assessing X-Ray Image Interpretation Competency Using X-Ray CAT

In a world of constant risk of terrorist attacks the need of a high level of security in transportation has to be satisfied. One important field is airport security. It is now recognized more and more that well trained aviation security screeners are essential in order to achieve and maintain a high level of security and efficiency at airport security checkpoints. The most expensive equipment is of little value if the humans who operate it are not selected and trained appropriately. Screeners have to know which objects are prohibited and what they look like in x-ray images of passenger bags in order to detect them effectively within a few seconds of inspection time. In order to meet this requirement, several countries use now adaptive computer-based training (CBT) such as X-Ray Tutor (XRT) for initial and recurrent training of screeners ([1], [2]).

A study by [3] showed that recognition of unfamiliar object shapes (e.g., a self-defense gas spray) in x-ray images is poor for non-screeners and much higher for well-trained aviation security personnel. In addition to such knowledge-based factors also image-based factors such as effect of viewpoint, superposition by other objects, and bag complexity produced by the number and type of other objects ([3], [4]), are considered in XRT. With this CBT screeners can be trained very effectively and efficiently using an individually adaptive algorithm ([5], [6] [7]). XRT contains thousands of x-ray images of bags and of prohibited items depicted in many different viewpoints (see [2] for details). During training, aviation security screeners are exposed to x-ray images of passenger bags containing a prohibited item (threat image) and to harmless bag images. Images are displayed for 10 seconds, and the trainees have to judge for each bag whether it is OK (contains no prohibited item) or NOT OK (contains a prohibited item). XRT is individually adaptive. It starts with threat items depicted in easy views and increases image difficulty by showing threat items in more difficult views in more complex bags and with increasing superposition by other objects. In order to prevent screeners to memorize images of bags, combinations of images of bags and threat objects are created at the point of use. This approach considers the individual training level and visual-cognitive abilities of each screener. It starts with easy images and then increases image difficulty for each individual trainee by showing threat objects in more difficult views in more complex bags and with more superposition by other objects. To conduct periodical measurements of individual screener performance the X-Ray Competency Assessment Test (X-Ray CAT) was developed and integrated in the XRT training system. In X-Ray CAT screeners have to visually search x-ray images of passenger bags for forbidden objects. The visual appearance of x-ray images is the same as during training with XRT. A scientifically reliable and standardized test is essential for individual competency assessment and certification of screeners. Like XRT, the X-Ray CAT has been developed considering scientific findings of threat detection in x-ray images of passenger bags ([3], [4]). As mentioned above, recognition of threat objects in passenger bags is dependent on the viewpoint in which the threat object is seen, the degree of superposition and the degree of clutter in the bag. Superposition is a measure of how much a threat object is superimposed by other objects in the bag (for details see equation below). In X-Ray CAT effects of viewpoint are controlled by using images of threat objects depicted in two standardized rotation angles in easy and difficult view (see below). Images of objects are combined with images of bags in a way that the two views of an object show the same degree of superposition. The bags are chosen such that they

Manuscript received February 26, 2006.
S. Koller and A. Schwaninger are of the Visual Cognition Research Group, Department of General Psychology, University of Zurich, Switzerland (phone: 0041 44 254 38 50; fax: 0041 44 254 38 56; e-mail: a.schwaninger@psychologie.unizh.ch)

are visually comparable concerning the number and form of objects with which they are packed (i.e. the degree of clutter).

The test contains two sets of objects in which object pairs are similar in shape. This allows not only measuring general effects of training by comparing the test results prior and after training, it also provides the possibility to measure transfer effects. By only training the images of one set of the test the threat detection performance for the images of the other set in a next test session indicates whether training of certain objects benefits the recognition of similar other objects. When comparing screener performance on an individual basis, tests must be reliable. Cronbach Alpha and split half reliability analyses were conducted to measure the reliability of X-Ray CAT.

II. METHOD

A. Participants

107 aviation security screeners of a European airport conducted the X-Ray CAT 1.0.0 before starting with the training using XRT SE 2.0.

B. Materials and Procedure

Stimuli were created from Smiths-Heimann Hi-Scan 6040i colour x-ray images of prohibited items and passenger bags (Fig. 1 displays an example of the stimuli). Four categories of prohibited items were chosen based on the categorization of current threat image projection systems: guns, improvised explosive devices (IEDs), knives and other prohibited items (e.g., gas, chemicals, grenades etc.). Of each category 16 exemplars are used (8 pairs). Each pair consists of two prohibited items that are similar in shape. These were divided into two sets, set A and set B. Furthermore, every item is depicted in two different viewpoints. The easy viewpoint shows the object in canonical view (see gun in Fig.1); the difficult viewpoint shows it with an 85 degree horizontal rotation or an 85 degree vertical rotation. In each threat category half of the prohibited items of the difficult viewpoint are rotated vertically, the other half horizontally. Set A and B are equalized in regard to the rotations of prohibited objects.

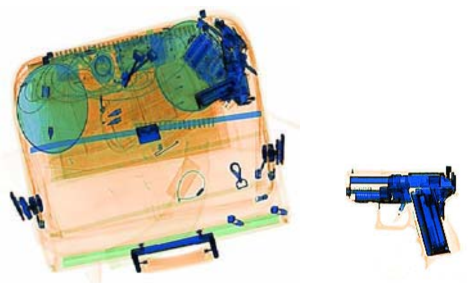


Fig. 1. Example of an x-ray image of a passenger bag containing the prohibited item depicted separately on the bottom right.

Every prohibited item was combined with a bag in a manner that the degree of superposition for one item is the same for both viewpoints. This was achieved using an image processing tool that calculates the difference of the brightness between the pixels of the two superimposed images using the following formula for superposition:

$$SP = \frac{\sqrt{I_{SN}(x,y) - I_N(x,y)}}{Threshold}$$

SP = Superposition; I_{SN} = Grayscale intensity of the SN (Signal plus Noise) image (contains a prohibited item); I_N = Grayscale intensity of the N (Noise) image (contains no prohibited item); Threshold: Number of pixels of the prohibited item where R, G and B are < 253

Using this equation the superposition value is independent of the size of the prohibited item. This value can be kept constant for the two views of a threat object, independent of the degree of clutter in the bags, when combining the bag image and the prohibited item. The bag images were checked by at least two aviation security experts to be sure they do not contain any other prohibited item. Clean bag images were assigned to the different categories and viewpoints of the prohibited items in a way that their image difficulty was balanced across all categories. This was done using the false alarm rate for each image based on a pilot study. Each bag is used twice, once containing a prohibited item (threat image) and once without (harmless image). Thus, the X-Ray CAT is composed of 256 test trials: 4 threat types (guns, IEDs, knives, other) * 8 (exemplars) * 2 (sets) * 2 (views) * 2 (harmless images vs. threat images).

The X-Ray CAT is integrated in the XRT training system and takes about 2-3 sessions of 20 minutes to complete. The visible appearance of the test is the same as during training. The task is to visually inspect the images and to judge whether they are OK (contain no prohibited item) or NOT OK (contain prohibited item). In this study, images disappeared after 10 seconds. In addition to the OK / NOT OK response, screeners had to indicate the perceived difficulty of each image on a 100 point scale (difficulty rating; 1=easy, 100=difficult). All responses are given by pressing buttons on the screen.

III. RESULTS

The two sorts of x-ray images used in the test (bags containing a prohibited item and bags containing no prohibited item) result in four different possible outcomes per trial: hits (correctly found threat objects), misses (missed threat objects), false alarms (incorrectly reporting a threat object) and correct rejections (correctly judged harmless bag as being OK). The hit rate alone is not a valid measure of detection performance in terms of sensitivity [10]. The reason is simple; a test candidate can achieve a high hit rate by simply judging most bags as NOT OK. Therefore, the false alarm rate has to be taken into account, in addition to the hit rate. A "non-parametric" measure of detection performance is A' which is calculated using the hit and false alarm rates adopting the following formula [8]:

$A' = 0.5 + [(H - F)(1 + H - F)]/[4H(1 - F)]$, whereas H is the hit rate and F the false alarm rate. If the false alarm rate is greater than the hit rate the equation must be modified [9]: $A' = 0.5 - [(F - H)(1 + F - H)]/[4F(1 - H)]$.

For details and other measures of x-ray screening performance see [10].

The results section provides detection performance, reaction time values, reliability measures and the results of an analysis of variance (ANOVA). The ANOVA was carried out to investigate whether detection performance varies across different threat categories (guns, IEDs, knives, others). In addition, potential differences and interactions regarding the two sets of images were examined.

A. *Detection Performance*

Fig. 2 shows the detection performance for each of the four categories of threat objects separately and for both views, averaged across all screeners. Detection performance is significantly higher for threat objects shown in canonical view than for threat objects that are rotated 85 degrees. No numeric values are shown since these are security sensitive data.

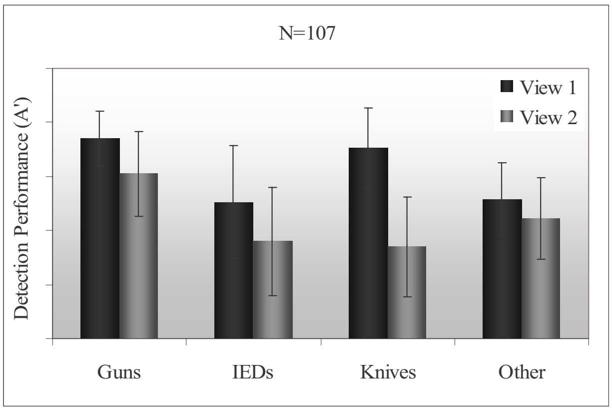


Fig. 2. Detection Performance A' broken up by category and unrotated (View 1) vs. 85° rotated objects (View 2). The thin bars are standard deviations. Pairwise comparisons showed significant viewpoint effects for all four threat categories (all p < .001).

B. *Reaction Times*

For each image, reaction times (RTs) were measured, i.e. the duration from image onset until an answer (OK or NOT OK) was given. The mean reaction times and standard deviations for each category and viewpoint are displayed in Fig. 3. In contrast to the results of detection performance, there were no viewpoint effects for the reaction times.

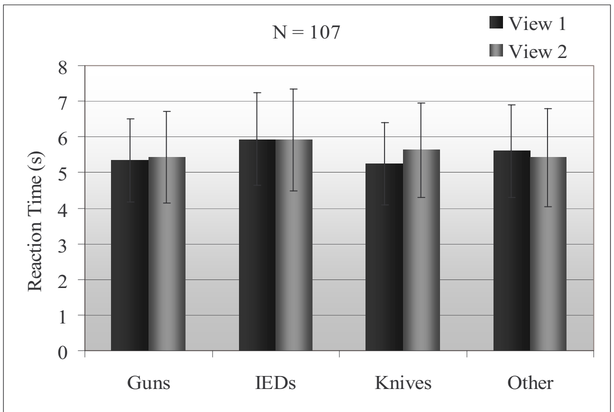


Fig. 3. Reaction times broken up by the four categories and unrotated (View 1) vs. 85° rotated objects (View 2). Thin bars represent standard deviations.

Bivariate Pearson correlations between A', RT and difficulty ratings were calculated. A' correlates with reaction times $r = -.353$ ($p < .001$) and with difficulty ratings $r = -.213$ ($p < .05$).

C. *Reliability of the X-Ray CAT*

Cronbach Alpha and Guttman split-half reliabilities were calculated based on hits and correct rejections (PC = percentage correct) as well as on difficulty ratings (DR). The analyses were made separately for test trials including a threat item (SN trials) and those not including a threat item (N trials). Table 1 shows the reliability coefficients.

TABLE 1. RELIABILITIES

RELIABILITY ANALYSES

Reliability Coefficients		PC SN	PC N	DR SN	DR N
X-Ray CAT	Alpha	.878	.914	.975	.984
	Split-half	.840	.903	.964	.980

Note. PC = Percent Correct, DR = Difficulty Ratings, SN = Bags containing a threat ("Signal plus Noise Trials"), N = Bags containing no threat ("Noise-Trials")

Additionally, a three-way ANOVA was conducted on A' scores with the three within-participants factors threat category, viewpoint and set. It showed significant main effects of prohibited items category (guns, IEDs, knives and other) with an effect size of $\eta^2 = .52$, $F(3, 318) = 112.61$, $MSE = .014$, $p < .001$, and viewpoint (canonical vs. rotated) $\eta^2 = .76$, $F(1, 106) = 343.39$, $MSE = .010$, $p < .001$, but no main effect of the set. The following two-way interactions were significant: Prohibited items category * viewpoint $\eta^2 = .35$, $F(3, 318) = 55.84$, $MSE = .009$, $p < .001$, prohibited items category * set $\eta^2 = .09$, $F(3, 318) = 10.46$, $MSE = .008$, $p < .001$, and view * set $\eta^2 = .05$, $F(1, 106) = 5.66$, $MSE = .006$, $p < .05$. These results indicate different detection performance for different prohibited items categories and higher detection performance for prohibited items in canonical view than for rotated prohibited items (effect of viewpoint, see also [3], [4]). This is consistent with results reported in the view-based object recognition literature (see [12] and [13] for reviews). The effect sizes were very large according to the conventions by [14]. The factor set shows no main effect and the effect sizes of the two interactions are small to medium ([14]). This indicates that the two sets are comparable in terms of image difficulty. Another fact supporting the comparability of the two sets is the overall correlation of $r = .715$ ($p < .001$) between them.

IV. DISCUSSION

The intention of this study was the development of a reliable test for the measurement of threat detection performance of airport security screeners. The X-Ray CAT contains images of the four prohibited items categories guns, IEDs, knives, and other prohibited items that are depicted in two different viewpoints (not rotated vs. rotated). High Cronbach Alpha coefficients showed that the X-Ray CAT is a reliable and useful tool for this important purpose (all $\alpha >$ .88). Further evidence for the good reliability of this test was revealed by the almost equally high Guttman split-half reliability (all $r > .84$). The two set composition of the test allows the investigation of transfer effects in training studies. An ANOVA on A' indicated no overall statistical difference between the two image sets. The aim of a transfer effect study would be to find out, if after training certain threat objects, a screener can transfer the acquired visual knowledge to similarly looking objects. At the moment this is being investigated in three European countries and in Canada using the X-Ray CAT. The importance of training not only canonical views of prohibited items but also rotated views is emphasized by the main effect of viewpoint in the ANOVA ($\eta^2 = .76$, $p < .001$). The results support results from earlier studies showing that detection performance can vary substantially depending on object rotation ([3], [4]). Another important result of this study was the fact that detection performance varies substantially depending on the prohibited items category. Detection performance was best for guns, followed by knives, other threat types, and IEDs. This pattern is consistent with the view that that knowledge and experience about which prohibited items exist and what they look like in x-ray images is essential for an aviation security screener to quickly and reliably perform the x-ray screening task.

Although RTs and detection performance A' correlate significantly, there were no viewpoint effects for RTs. This indicates that there is no speed-accuracy trade-off regarding effects of viewpoint. The lower detection performance for rotated objects can not be due to faster and thus less accurate responses of the screeners. The correlation between A' and reaction time values across all threat categories and both viewpoints could represent higher reaction times for more difficult threat objects on average. In other words, screeners need on average more time to respond to an image containing a more difficult threat object. This would correspond to findings in threat object recognition, that more time is spent for searching an image when a prohibited item is not recognized [2]. Search is discontinued as soon as a potential threat object is found and the answer NOT OK can be given. If no such object can be found, search can be prolonged until the image disappears from the monitor. The lower detection performance could result from a higher miss-rate (and thus lower hit rate) which could imply longer search time. This assumption finds confirmation in the significant correlation between the hit rate and average reaction time for each image containing a threat object ($r = -.683$, $p < .001$). This negative correlation signifies shorter reaction times for higher hit rates that is, the higher the probability of an image to be correctly judged as NOT OK (containing a prohibited item), the less time is needed for finding the threat object.

The significant correlation between A' and difficulty ratings (across all categories and both viewpoints) reveals the tendency of better recognized threat images to be estimated by the screeners as being easier to judge (although this correlation was rather small).

The X-Ray CAT can be used to measure effects of training by comparing the results of the test prior and after training. This is currently being done in other studies conducted in three European countries and Canada. Since the X-Ray CAT is implemented in the XRT training system, the test taking procedure is very easy. The X-Ray CAT is also a tool for quality control as it allows airports to obtain measures of the detection performance of their screeners, screening companies and checkpoints. These can be compared with each other as well as other airports or over time. Moreover, the high reliability of this test provides a solid basis for using this test in order to measure individual screener x-ray image interpretation competency and to conduct periodical certification.

ACKNOWLEDGMENT

This research was financially supported by Zurich Airport Unique, Switzerland. We are thankful to Zurich State Police, Airport Division for their help in creating the stimuli and the good collaboration for conducting parts of this study.

REFERENCES

[1] A. Schwaninger, "Computer based training: a powerful tool to the enhancement of human factors," *Aviation Security International* , FEB, pp. 31-36, 2004.

[2] A. Schwaninger, "Increasing Efficiency in Airport Security Screening," *WIT Transactions on the Built Environment*, vol. 82, pp. 405-416, 2005.

[3] A. Schwaninger, D. Hardmeier and F. Hofer, "Measuring visual abilities and visual knowledge of aviation security screeners," *IEEE ICCST Proceedings*, vol. 38, pp. 258-264, 2004.

[4] A. Schwaninger, "Evaluation and selection of airport security screeners," *AIRPORT*, vol. 2, pp. 14-15, 2003.

[5] A. Schwaninger, "Training of airport security screeners," *AIRPORT*, vol. 5, pp. 11-13, 2003.

[6] A. Schwaninger, "Computer based training: a powerful tool to the enhancement of human factors," *Aviation Security International*, pp. 31-36, February 2004.

[7] A. Schwaninger and F. Hofer,, "Evaluation of CBT for increasing threat detection performance in X-ray screening,"in *The Internet Society 2004, Advances in Learning, Commerce and Security,* K. Morgan and M. J. Spector, Eds. Wessex : WIT Press, 2004, pp. 147-156.

[8] J.B. Grier, "Nonparametric indexes for sensitivity and bias: Computing formulas," Psychological Bulletin, vol. 75, 424-429, 1971

[9] D. Aaronson and B. Watt, "Extensions of Grier's computational formulas for A' and B'' to below-chance performance," Psychological Bulletin, vol. 102, pp. 439-442, 1987

[10] F. Hofer and A. Schwaninger, "Reliable and valid measures of threat detection performance in X-ray screening," *IEEE ICCST Proceedings*, pp. 303-308, 2004.

[11] A. Schwaninger, "Reliable measurements of threat detection," *AIRPORT*, vol. 1, pp. 22-23, 2003.

[12] M. J. Tarr and H. H. Bülthoff, "Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993)," *Journal of Experimental Psychology: Human Perception and Performance,* vol. 21, pp. 1494-1505, 1995.

[13] M. J. Tarr and H. H. Bülthoff, "Image-based object recognition in man, monkey and machine," in *Object recognition in man, monkey and machine,* M. J. Tarr and H. H. Bülthoff, Eds. Cambridge, MA: MIT Press, 1998, pp. 1-20.

[14] J. Cohen, *Statistical power analysis for the behavioral sciences*. New York: Erlbaum, Hillsdale, 1988.

The Influence of Age and Gender on Detection Performance and the Response Bias in X-ray Screening

J. Riegelnig and A. Schwaninger

Abstract— **X-ray screening is an essential component of airport security. In this study we investigated influence of age and gender on detection performance (sensitivity) and response bias in x-ray screening. Airport security screeners were tested with the Object Recognition Test (ORT). The ORT consists of x-ray images of passenger bags that contain guns and knives in different views with different superposition and complexity levels. The results show that age and detection performance (sensitivity measured using A') correlate negatively. Moreover, analysis of response bias (criterion measure B'') showed that women tended to be more risk averse than men. On average, women had higher hit rates but also higher false alarm rates. However, the detection performance (sensitivity) was similar between man and women. Only a small difference was found for one screener group (CBS) favoring males, whereas for the other group (HBS) no gender difference was found regarding detection performance.**

Index Terms—**airport security, x-ray screening, age, gender differences, detection performance, sensitivity, response bias**

I. INTRODUCTION

THE relevance of security procedures has strongly increased in recent years. Aviation security has become extremely important since September 11, 2001. X-ray screening is an important technology in this field and it has improved substantially to match the increased security requirements. But the technical equipment is only one aspect of security procedures. The humans who operate the x-ray screening systems take the final decision whether a passenger bag is ok or whether there is a forbidden object in it. X-ray screening is a challenging and responsible task and screeners should be carefully selected and well trained for it. It is therefore important to identify the variables that have an influence on the detection ability of screeners. For this purpose it is important to understand the underlying cognitive processes that are involved in x-ray screening.

X-ray baggage screening is a visual cognition task that consists of several processes. A screener who looks at an x-ray image of a bag processes different objects using visual search. Whether an object is recognized as a prohibited item or not depends on image-based factors [1] such as superposition (whether the object is superimposed by other objects), rotation (how much the object is rotated) or bag complexity (whether there are many objects in the bag). Interestingly, there are large differences between people in terms of their ability to cope with these image-based factors [2]. Different models on object recognition have been proposed (for reviews see for example [3]. However, most of them agree with the following basic representational and procedural assumptions. Depending on the inspected area, an input representation is formed which is compared to what is stored in visual memory. A screener needs to be familiar with the forbidden objects in order to have developed visual memory representations of them. Most people might have a memory representation of a gun whereas they would not recognise other forbidden objects easily because they rarely see them in everyday life (e.g., an electric shock device) or because the items look quite different in an x-ray image such as for example a teaser weapon. The input representation is compared to visual memory representations which could involve mental transformations such as mental rotation. Depending on the perceived similarity of the two representations a decision is made about whether the object is harmless or forbidden and therefore whether the bag is ok or not. The x-ray screening task involves a number of cognitive abilities: visual-spatial abilities (figure-ground segregation, mental rotation), working memory capacity as well as specific knowledge about what x-ray images of prohibited items look like.

However, while the performance in x-ray screening is depending on visual abilities, the decisions that are made are also influenced by other various factors such as probability of a target stimulus, payoffs or motivation. For instance, when there is a long waiting line before a checkpoint at the airport, screeners are perhaps more likely to accept a bag as ok even though they would perhaps have it hand-searched would there be no waiting line. According to signal detection theory [4] these factors affect the subjective response bias but not detection performance in terms of sensitivity. The latter depends largely on training, e.g., [5]. In this study, we investigated whether age and gender affect response bias and

Manuscript received February 25, 2006. This work was supported in part by Airport Zurich. J. Riegelnig and A. Schwaninger are with the University of Zurich, Department of Psychology, Switzerland (corresponding author to provide phone: +41 44 254 38 50; fax: +41 44 254 38 56; e-mail: a.schwaninger@psychologie.unizh.ch).

detection performance in terms of sensitivity in x-ray screening.

A. Aging and Cognition

There is evidence that many aspects of cognition show a decline with aging (for a review see [6]). Some of the domains that could be relevant in x-ray screening and are associated with age-related deficits are: working memory [7], attention [8], perception [9], [10] and visual cognition [11].

1) Memory & Attention: One of the functions that are especially important in cognitive tasks such as x-ray screening is working memory. It is described as the central executive that processes information at a conscious level [12]. Working memory has been associated with age-related declines in various studies, e.g., [13].

Attention is also essential in x-ray baggage screening. Screeners have to be constantly vigilant which requires sustained attention (i.e. maintenance of attention across time). The relationship between age and sustained attention is not clear. For some tasks an age-related difference has been found, for others not [8].

2) Perception & Spatial cognition: There is evidence that declines in perception occur with aging. Age-related changes of eye structures lead to decreases in sharpness, contrast and brightness of visual stimuli [9]. The age-related loss of photoreceptors leads to a deficit of colour sensitivity [10]. The smaller number of receptors also influences visual processing in the periphery.

Several studies showed that younger subjects obtained higher scores in mental image rotation tasks than older subjects, e.g., [14]. There is evidence that endogenous sex hormones influence spatial cognition. Androgen levels change across the lifespan. In males testosterone levels for example decline with aging [15]. Some researchers report beneficial effects of androgen supplementation on spatial cognition in elderly people [16]. However, other studies do not provide any evidence for a beneficial effect of higher androgen levels [17].

B. Gender differences in spatial cognition

Research on gender differences suggests that men and women differ on cognitive tasks (for a review see [18]). The largest gender difference was found for spatial tasks with males performing better than females [14], [19], [20]. Men especially excel on mental rotation, the ability to rotate a two- or three-dimensional object [21]. Women, on the contrary perform better on perceptual speed [19].

There are different approaches to explain gender differences in spatial ability tasks: Hormonal influence [19], [22] socialization factors [21], evolutionary theories [23] and use of different strategies. In the following we discuss some of these theories.

1) Hormonal influence: Several researchers found a connection between hormones and spatial performance, e.g., [16], [22], [24]. Different levels of androgens seem to influence the organization of brain in early development and therefore affect certain cognitive abilities in adulthood. Activational fluctuations in the androgen level have an influence as well [19]. A nonlinear relationship between spatial ability and testosterone has been found. There is obviously an optimal level of testosterone that leads to best performance in spatial tasks. This level is in the low male range, which means that for women high levels of testosterone result in a better performance whereas for men lower levels seem to be more beneficial [16].

2) Evolutionary theories: There is a large number of evolutionary theories explaining gender differences in spatial cognition [25]. Most of them refer to a former greater space use of males than females as a cause for the development of better spatial abilities [19]. However, only few of these hypotheses are supported by data. The range size is the most plausible hypothesis. It proposes that males had to cover a larger area to maximize their reproduction and therefore developed better spatial abilities [25].

3) Socialization factors: Socialization factors also play a role in cognitive abilities. It can be assumed that predisposition and social influence interact with each other. Even though there might be a biological predisposition for certain abilities, socialization influences to what extent we use our potential [21].

C. Relevance to x-ray screening

The first research question is whether age has an influence on x-ray screening. As several studies have shown, aging leads to cognitive declines in a number of domains. As x-ray screening is a demanding cognitive task we therefore expect a negative relationship between age and x-ray screening performance. It is imaginable though that older screeners might be able to compensate for their deficits by using their domain knowledge.

The second research questions examines if there is a gender difference in x-ray screening. Spatial abilities seem to be rather important in x-ray screening and we therefore could assume that males might outperform females. However, as x-ray screening involves different cognitive processes and spatial cognition is probably only one of the required abilities, the relationship could be more complicated. Interestingly, we have heard several times from different airport supervisors that they believe women are better in x-ray screening than men, because they usually find more prohibited items.

The goal of this study was to investigate whether age and gender affect performance in x-ray screening. Performance in x-ray screening was measured with the X-Ray Object Recognition Test, X-Ray ORT [1], [2]. It has been developed to measure the screener's ability to cope with image-based factors such as rotation, superposition and bag complexity. The X-Ray ORT contains images of bags with either a gun or a knife or no prohibited item in it. The research questions are (1) whether an influence of age on detection performance (sensitivity) and response bias can be found, and (2) whether there is a gender difference in these two measures.

II. METHOD

A. Participants

A total of 410 screeners of Zurich Airport Unique at the age of 24 to 65 completed the test. Screeners were part of two different groups: cabin baggage screeners (CBS) and hold baggage screeners (HBS).

We only used data of screeners who worked more than an average of 30 percent in the past four years to make sure that all participants were working on a regular basis. With this constraint the CBS group consisted of 308 screeners (96 men, 212 women). The total number of HBS screeners was 84, i.e. 27 men and 57 women.

B. Materials and Procedure

As explained above, the X-Ray ORT, a computer-based x-ray screening test, was used in this study [1], [2]. The task in this test consists of finding weapons in x-ray images of bags. A total of 16 weapons is used, eight images of guns and eight images of knives. Before test start, all guns and knives are shown for 10 seconds on the screen, respectively. This is done in order to make sure that people are familiar with the weapons they have to look for in the test trials. There is an introductory phase with two guns and two knives that are not used in the test phase. Every threat item is shown in a rotated and in a non-rotated view. The images also vary systematically in bag complexity and superposition. Images are in black-and-white. There are a total of 256 trials: 2 weapons (guns, knives) * 8 (exemplars) * 2 (views) * 2 (bag complexities) * 2 (superpositions) * 2 (harmless vs. threat images). The trials are divided into four blocks with 64 trials each. The blocks are counterbalanced across participants using a Latin Square. Within each block the trials are presented in random order.

III. RESULTS

To measure the performance in the X-Ray ORT we used the measures A' and d' from signal detection theory [4] which are calculated from the hit and false alarm rate of each screener.

A' is calculated with the following formula:

If H (hit rate) > F (false alarm rate):

$A' = 0.5+(H-F)*(1+H-F)/[4*H*(1-F)]$

If F > H:

$A' = 0.5-(F-H)*(1+F-H)/[4*F*(1-H)]$

In the following we report only A' scores unless the two measures differ in their levels of significance. To measure the subjective response bias we used the criterion measure B'' which is also calculated from hit and false alarm rate as following:

If H (hit rate) > F (false alarm rate):

$B'' = [H(1-H)-F(1-F)]/[H(1-H)+F(1-F)]$

If F > H:

$B'' = [F(1-F)-H(1-H)]/[F(1-F)+H(1-H)]$

An outlier analysis was performed and we excluded values higher or lower than two standard deviations from the mean. This resulted in an exclusion of 14 CBS and 4 HBS participants for the analyses of detection performance and 13 CBS and 5 HBS screeners for the analyses of the response bias, which is about 5% of the data. Effect sizes η^2 are reported and can be judged based on Cohen [26].

A. Age

An effect of age on detection performance was found. To extract the influence of how long people were employed as screeners we used partial correlations taking hours on job into account (working experience). Partial correlations between A' and age controlling for working experience showed that older people performed significantly worse both in the CBS group ($r(294) = -.32$, $p < .01$) and the HBS group ($r(80) =-.35$, $p < .01$)[1].

No influence of age was found concerning the subjective response bias. Partial correlations controlling for working experience were not significant for neither screener group (CBS: $r(296) = -.03$, $p = .61$; HBS: $r(79) = .18$, $p = .12$).

B. Gender differences

CBS men had an average detection performance of A' = .88 whereas the average detection performance of women was A' = .87. Both groups had a standard deviation of $SD = .03$. An analysis of covariance (ANCOVA) with gender as between-participants factor, age and working experience as covariates, showed a small significant mean difference of gender with an effect size of $\eta^2 = .02$, $F(1, 291) = 5.46$, $MSE = .001$, $p < .05$ for the CBS group. In the HBS group both males and females had an average detection performance of A' = .91. The ANCOVA with age and working experience as covariates was not significant and had an effect size of $\eta^2 = .003$, $F(1, 76) = 0.26$, $MSE = .001$, $p = .62$. For CBS, men had a positive mean response bias ($M = .08$, $SD = .30$) whereas in women the response bias was negative ($M = -.08$, $SD = .29$). An ANCOVA with gender as between-participants factor, age and working experience as covariates showed that this difference was significant ($\eta^2 = .06$, $F(1, 292) = 18.76$, $MSE = .09$, $p < .001$). There was also a very similar gender difference regarding the response bias for the HBS group (men: $M = .07$, $SD = .21$; women: $M = -.04$, $SD = .30$, analysis of covariance: $\eta^2 = .07$, $F(1, 75) = 5.54$, $MSE = .07$, $p < .05$).

IV. DISCUSSION

In this study we examined the influence of age and gender on the detection performance (sensitivity) and the subjective response bias in x-ray screening. For this purpose a computer-based test, the X-Ray ORT was used (for details see [2]).

A negative effect of age on performance was found. Older screeners performed significantly worse in the x-ray screening test. This result coincides with literature on cognition and aging, e.g., [9], [14].

Gender had a small effect, but only in the CBS group. Men performed slightly better in the X-Ray ORT. This result also corresponds with previous research in this field, e.g., [14], [19], [20], [21]. However, the effect is small, as the effect size

[1] Partial correlation between d' and age: $r(78) = -.25$, $p < .05$

of η^2 = .02 clearly shows. In the HBS group no gender difference was found (η^2 = .003).

Women had a significantly lower response bias in both groups, CBS and HBS. This means that on average, women tend to give more often the answer that the bag is not ok. This results in a higher false alarm rate but also in a higher hit rate. If the same happens when working at the checkpoint, they would open more bags and therefore find more prohibited objects. This would explain why supervisors at airport security checkpoints sometimes claim that women find more prohibited items. Some earlier studies have indeed reported that women tend to be more risk averse than men (see [27] for a meta-analysis). This would account for the difference in the response bias between female and male screeners.

The results of this study correspond for the most part with results of cognition-related research. As explained in the introduction, X-ray screening involves several cognitive functions. Earlier research has shown that certain cognitive functions are affected by aging [6], [9]. In this study we also found a decrease of detection performance in terms of sensitivity. But we do not yet know which functions are most important for detection performance. Obviously older people could not compensate for their deficits. It seems that they had no more domain knowledge than younger participants or that this domain knowledge could not be used in order to compensate for impaired detection performance. However, we must be cautious regarding the implications for real life screening operations of these results. Compared to training effects, the influence of age is rather small. Ghylin, Drury and Schwaninger reported huge training effects when screeners used an individually adaptive computer based training system, X-Ray Tutor. The effect size was η^2 = .505, which was 50 times larger than the age effect (η^2 = .011) [28].

As we have seen gender differences are also rather small. Research about cognitive gender differences suggests that men perform better in spatial tasks [20]. X-ray screening is quite a complex process and requires probably a variety of different abilities and maybe not only spatial cognition abilities.

It is important for aviation security to know on what factors detection in x-ray screening depends. The cognitive processes which underlie this task must be understood. If we know what abilities are essential, how these abilities are influenced, and how they can be improved, selection and training of screeners can be as effective as possible.

ACKNOWLEDGMENT

This research was financially supported by Zurich Airport Unique, Switzerland. We are thankful to Zurich State Police, Airport Division for their help in creating the stimuli and the good collaboration for conducting this study. We thank Diana Hardmeier for her help in data collection.

REFERENCES

[1] A. Schwaninger, "Evaluation and selection of airport security screeners," *AIRPORT,* pp. 14-15, Feb. 2003.

[2] A. Schwaninger, D. Hardmeier, and F. Hofer, „Aviation security screeners visual abilities & visual knowledge measurement," *IEEE Aerospace and Electronic Systems*, vol. 20, no. 6, pp. 29-35, 2005.

[3] S. M. Kosslyn, *Image and Brain. The resolution of the imagery debate.* Cambridge, Massachusetts: MIT Press, 1994.

[4] D. M. Green, and J. A. Swets, *Signal detection theory and psychophysics*. New York: Wiley, 1996.

[5] A. Schwaninger, and F. Hofer, "Evaluation of CBT for increasing threat detection performance in X-ray screening," in: K. Morgan, and M. J. Spector, *The Internet Society 2004, Advances in Learning, Commerce and Security*, Wessex: WIT Press, pp. 147-156, 2004.

[6] F. I. M. Craik, and T. A. Salthouse, *The handbook of aging and cognition*. Mahwah NJ: Erlbaum, 2000.

[7] L. Bäckman, B. J. Small, and Å. Wahlin, "Aging and Memory: Cognitive and Biological Perspectives," in: J. E. Birren and K. W. Schaie, *Handbook of The Psychology of Aging,* San Diego: Academic Press, 2001.

[8] W. A. Rogers, and A. D. Fisk, "Understanding the Role of Attention in Cognitive Aging Research," in: J. E. Birren and K. W. Schaie, *Handbook of The Psychology of Aging,* San Diego: Academic Press, 2001.

[9] R. Cabeza, "Functional Neuroimaging of Cognitive Aging," in: R. Cabeza and A. Kingston (Eds.). *Handbook of Functional Neuroimaging of Cognition*, Cambridge: MIT Press, MA, 2001.

[10] J. L. Fozard, and S. Gordon-Salant, "Changes in Vision and Hearing with Aging," in: J. E. Birren and K. W. Schaie, *Handbook of The Psychology of Aging,* San Diego: Academic Press, 2001.

[11] M. J. Sharps, "A developmental approach to visual cognition in the elderly," in: T. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization,* North-Holland: Elsevier Science Publishers, 1990.

[12] A. Baddeley, *Working memory*, New York, Oxford University Press, 1986.

[13] K. E. Cherry, and D. C. Park, "Individual difference and contextual variables influence spatial memory in younger and older adults," *Psychology and Aging*, vol. 8, pp. 517–526, 1993.

[14] A. Campos, M. J. Pérez-Fabello, and R. Gómez-Juncal "Gender and age differences in measured and self-perceived imaging capacity," *Personality and Individual Differences*, vol. 37, pp. 1383-1389, 2004.

[15] A. Vermeulen, "Androgen Replacement Therapy in the Aging Male – A Critical Evaluation," *The Journal of Clinical Endocrinology & Metabolism*, vol. 86, no. 6, pp. 2380-2390, 2001.

[16] J. S. Janowsky, S. K. Oviatt, and E. S. Orwoll, "Testosterone Influences Spatial Cognition in Older Men," *Behavioral Neuroscience*, vol. 108 no. 2, pp. 325-332, 1994.

[17] O. T. Wolf, and C. Kirschbaum, „Endogenous Estradiol and Testosterone Levels Are Associated with Cognitive Performance in Older Women and Men," *Hormones and Behavior*, vol. 41, pp. 259-266, 2002.

[18] D. F. Halpern, *Sex Differences in Cognitive Abilities*, edn. 2. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1992.

[19] D. Kimura, "Sex, sexual orientation and sex hormones influence human cognitive function," *Current Opinion in Neurobiology*, vol. 6, pp. 259-263, 1996.

[20] D. Voyer, S. Voyer, & M. P. Bryden, "Magnitude of Sex Differences in Spatial Abilities: A Meta-Analysis and Consideration of Critical Variables," *Psychological Bulletin*, vol. 117, no. 2, pp. 250-270, 1995.

[21] M. B. Casey, "Gender, Sex, and Cognition: Considering the interrelationship between biological and environmental factors," *Learning and Individual Differences*, vol. 8, no. 1, pp. 39-53, 1996.

[22] I. Silverman, D. Kastuk, J. Choi, and K. Phillips, "Testosterone levels and spatial ability in men," *Psychoneuroendocrinology*, vol. 24, pp. 813-822, 1999.

[23] D. C. Geary, "Sexual Selection and Sex Differences in Spatial Cognition," *Learning and Individual Differences*, vol. 7, no. 4, pp. 289-301, 1995.

[24] L. S. Liben, E. J. Susman, J. W. Finkelstein, V. M. Chinchilli, S. Kunselman, J. Schwab, J. Semon Dubas, L. M. Demers, G. Lookingbill, M. R. D'Arcangelo, H. R. Krogh, and H. E. Kulin, "The Effects of Sex Steroids on Spatial Performance: A Review and an Experimental Clinical Investigation," *Developmental Psychology*, vol. 38, no. 2, pp. 236-253, 2002.
[25] C. M. Jones, V. A. Braithwaite, and S. D. Healy, "The Evolution of Sex Differences in Spatial Ability," in: *Behavioral Neuroscience*, vol. 117, no. 3, pp. 403-411, 2003.
[26] J. Cohen, *Statistical power analysis for the behavioral sciences*, New York, Erlbaum, Hillsdale, 1988.
[27] J. P. Byrnes, D. C. Miller, and W. D. Schafer, „Gender Differences in Risk Taking: A Meta-Analysis," *Psychological Bulletin*, vol. 125, no. 3, pp. 367-383, 1999.
[28] K. M. Ghylin, C. G. Drury , and A. Schwaninger, "Two-component model of security inspection: application and findings," in Press.

Doctoral Symposium

Enterprise Risk Management Perceptions in Airports of Turkey

Ayse Kucuk Yilmaz, PhD Canditate

Abstract— **Organizations operate in an environment which includes many factors resulting in very rapid, intense and continuous changes such as globalization, regulations, re-engineering, changing markets and competitive structures. These changes increase the uncertainties, major source of risks in future decisions. Sources and management methodologies of risk are closely related with business strategy. Risk management is an essential discipline which will increase the success of this strategy. Air transportation, catalyst of the entire economy, have two primary concerns: ensuring strategic advantages and competitive superiority, and obtaining maximum shareholder value in this volatile environment like any other business. Survival of the enterprises strongly depends on their ability of managing entire corporate risks. Implementation of Enterprise Risk Management (ERM) can offer an efficient business framework for managing these risks.**

***Index Terms*— Airport Business, Enterprise Risk Management, Risk Management Perception.**

I. INTRODUCTION

THE objective of this study is to determine existing risk infrastructures and strategies, organizational structures, ERM approach and perception at airports in Turkey, a typical example for developing countries. Allocation and use of limited resources through an effective and systematic risk management effort are critical in Turkey as well as other developing countries. Determination of risk approaches and perceptions can provide a better understanding in establishing necessary infrastructure for ERM.

Commercialized and privatized airports are improving and increasing in terms of service concepts, quantity and network in Turkey. These means those, airport business are in the process of rapid development. The process is risky and risks are must managed. Inflation has become a major obstacle for the improvement of financial markets in Turkey for the last 30 years. Financial crises had a negative impact on air transportation industry as well as entire economy in Turkey in the last decade. Meanwhile, international developments gained speed because of globalization and technology.

The study consists of three sections: the importance and necessity of ERM and ERM perceptions in airports as well as other organizations are explained in the first section. In the second section, the results of a questionnaire implemented to the prominent airport in turkey, a representative developing country, are presented. In the final section, future works are presented with suggestions.

II. IMPORTANCE AND NECESSITY OF ERM FOR AIRPORTS IN TURKEY

Air transportation industry one of the world's most important industries. Air transportation plays a vital role in developing and facilitating economic growth, particularly in developing countries like Turkey. Its development and technical and service achievements make it one of the greatest contributors to the advancement of the modern society.

Airports are an interface between air and ground transportation system. This is an unchanging fundamental for airports. Therefore society becomes prosperous through the flow of people and goods. Airports are critical and fundamental element of air transportation system. Additionally, airport business practices play a critical role in shaping airline competition [1].

Development and changes in the air transportation such as deregulation, liberalization, open skies, privatization and commercialization, and competitive structure, have an influence over airports as a part of air transportation system [2]. Uncertainty and risks are increasing and driving due to this developments and changes. ERM is essential for survival of airport business. Risk perceptions must be developed before managing risks. Risks are unmanageable without a right perceptions and definition.

Though many organizations require understanding how ERM can contribute their value, implementations regarding this issue are very limited in both quantity and context. Therefore, many problems are appearing in ERM perceptions and implementations. This study is prepared to answer this need and establish an infrastructure for the studies regarding ERM understanding and perceptions.

Ayse Kucuk Yilmaz is with School of Civil Aviation, Anadolu University, Eskisehir, CO 26470, Turkey (phone: 90-222-321-3550 (ext. 6812); fax: 90-222-322-1619; e-mail: akucukyilmaz@ anadolu.edu.tr).

III. SURVEY HYPHOTHESIS, OBJECTIVES AND METHODOLOGY

Airport businesses have more risks than any other industries. These risks appear suddenly and have strong impacts. Air transportation risks results are suddenly and high impact.

A. Hypothesis of this study;

- ERM provides more effective business in a risky environment. This is an important advantage for sur-petition in environments filled with risks, and in volatile markets such as Turkey
- ERM is a fundamental enterprise function and ERM have influences on business vision, mission and strategy
- All organizations –profit or non-profit- face uncertainty. ERM framework is a support for the organizations about these challenges.

B. Objectives of this study are;

Determine the perception and approach about risk and risk management in airport management in Turkey. So, we can determine ability of risk management implementation in the current airports.

C. Methodology of this study is;

The qualitative research technique is used in this survey. The questionnaire, including structured and unstructured question sets is prepared and applied to the major airports in order to investigate these issues. This questionnaire is applied to airports servicing in domestic and international. 60 Questions are prepared to demonstrate elements of effective ERM. Also, the questionnaire is prepared to determine the situation of governance since ERM interrelated with organizational objectives and mission. Questionnaires are sent to all major airport managers in Turkey. Interviews are completed. Airports are highly interested in this questionnaire about risk management perceptions.

The questionnaire is structured in the following sections:

1. General Information about Organization
2. Organizational Objectives; Organization and culture
3. Understanding risk and risk management; risk perceptions, risk identification
4. Risk management policy

IV. QUESTIONNAIRE RESULTS

Numbers of airport business are very limited(5) in Turkey. Statistical analyses don't used for assessment in answers of questions on questionnaire. Statistical analysis isn't give meaningful results as number of airports is very low. Qualitative method is used for evolutions of questionnaire answers of questionnaire in this study. Also, questionnaire is a type of qualitative analysis. Our aim was gathered approach and opinions of airports managers by apply of questionnaire.

In Turkey, major airports (such as Atatürk Airport, Antalya Airport, Esenboğa Airport, Adnan Menderes Airport) have been operated by private sector while the rest of them have been operated by government body: DHMI. The DHMI, the Air Navigation Service Provider of Turkey, provides high quality services in a safe ATM environment. DHMI operates 34 airports and aerodromes [3].

The response level was high; we sent out 5 questionnaires and received 5 back. Also we are interviewed with all airport operator managers. Also, all of them currently interesting ERM, The responses provide strong evidence that the board is becoming more interested in risk management and want to implement it in their organizations. In terms of ERM, the respondents indicated the fact that there is an increasing awareness of non-operational risks by security, safety and financial risk management personnel, and vice-versa, also suggested that companies are moving toward an enterprise-wide view of risk.

Results of questionnaire are shows that the airport managers are perceived risk and ERM within their own managerial and cultural framework. Their perceptions of risk and ERM are very similar and narrow with traditional risk management at airports and what executives care most about in terms of the risks that they face varies according to their area of business: security, safety and financial. Our study's results show that airport managements think more seriously about security than they did previously, but gaps remain. ERM required that corporate risks must consider with holistic approach. In addition, some airports managers want to establish ERM systems with regard to corporate risks that they face in their entire business and operations, including their ability to compete, achieve to strategic objectives and comply with new regulations and global business environment.

There are concise results of questionnaire analysis. Answers of questionnaire's questions show that following statements and implementations exist;

- The organization's overall aims are clearly set out and published in a manner that can be understood easily by executive management. Staff understand how the aims and objectives of the organization link to their personal [work related] objectives.
- Risk Management philosophy is different from each other. Some organizations are managing risk as system-based. Other organizations are managing risk as unit-base. Other organizations are managing risk context of system and portfolio.
- Effective risk management is important to the achievement of their organization's objectives. It is accepted by airports that "Effective risk management can improve organization's performance."
- Organizations are able to allocate appropriate resources in support of risk management policy and practice. But

perceptional and cultural deficiencies are exists related with risk management perceptions.

- Training has been provided for management by whole organizations on: risk, risk policy, procedures and practices, risk taking. But it is not enough.
- A common definition of business risk isn't used throughout the organizations. Generally, organizations have not a documented risk management policy.
- The cultures of organizations are tending to reflect a risk averse attitude.
- Risk management is partly integrated into the organizational objectives, their strategy, initiatives and efforts in their organization at high degree.
- In identifying risks, organizations are considered the following sources of risk: strategic environment and conditions; organizational environment and conditions; risk Management environment and conditions.
- Organizations are identified risks in terms of: "What can happen?; How and why risks arise?; Area of impact? ; The source of the risk?"
- Generally, boards of directors, managers and staff have the responsibility for managing risks. Various risk management standards, guidelines, policies, internal and external reports are known by airport managers.
- Following tools and techniques are used by these organizations for identifying their risks: " Audits or physical inspection, Brainstorming, Examination of local/overseas experience, SWOT analysis, Interview/focus group discussion, Judge-mental, Surveys/questionnaires, Scenario analysis, Operational modeling, Past organizational experience and Process analysis".
- The responsibility for risk management is not clearly set out and understood throughout the entire organizations. The organization's risk management procedures and processes are not documented and provide guidance to staff about managing risks

Consequently, Risk management perceptions do not exist in airport managements and staff as a concept of ERM. Their implementations are restricted with threat and financial risk management. They do not have practice of enterprise-wide risk management implementations. However, these airport managers are willing to increase the awareness for ERM at level of corporate. Their efforts are improving across their organizations in this field. Besides, these organizations try to make their staff conscious for global crisis and grow which is directly affected to risk management in aviation. Risk culture partly exists in these organizations and their efforts are try to increase it. Risk management is very limited integrated with organizational strategy, initiative and efforts in these organizations.

Interviews are going on with many of these organizations. They request for us following studies; "Defining standards, Defining risk management organizations, determining roles end responsibilities, Establishing risk management framework for their organizations, Assessment of risk management framework." The output of this study lent support to this process and assist managers with preparing their frameworks for managing enterprise risk in a way that fits the operating style of their organization and builds on good practice and practices already in place in the organization.

V. Conclusions and Benefits of Study and Future Works

I am preparing doctoral thesis about "airport enterprise risk management and model suggestion for Turkey airport". The questionnaire results will be used in this doctoral dissertation. Existing situation will determine and analyze related with the applicability of ERM in airport at Turkey. In these concepts, model suggestions will be prepared with University-private sector cooperation. Most of these organizations want to establish ERM by collaborative study with us. This is a strong and important indicator for changing traditional risk management perceptions and approach. Finally, our questionnaire is reached to our aim which is to determine risk management perceptions.

My study is important and beneficial for academic areas and airport business sector for the following reasons;

- Both ERM concept and practices are very limited in Turkey. This study and dissertation will be answer for this deficiency.
- This study and dissertation will be guidelines for scientific studies and airport business management for determining existing situation in airports at Turkey related with applicability of ERM
- This study and dissertation will provide various benefits for aviation sector across world, especially for Turkey.

ERM is becoming important day by day. ERM is slowly having some attention. This questionnaire and doctoral thesis are contributed for increasing of attention and consideration at ERM. It will also provide a common language regarding to an effective ERM process. The success of policies and procedures depend critically upon a positive risk culture and risk perceptions. Every organization can benefit from developing a risk culture since no entity operates in a risk free environment. ERM simply enables management to operate more effectively in environments filled with risks, and in volatile markets such as Turkey.

Appendix

Questionnaire sources

Questions of Questionnaire prepared to benefit from following sources;

- The Audit Office of New South Wales, "Risk Management Questionnaire", www.audit.nsw.gov.au/publications/questionnaires/RiskSurveyInstrument.pdf , June 2002.
- The Institute of Internal Auditors "Enterprise Risk Management Questions Management and the Board Should

Consider”, Corporate Board Member Magazine, August 2003.

REFERENCES

[1] FAA/OST Task Force Study, “Airport Business Practices and Their Impact on Airline Competition”, http://ostpxweb.dot.gov/aviation/Data/airportsbuspract.pdf, October 1999, p.iii.
[2] Kaya, Ergun. “ Build-Operate- Transfer practices in Airport in Turkey”, T.C. Anadolu University Publications; No.1649; Publication of Scholl of Civil Aviation; No.13; Eskisehir, Turkey, 2005, p.19.
[3] www.dhmi.gov.tr, February 2006.

BIOGRAPHY

Ayse Kucuk Yilmaz. School of Civil Aviation, Anadolu University, Turkey in 2001, Master MSME, Anadolu University Social Science Institute, Civil Aviation Management, 2003. She is an aircraft technician and PhD candidate at Anadolu University-School of Civil Aviation in Eskisehir, Turkey.

A multilevel modeling approach for air transportation systems

Aïcha Alou Oumarou, Félix Mora-Camino, Marc de Coligny, Amadou Mahamadou Handou

***Abstract*— In this communication Graph Theory is used to propose a full structural model for an air transportation system operating over a given geographical area so that a global framework is made available for decision making at planning and operations levels by the different actors of the system: ATM authorities, airports, airlines and last but not least, passengers and freight.**

***Index Terms*— Air transportation planning, flows in networks, graphs, modeling, multi agents systems.**

I. Introduction

Along the last decades, many studies in the fields of Operations Research, Systems Management and Applied Economics have been devoted to air transportation planning and operations related issues. In general the analysis which have been performed are limited to direct effects so that the scope of the adopted models are in general too limited. This implies that feedback phenomena between the different actors and involved air transportation activities cannot be fully taken into account to perform a comprehensive analysis and to design efficient plans and policies .

In this communication a global framework, based on graph theory and flow modeling is proposed. This framework is composed of a set of graphical models associated to the different perceptions each interacting organization inside the air transportation system has of its own environment. To insure global coherency between these models, compatibility constraints are introduced. These constraints are mainly relative to interactions between physical flows and available capacities at different levels. In this communication the analysis is limited voluntarily to physical flows and does not consider costs and price structures as well as financial and information flows between the different actors of the system.

In a first step, structural analysis is performed top-down, starting from the macro level of controlled airspace and resulting traffic, then going to the level of airlines fleet networked operations and finally reaching the level of end users (passengers and freight) structured demand and markets. Then, in a second step, the compatibility constraints are introduced in a bottom-up approach, starting from passengers flows, going through fleet conservation flows and ending up with air traffic flows constraints. Thus, a full inventory of the physical flows constraints involved in air transportation is achieved.

The proposed framework should be suitable to promote studies related with such different issues as: passengers/freight demand analysis and forecast; fleet dimensioning, composition and assignment; airlines competition and network revenue management; air traffic capacity analysis and air traffic management; pricing for airports and ATC services.

II. Structural Modeling of Distributed Organizations

A. *The geographical area considered*

The geographical area is covered by the same air traffic control organization and includes n cities whose associated airports compose a set of airports. Let be the set of local airlines that serve this area. In this study competition with ground transportation modes is not directly considered and so, potential demand for air travel is taken as given.

The period of time retained for this study is the year. It is also considered that the set of local airports $\hat{A}$ is connected to a set of external airports $\breve{A}$ located out of the geographic area of study and that the corresponding connections can be operated either by local or by non-local airlines. Then, the total set of airports of interest is given by:

$$A = \hat{A} \cup \breve{A} \quad (1)$$

When external airlines operate a same connection between a local airport and an external airport they will be taken as a single airline. Let $\breve{C}$ be the set of external airlines connecting the area to the outside world, then the total set of airlines of interest is given by:

$$C = \hat{C} \cup \breve{C} \quad (2)$$

A small scale example is displayed in figure1, where the area of study is composed of nine airports while two outside

Manuscript received on February 21, 2006.

A. Alou Oumarou is with the Laboratoire d'Automatique et de Recherche Opérationnelle - LARA, ENAC, Toulouse, (33-56217-4693; fax: 33-56217-4403; e-mail: aicha.alou-oumarou@ enac.fr)

F. Mora-Camino is with the Laboratoire d'Automatique et de Recherche Opérationnelle - LARA, ENAC, Toulouse, (e-mail: felix.mora-camino@enac.fr)

M. de Coligny is with the MARTA, University Toulouse II, Toulouse, France (e-mail: coligny@univ-tlse2.fr).

A. M. Handou is with the Department of air transportation, EAMAC/ASECNA, Niamey, Republic of Niger

are connected to this area.

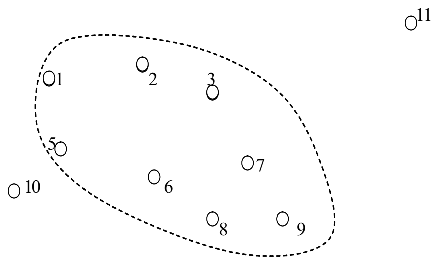


Fig. 1. An example of geographical area

Here, the different sets of airports are given by:

$$\widehat{A} = \{1,2,3,4,5,6,7,8,9\},\ \breve{A} = \{10,11\} \quad (3)$$

B. Air traffic and control organization modeling

The level of modeling adopted here is quite rough and retains only the TMA's associated to the airports of set $\widehat{A}$, the upper space control sectors covering the whole area and the airways crossing it. So the airspace is divided in T sectors, each of them having a limited air traffic control capacity while air traffic flows are supposed to follow predetermined routes connecting the airports.

Let $G_T = [A_T, U_T]$ be the connected graph associated to the airways and the airports, where U_T is the set of.
arcs composed of segments of airways and A_T is given by:

$$A_T = A \cup A_X \quad (4)$$

where A_X is the set of intersecting points of the different airways .

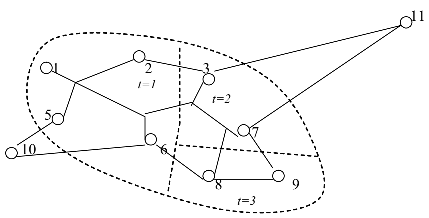


Fig. 2. Example of simplified airspace structure

The resulting graph can correspond to situations so different as direct point to point connections (corresponding for instance to a free flight situation) resulting in many intersections, or as a concentrated traffic over a maximal tree connecting the different airports and leading to a reduced number of flows crossings.
On figure 2, each edge represents two opposite segments of airway to which different flight levels are assigned in accordance with air traffic regulations. This graph could be refined by introducing the intersections of the airways with the limits of the control sectors of the area under study. Only airways intersection points within this area will be taken into account. Then graph G_T is composed of four types of arcs (*i*, *j*) :

$$u(i,j) \in U_T \Rightarrow (i \in A,\ j \in A)\ or\ (i \in A,\ j \in A_X)\ or\ (i \in A_X,\ j \in A_X)\ or\ (i \in A_X,\ j \in A) \quad (5)$$

To each arc α of U_T is associated its length d_α and to each vertex *i* of A_T is assigned its traffic area written $\tau(i)$:

$$i \to \tau(i) \in \{1, 2, \cdots, T\} \quad (6)$$

Let n_t be the number of vertices of G_T present in that air traffic control sector t, $t \in T$. A bipartite graph G_{CT} can be built between traffic sectors and the edges of graph G_T that intercept them :

$$G_{TC} = [C \cup U_T, \Gamma_{CT}] \quad (7)$$

Each arc α of this bipartite graph can be characterized by the number of times it intercepts the limits of sector *t*, $p(t, \alpha)$. Then, to each arc $\alpha = (i, j)$ of the graph G_T, can be assigned the proportion a_{ij}^t of the section of route i-j which belongs to control sector *t*.

These proportions are linked together by the two following sets of relations:

$$\left.\begin{array}{l}\sum_{t \in T} a_{ij}^t = 1 \\ 0 \le a_{ij}^t \le 1, t \in T\end{array}\right\} \quad \forall i \in \widehat{A} \cup A_X, \forall j \in \Gamma(i)\ and\ j \in \widehat{A} \cup A_X \quad (8)$$

where $\widehat{A}$ is the set of airports located inside the considered geographical area,
and

$$\left.\begin{array}{l}\sum_{t \in T} a_{ij}^t < 1 \\ 0 \le a_{ij}^t < 1, t \in T\end{array}\right\} \quad (\forall i \in \widehat{A} \cup A_X, \forall j \in \Gamma(i)\ and\ j \in \breve{A}) \quad or \quad (\forall i \in \breve{A}, \forall j \in \Gamma(i)\ and\ j \in \widehat{A} \cup A_X) \quad (9)$$

where $\breve{A}$ is the set of airports located outside the considered geographical area.
The length of arc $\alpha = (i, j)$ that is in traffic sector *t* is given by :

$$\Delta_{ij}^t = d_{ij}\ a_{ij}^t \quad (10)$$

C. Airlines networks modeling

It is assumed that airlines, always adopt, on a direct flight between an origin airport *i* and a destination airport *j*, the path of minimum length over the air traffic graph G_T. It is assumed that this minimum length path, π_{ij} , is unique for each o-d pair. Let δ_{ij} denotes its length. It is then possible to define incidence matrices between o-d pairs and arcs of G_T :

$$\begin{cases} \omega_{ij}(\alpha) = 1 & if\ \alpha \in U_T\ and\ \ \alpha \in \pi_{ij} \\ \omega_{ij}(\alpha) = 0 & else. \end{cases} \quad (11)$$

and the o-d length for the airlines is given by:

$$\delta_{ij} = \sum_{\alpha \in U_T} \omega_{ij}(\alpha)\ d_\alpha \quad (12)$$

In this study it is assumed that each company *c* can have no more than two types of aircraft: a number L_c of short/medium range aircraft and a number of S_c medium/long range aircraft .
A local company *c* is characterized by its set of airport bases $B_C \subset \widehat{A}$, from where she builds its network. According to the length δ_{ij} of a direct flight from airport *i* to airport *j* , the

company can operate on this direct flight either a short/medium range aircraft or a medium/long range aircraft . Before considering the set of airports served by an airline, it is suitable to define the non oriented graph G giving support to the effective network operated by the company with a fleet of short/medium range aircraft and medium/long range aircraft:

$$G = [A, \Gamma] = [A, \Gamma_L] \cup [A, \Gamma_S] \tag{13}$$

where $\Gamma_L(i)$ is the set of airports that can be reached in A from airport i in $\widehat{A}$ with a medium/long range aircraft and where $\Gamma_S(i)$ is the set of airports that can be reached in A from airport i in $\widehat{A}$ with short/medium range aircraft. Then:

$$\Gamma = \Gamma_L \cup \Gamma_S \tag{14}$$

The network operated by the fleet of a local airline c from its bases is such that :

$$G_c = G_c^L \cup G_c^S \tag{15}$$

with for the short/medium range network:

$$G_c^S = [A_c^S, \Gamma_c^S] \tag{16}$$

where $$A_c^S \subset B_c \cup \Gamma_S(B_c) \cup \cdots \cup \Gamma_S^{(K_S)}(B_c) \tag{17}$$

Here, K_S is an integer (a possible maximum value can be taken as 5). Here, any airport out the base of A_c^S is reachable from an airport of this base by at least a path of length inferior or equal to K_S and there is at least a path of length inferior or equal to K_S between this airport and an airport of the base. Γ_c^S is hence the restriction of Γ_S to A_c^S and insures that the aircraft rotations, in a periodic operational framework, have a limited value.

The medium/long range network operated by airline c is such that:

$$G_c^L = [A_c^L, \Gamma_c^L] \tag{18}$$

where

$$A_c^L \subset B_c \cup \Gamma_L(B_c) \tag{19}$$

Γ_c^L is therefore the restriction from Γ_L to A_c^L .

Fig. 3 and *Fig. 4* display an example of airlines networks for the area considered in *Fig. 1*.

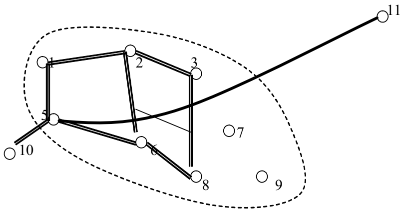


Fig. 3. Network operated by local airline c_1=1.
Here the set of airline bases is $B_{c_1} = \{1, 5\}$, and the short/medium range network is given by : , while the medium/long range network is given by:

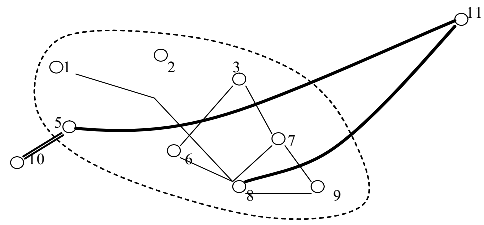


Fig. 4. Network operated by the other airlines.

Airline c_2 has a base at airport 8 and operates a small/medium range network : , $\breve{c}_1$ operates a medium/long range network between airport 11 and airports 5 and 8, airline $\breve{c}_2$ operates a small/medium range link between airports 10 and 5.

By comparing the networks displayed in figure 3 and 4, it appears that airline c_1 is in competition with airline $\breve{c}_2$ on the o-d's (5,10) and (10,5), and with airline c_2 on o-d's (6,8) and (8,6), and with airline $\breve{c}_1$ on o-d's (5,11) and (11,5).

To local airline c it is possible to associate two networks, one relative to its short/medium range activity and one relative to its medium/long range activity.

$$R_c^L = [G_c^L, [f_{ij}^{cL}]] \tag{20}$$

where $[f_{ij}^{cL}]$ is the flow of aircraft between airports i and j on the medium/long range network of airline c , and :

$$R_c^S = [G_c^S, [f_{ij}^{cS}]] \tag{21}$$

where $[f_{ij}^{cS}]$ is the flow aircraft between airports i and j on the short/medium range network of airline c . The total network operated by airline c is then given by:

$$R_c = [G_c, [f_{ij}^c]] \tag{22}$$

with

$$G_c = [A_c, \Gamma_c] \quad \text{where} \quad A_c = A_c^L \cup A_c^S \tag{23}$$

$$\text{and} \quad \Gamma_c = \Gamma_c^L \oplus \Gamma_c^S \tag{24}$$

G_c^S is such that:

$$\underline{l}_{\min} \le \delta_{ij} \le \underline{l}_{\max} \tag{25}$$

where $\underline{l}_{\min}$, $\underline{l}_{\max}$, $\bar{l}_{\min}$ and $\bar{l}_{\max}$ are threshold distances associated to the different aircraft categories (short/medium and medium/long ranges).

In the case of a non local company, there are short/medium range or medium/long range point to point connections between an airport of $\widehat{A}$ and an airport of $\breve{A}$.

To the set of non local airlines $\breve{C}$ corresponds therefore $2|\breve{C}|$ connections (one in each way).

$$\forall c \in \breve{C}, \exists!\, i_c \in \widehat{A} \ et\ \exists!\, j_c \in \breve{A} \ such \ \ as \ A_c = \{i_c, j_c\} \tag{26}$$

and according to the distance between i_c and j_c, the connection will be of the L type or of the S type. Then, either :

$$G_c^L = [A_c^L, \Gamma_c^L] \quad \text{where} \quad A_c^L = \{i_c, j_c\} \tag{27}$$

$$\text{with} \quad \Gamma_c^L(i_c) = j_c, \Gamma_c^L(j_c) = i_c \tag{28}$$

or:

$$G_c^S = [A_c^S, \Gamma_c^S] \quad \text{where} \quad A_c^S = \{i_c, j_c\} \tag{29}$$

$$\text{with} \quad \Gamma_c^S(i_c) = j_c, \Gamma_c^S(j_c) = i_c \tag{30}$$

The network associated to passengers/freight flows carried by the company between the different airports is then given by:

$$R_{pax}^c = [G_c, [\varphi_{ij}^c]] \tag{31}$$

Here, φ_{ij}^c denotes the passengers/freight flows carried by airline c between airports i and j.

D. Passengers networks modeling

It is assumed that, when possible, an end user takes only one airline to perform an air journey. If this is not possible, it is considered that his journey is composed of several air trips done with an airline at a time, each trip feeding the demand of the corresponding airline. This analysis could be completed by introducing travel associations and agencies.

Let D_{ij}^c be the potential demand for airline c between two airports i and j. This potential demand represents the flow of passengers/freight that is expected to ask to airline c for air transportation between i and j during the reference year . To perform a trip between airports i and j, through the network operated by a given airline c, a end user may have the choice between a direct flight and a set of connected flights. Between each pair of airports i and j, a set of paths can be defined by:

$$Ch_{ij}^c = \bigcup_{p=1}^{p_{\max}} Ch_{ij}^{cp} \quad i \in A_c,\ j \in \Gamma_c(i) \tag{32}$$

where Ch_{ij}^{cp} represents the set of n_{ij}^{cp} paths going from i to j in G_c and which are composed of p successive flights.

When considering passengers flows, for comfort and safety reasons, it is considered that a traveler will seldom choose a flight with more than two connections. Then in this case, this consideration leads to choose $p_{\max} = 3$, which corresponds for instance to a flight from a regional airport to a central airport (a *hub*), then to a flight to another central airport (another *hub*) and finally to a flight towards a regional airport at destination.

In figure 5, the set of paths which connect airport 1 to airport 8 by flights from airline c_1 is given by:

$$Ch_{1,8}^1 = \{(1-2-3-8), (1-2-6-8), (1-5-6-8)\} \tag{33}$$

The paths of Ch_{ij}^c can be arranged by increasing length. Their number n_{ij}^c is given by relation (33):

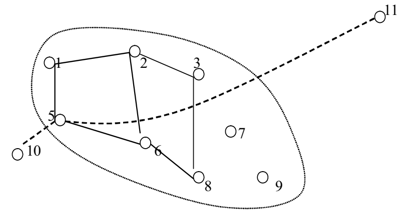


Figure 5 : Edges supporting the paths operated by airline c1 and linking airports 1 to 8

$$n_{ij}^c = \sum_{p=1}^{p_{\max}} n_{ij}^{cp} \tag{33}$$

Let Ch_{ijk}^c be k^{th} path of Ch_{ij}^c, its length is denoted $\tilde{l}_{i,j}^{ck}$ and in the case of $k = 1$:

$$\tilde{l}_{i,j}^{c1} = \delta_{i,j} \tag{34}$$

The flow of passengers who choose the k^{th} path between i and j is written φ_{ijk}^c.

A paths/arcs incidence matrix $[\alpha_c(i,j,k,h,l)]$ can be introduced as follows :

$$\alpha_c(i,j,k,h,l) = 1 \text{ if the arc } (h,l) \text{ of } G_c \text{ belongs to } Ch_{ijk}^c$$

$$\text{and} \quad \alpha_c(i,j,k,h,l) = 0 \quad \text{else.} \tag{36}$$

The flow of end users φ_{ij}^c on direct flights between airports i and j served by airline c are then such as:

$$\varphi_{ij}^c = \sum_{h \in A_c} \sum_{l \in A_c, l \neq h} \sum_{k=1}^{n_{ij}^c} \alpha_c(h,l,k,i,j)\, \varphi_{hlk}^c \quad i \in A_c, j \in \Gamma_c(i) \tag{37}$$

III. INTRA AND INTER ORGANIZATIONAL FLOWS CONSTRAINTS

In the next sub sections the different constraints relating the different flows and available capacities between the air traffic network, the airlines networks and the air transportation markets seen as users' networks, are introduced. These constraints represent consistency conditions which characterize feasible global amounts and spatial distributions of activities for the different actors of the air transportation system.

A. Passengers demand levels constraints

Total demand Dij between two airports i and j is distributed between competing airlines so that:

$$\sum_{c=1}^{C} \varepsilon_{ij}^c D_{ij}^c = D_{ij} \quad i \in \bigcup_{c=1}^{C} A_c, j \in \bigcup_{c=1}^{C} A_c \ et\ j \neq i \tag{38}$$

where $\varepsilon_{ij}^c = 1$ if $n_{ij}^c \geq 1$, $\varepsilon_{ij}^c = 0$ otherwise.

Then the demand to airline c for connection i-j is distributed along the different paths it provides:

$$\sum_{k=1}^{n_{ij}^c} \varphi_{hlk}^c = D_{hl}^c \quad h \in A_c, l \in A_c, l \neq h \tag{39}$$

The prediction of total demand D_{ij} between every pair of airports is basic for any planning or operations decision making problem, so many studies have been devoted to the development and improvement of o-d demand estimation tools.

B. Compatibility constraints between demand levels and available airlines capacity

The end users flows obey to capacity constraints relative to the capacity supplied by the fleets of the respective airlines:

$$0 \le \varphi_{ij}^{c} \le Q_{ij}^{c} \quad i \in A_C, j \in \Gamma_C(i) \tag{40}$$

with

$$Q_{ij}^{c} = Q_c^L \; f_{ij}^{cL} + Q_c^L \; f_{ij}^{cS} \tag{41}$$

if (i, j) is an edge of G_c^L and of G_c^S and,

$$Q_{ij}^{c} = Q_c^L \; f_{ij}^{cL} \tag{42}$$

if *(i, j)* is only an edge of G_c^L and finally:

$$Q_{ij}^{c} = Q_c^S \; f_{ij}^{cS} \tag{43}$$

if *(i, j)* is only an edge of G_c^S,

where Q_c^L and Q_c^S are respectively the capacities of the medium/long range aircraft and of the short/medium aircraft operated by airline c.

C. Consistency constraints between aircraft flows

It is assumed that during the considered period of time, the fleet composition of each airline remains constant so that flows conservation constraints over this time period must be introduced:

$$\sum_{j\in\Gamma_c^L}^{n} f_{ij}^{cL} = \sum_{k\in(\Gamma_c^L)^{-1}}^{n} f_{ki}^{cL} \qquad \forall i \in A_c^L \tag{44}$$

and

$$\sum_{j\in\Gamma_c^S}^{n} f_{ij}^{cS} = \sum_{k\in(\Gamma_c^S)^{-1}}^{n} f_{ki}^{cS} \qquad \forall i \in A_c^S \tag{45}$$

All considered flows in this study must take in theory integer values, this is particularly the case for aircraft flows, so that, integrity constraints such as :

$$f_{i,j}^{cL} \in IN \quad \forall i \in A_c^L, \forall j \in \Gamma_c^L(i) \tag{46}$$

and

$$f_{i,j}^{cS} \in IN \quad \forall i \in A_c^S, \forall j \in \Gamma_c^S(i) \tag{47}$$

could be introduced.

The fleet operation by airlines faces time availability constraints such as:

$$\sum_{i\in A_c^L} \sum_{j\in\Gamma_c^L(i)} f_{i,j}^{L}\theta_{i,j}^{L} \le \Theta_c^L . L_c \tag{48}$$

and

$$\sum_{i\in A_c^S} \sum_{j\in\Gamma_c^S(i)} f_{ij}^{S}\,\theta_{ij}^{S} \le \Theta_c^S . S_c \tag{49}$$

where $\theta_{i,j}^L$ and $\theta_{i,j}^S$ represent respectively the block times for medium/long range and short/medium range aircraft of airline *c* when performing a direct flight between airports *i* and *j*. Here Θ_c^L and Θ_c^S represent respectively the mean time availability for medium/long range and short/medium range aircraft of airline *c* during the whole time period. For a non local airline , those parameters represent the proportion of the total time availability of their aircraft assigned to the connection it operates. This is because their aircraft can be also used by the non local airline to perform external connections. Here, the relation between the mean duration of aircraft rotations and the levels of airport and air traffic congestion could be introduced.

D. Compatibility constraints between airlines demand levels and available ATC/ATM capacity

Total airlines activity levels at a given airport have to be in accordance with the number of available time slots in this airport. Then constraints such as :
for departures:

$$\sum_{c|i\in A_c^L} \Big(\sum_{j\in\Gamma_c^L(i)} f_{ij}^{cL} \Big) + \sum_{c|i\in A_c^S} \Big(\sum_{j\in\Gamma_c^S(i)} f_{ij}^{cS} \Big) \le K_i^d \qquad \forall i \in A \tag{50}$$

for arrivals:

$$\sum_{c|j\in A_c^L} \Big(\sum_{i\in\Gamma_c^{L-1}(j)} f_{ij}^{cL} \Big) + \sum_{c|j\in A_c^S} \Big(\sum_{i\in\Gamma_c^{S-1}(j)} f_{ij}^{cS} \Big) \le K_i^a \qquad \forall j \in A \tag{51}$$

Here, K_i^d and K_i^a represent respectively the number of slots allocated at an airport *i* to departures and arrivals.

With respect to the modeling of the capacity of the air traffic control sectors, two alternative choices can be made here :

1) either capacity is taken as a hard constraint for the traffic crossing in each sector,

or it is taken as a soft constraint and when traffic exceeds this capacity, it means that the resulting safety level is reduced in the concerned air traffic control sector.

It is a very complex matter to try to quantify with a single value the capacity of an air traffic control sector. However, the activity level of an air traffic control sector can be described with several indexes such as:

a) The total volume of traffic inside sector *t* and expressed in *Km* X *aircraft*, which is given by:

$$V_t = \sum_{\alpha\in U_T} (\Delta_\alpha^t \cdot \phi(\alpha)) \tag{52}$$

where $\phi(\alpha)$ the total flow of air traffic that goes through the arc α of G_T. It is given by relation:

$$\phi(\alpha) = \sum_{c\in C} \Big(\sum_{i\in A_c^L} \sum_{j\in\Gamma_c^L(i)} \omega_{ij}(\alpha) f_{ij}^{cL} + \sum_{i\in A_c^S} \sum_{j\in\Gamma_c^S(i)} \omega_{ij}(\alpha) f_{ij}^{cS} \Big) \tag{53}$$

b) The total volume of traffic crossing the limits of sector *t* , it is given by relation :

$$\Phi_t = \sum_{\alpha \in \Gamma_{CT}} p(t,\alpha)\,\phi(\alpha) \tag{54}$$

c) For each crossing point inside a sector, (airport or intersection of routes at way points), the volume of the different converging flows :

$$\forall j \in A_T \ \ and \ \ \tau(j) = t : \ \left\{ \phi(\alpha(i,j) \,\middle|\, i \in \Gamma_T^{-1}(j) \right\} \tag{55}$$

IV. A SYNTHETIC VIEW OF AIR TRANSPORTATION SYSTEMS

A. A multi agents organization

From the above it appears that the whole air transportation system can be perceived as a three level organization: the end users level (passengers and freight), the airlines (aircraft fleets) level and the air traffic management (ATC and airports) level. Many differences exists between these components and subsystems:

While end users are present in large numbers, even when aggregated by o-d pairs, airlines are few and ATC is (or should be) unique. And while end users are interested, while traveling, in maximizing their utility built up from a mix of comfort and costs considerations, airlines are involved in maximizing their profit by providing attractive offers to end users and ATC tries to guarantee global security levels to the generated air traffic flows while not impairing airlines operations efficiency.

The behavior of end users is highly variable, it depends widely on external factors such as short term economic conjuncture and political events, but tariffs applied by airlines and other concurrent transportation modes are also a strong determinant in their choice. Then end users demand levels present a notable degree of uncertainty while a feedback loop operates between demand levels and airlines tariffs.

The supply behavior of airlines is also reactive to external factors such as fuel costs, but since their activity immobilizes large capital amounts (mainly their fleet), their operations behavior is more stable than the short term fluctuating levels of demand and its easier answer to demand fluctuations is by short term adaptation of tariffs. The development of airlines activities goes along with current medium and long term demand predictions and with the foreseen evolution of their share of markets where competition is not limited to tariffs."

The dimensioning of ATC capacity is not affected by short term fluctuations in air traffic levels resulting from the airlines commercial activities.

Its development should anticipate medium to long term air traffic trends so that at a given time, available ATC capacity remains well above traffic demand levels. At that point, it could be considered that the financing sources for investment in this sector should be different depending if traffic demand is in an increasing phase or is already stabilized

B. Connecting activity levels and charges

There is no direct connection between end users and ATC operations: end users are reactive to the characteristics of air transportation services supplied by the airlines . To get one of the many transportation services offered by airlines they have to choose a service according to complex time tables and pricing structures. These pricing structures are dependent of the airlines market strategies and are such that airlines investments and operations costs, including ATC charges and cooperative avionics (CNS technology) equipments, are covered in an acceptable part.

It appears that, although complex, the relation between airlines levels of service and end users budgets is quite direct. It is not the case between ATC level of service and charges applied to airlines since ATC capacity is a concept difficult to be quantified and market laws are not straightforward in this case. However, ATC fees applied to airlines activities can have a significant influence over the end users demand levels since these extra charges will be *in fine* largely supported by the end users. This argument can be illustrated in two extreme cases:

-If national or international policy is to consider that growth of air transportation must be promoted, the incidence of ATC charges over air tickets tariffs should remain small.

-In the case where air traffic demand is at saturation levels for ATC capacity, ATC fees could be maintained high enough to stabilize on the medium term, end users demand levels.

Observe that aircraft technology improvements, such as enlarged capacity aircraft and minimized fuel consumption, may have a significant influence on the levels and spatial structure of activities and returns of an air transportation system.

V. CONCLUSIONS AND PERSPECTIVES

This paper proposes a modest but original contribution to air transportation studies. What is new here is the multilevel approach of the modeling of an air transportation system which allows to consider explicitly indirect relations between end users and high level service providers. No specific decision problem has been considered here but the proposed framework should be suitable to promote studies related either with physical flows issues, or prices and charges issues or even both. The displayed modeling approach makes use of basic mathematical concepts from Graph Theory and Algebra which are compatible with classical primal-dual activity and pricing optimization techniques. Refining the analysis of air transportation systems, more complex mathematical structures and tools could be helpful to optimize decision making in this field of application.

REFERENCES

[1] A. A Assad,.,(1978), “Multicommodity network flows-a survey”, Networks, 8, pp. 37-91.
[2] C. Barnart, N. Boland, L. Clark, E. Johnson, G. Nemhauser and R. Shenoi, “Flight string models for aircraft, fleeting and routing”, Transportation Science, 32, pp.208-220.
[3] C. Berge, (2001),”The theory of graphs”, New York, Dover.
[4] C. Berge and A. Ghouila-Houri, “Programming, games and transportation networks”, New York, John Wiley & Sons, Inc.
[5] D. J. Bertsimas and S. Stock, (1998), “The air traffic management problem with en route capacities”, operations research, 46, pp.406-422..
[6] L. R. Ford and D. R. Fulkerson, (1962), “Flows in networks”, Princeton, Princeton university Press.
[7] R. J. Wilson, (1972), “Introduction to graph theory”, New York, Academic Press..

A. Alou Oumarou, is native from Republic of Niger in West Africa, She obtained a BSc in Industrial Engineering (2002) and MSc in Applied Economics (2004) from Toulouse University I. She is presently a PhD student at MARTA/LARA a common Air Transportation laboratory between Toulouse University II and ENAC, the French Civil Aviation Institute. Her main fields of interest are pricing and game theory applied to Air Transportation.

Modeling ATM as a Complex System

Soufian Ben Amor
Complex Systems Modeling and Cognition Lab
Eurocontrol and EPHE Joint Research Lab
41 rue G. Lussac, F75005 Paris
Email :{sofiane.benamor@ephe.sorbonne.fr}

***Abstract*— In this research project we aim to apply a *holistic* approach to model the Air Traffic Management system (ATM). Precisely, our purpose is to investigate if percolation theory is efficient to model the ATM complexity and to understand how local availability/unavailability may affect the whole state of the system.**

***Index Terms*— Complex Systems, ATM, availability, percolation theory, diffusion theory, random graph theory, pretopology theory, phase transition.**

I. Introduction

The whole ATM system can be decomposed into three major interacting subsystems : the societal subsystem that contains the society constraints such as airlines, the technical subsystem that contains all technical infrastructures supporting the functioning of the ATM system, and human subsystem that includes all human components of the whole system, from flow managers to supervisors, pilot and controllers. ATM can be modelled as a set of components of different subsystems in mutual interactions in order to accomplish the mission of simultaneously maintaining safety and sustaining growth.

The ATM system can be defined as a complex system because its behavior depends on a complex combination of various sub-systems performing complicated functions. The evaluation of the impact of each function on the overall ATM system cannot be performed unless a specific approach is used. As complex systems need to be studied using a *systemic* approach, we aim to investigate if *percolation theory* which studies the general properties of random structures is a able to model efficiently ATM. Two basic aspects of the problem must be specified in order to perform a coherent modeling of the ATM behavior :

- proposing an adapted mathematical approach to model the ATM from a systemic point of view and define its complexity using the appropriate theories.
- abstracting the capacity concept to the availability concept.

Understanding the mechanisms by which complexity may be reduced in the particular domain of ATM may provide important solutions to optimize the system and its structure.

II. Related work

Currently the air traffic complexity definition concerns especially the Air Traffic Control (ATC) subsystem and it is based on the number of aircraft in a sector, or region of airspace. In 1998, Sridhar *et al.* showed that operational capacity of a sector depends on additional factors beyond the number of aircraft present in it [8]. Other researches have identified many factors that appear to influence the complexity of an air traffic situation but are not taken into account by the existing metrics. These factors include the distribution of aircraft in the sector, the shape of the sector, and the number of aircraft making vertical transitions.

The study of complexity in ATC has specifically focused on the identification of influent factors in an air traffic situation more or less complex. These factors are determined using basically two approaches :

- direct techniques using the results obtained by of verbal reports, questionnaires, and interviews.
- indirect techniques using statistical techniques analyzing controller judgements of the relative complexity of different air traffic situations.

Several metrics have been developed in response to the need for more objective and precise measures of complexity including sector size, sector shape, the configuration of airways, the location of airway intersections relative to sector boundaries, and the impact of restricted areas of airspace. A new air traffic complexity metric was introduced by Delahaye and Puechmorel in [4]. This metric is based on non-linear dynamical systems and expresses the complexity of an air traffic situation by identifying the organization of the aircraft trajectories and their clustering structure. It consists in computing the Lyapounov exponent corresponding to the non-linear dynamical system model that fits the real observations.

These complexity measures are very important to understand local behavior and dynamics of specific elements of the ATM system. However, these approaches lack of global view of the total dynamics of the system. Dynamics of ATM can not easily be deduced from the local behavior of its components. Thus, a top-down approach focusing on the properties and behavior of the whole system while integrating information provided by local studies will provide a better understanding of the underlying mechanisms in ATM complexity and efficient tools to optimize its functioning.

III. ATM complexity

Every organized human activity exhibits an opposition between tow basic requirements : the division of labor into different tasks to be performed, and the coordination of these tasks to accomplish a mission. The complexity of the system is related to the efficiency of its structure and dynamics in accomplishing the global mission. According to the New England Complex System Institute, complex systems *is a new field of science studying how parts of a system give rise to the collective behavior of the system, and how the system interacts with its environment.* [1]

This definition can be used to describe the ATM system and induces the fact that it should be studied using the mathematical techniques dealing with criticality and phase transition phenomena which are the main properties of complex systems and express the network effect phenomenon. The phase transition is an abrupt change in the behaviour (or the state) of the system which occurs around a critical value of a key parameter called transition threshold.

The ATM system is a complex network composed of several heterogeneous and mutually interacting subsystems. The complexity of the ATM can be related to the following factors :

- System size
- Diversity of users
- Safety constraints
- Uncertainty (weather, human factor, technical factor...)

This complexity can be also related to the Air Traffic Control subsystem representing the rigidly structured airspace and the largely centralised, human operated control hierarchy. For this reason, new approaches based on a more flexible organization of airspace are investigated by the aviation community. One of the interesting approaches is based on the concept of *Free Flight* defined by the Federal Aviation Aadministration[2] as

> *a distributed system that allows pilots, whenever practical, to choose their own route and file a flight plan that follows the most efficient and economical route.*

However this approach implies the use of precise algorithms to generate safe trajectories in order to maintain safe separation between aircrafts and to avoid the effect of hazardous weather. In [1], an *ant colony optimization based weather avoidance algorithm* is described. It generates optimal weather avoidance trajectories in Free Flight airspace. As in computer science, distributed systems are, for different applications, more powerful than centralized schemes. Our approach may provide a theoretical framework to compare the efficiency of different air navigation conceptions and to show if a distributed approach is better than a centralized one and in which conditions.

A. Objective of the research

Our objective in this research project is to investigate the influence of the structure and dynamics on the ATM network in order to understand how local situations (availability or unavailability of the components) influences the total state of the system. Two important and dependent aspects are to take into account : the structural complexity, which is related to the topological properties of the graph modeling the network and the dynamical complexity that reflects the evolution of the availability of the system according to the availability of its components. The relationship between the structural complexity and dynamical complexity is very intricate : the structure properties influence the dynamical parameters and dynamics may change the topology of the structure. As we can see in figure 1, the initial demand of companies exceeds the declared capacity of a sector at a specific moments. The Central Flow Management Unit (CFMU), an operational unit of EUROCONTROL, manages the air traffic in order to avoid the congestion due to this difference and to optimize the utilization of resources. Despite the regulation and the planning made by the CFMU there are also differences between planned and real traffic. These differences are specific to each sector and the resulting effect of these local states and their interactions between them on the whole availability of airspace are difficult to determine.

We think that by considering the densely interconnected system of ATM as a network (that may be a hierarchical network) where components properties are heterogeneous and individual and by applying appropriate theories we are able to model the emergence of global properties in the system from the local behavior of its component (typically the availability and congestion of the airspace). In this way we can take into account the coordination requirements representing the interactions between controllers in adjoining sectors, which is an important factor in ATC complexity [8]. It is also important to note that these interactions are closely dependent of the airspace design. In fact the topological structure of airspace defines the structure of the sectors coordination network. The space-time analysis that we propose in this research project is a general approach focusing on the intricate relation between these tow fundamental aspects.

B. Hypthesis

1) Structural complexity: Most of "real world" networks can be considered as complex from the topological point of view. They have structural properties that do not exist in random graphs, such as a heterogeneous degree distribution, a high clustering coefficient, assortativity or disassortativity among vertices, community structure at many scales and a hierarchical structure. The two most well-known examples of complex networks are those of *scale-free networks* and *small-world networks*. Both are specific models of complex networks discovered in the late 1990s by physicists.

Scale-free networks: A scale-free network is a network with a power law degree distribution :

$$P(k) \approx k^{-\alpha}$$

These networks have no characteristic scale because of the heterogeneity of the vertices degrees : a few number of vertices called "hubs" have a high degree and the rest of the vertices have a low degree (figure 2.b, 2.d, 2.f). This characteristic distribution is observed in many real world networks such as the World Wide Web, the network of Internet routers, protein interaction networks, email networks, social networks etc.

Random graphs are more adapted to model regular networks having a homogeneous degree distribution : all the nodes have

[1] http://necsi.org
[2] http://www.faa.gov/freeflight/

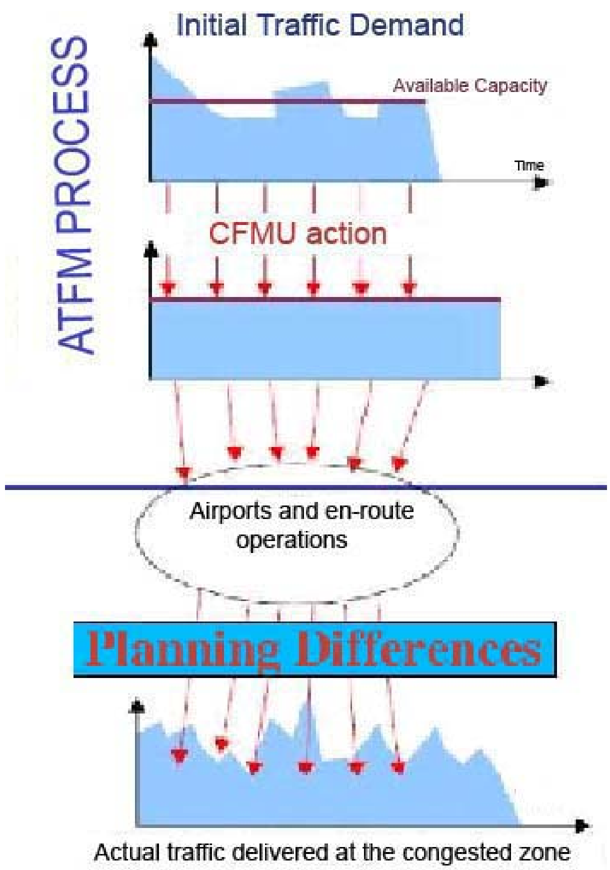


Fig. 1. Real *versus* planned traffic in a congested zone. Source: [6]

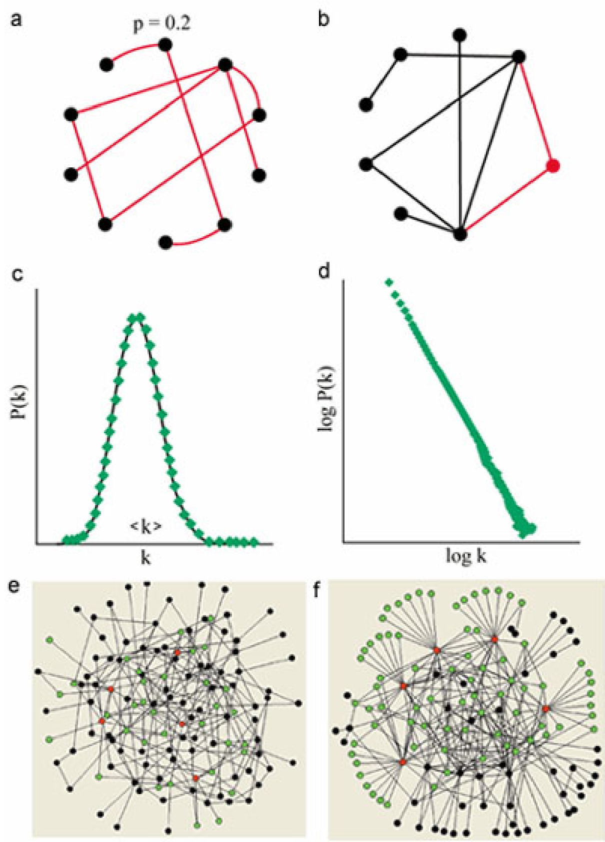


Fig. 2. Scale-free networks *versus* random networks.

almost the same degree (figure 2.a, 2.c, 2.e). In fact, the degree distribution in a random graph follows a Poisson law :

$$P(k) \approx e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

This structural difference (figure 3) have an important effect on the dynamics in the network. For example, a random node failures have very little effect on a scale-free network's connectivity, but deliberate attacks on the "hubs" are able to dismantle easily the whole network. In the air transportation context, free-scale networks are present in different subsystems. For example, the american air transportation network have a free-scale structure (figure 4).

Thus, we assume basically that ATM system have a complex structure if it is a scale-free network.

2) Dynamical complexity: As we can see random graph theory deals specifically with the topological properties of networks. Concerning the dynamical aspect, we assume that percolation theory is able to characterize network effects in the ATM system from a temporal point of view.

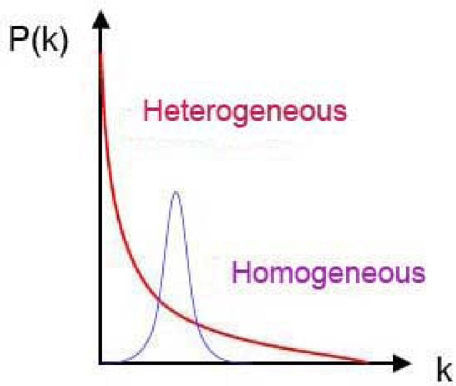


Fig. 3. Homogeneous and heterogeneous distributions.

C. Theoretical framework

As complex systems present generally a phase transition phenomenon according to a certain property and their global behavior can not be easily deduced from the behavior of their components, it is important to study them using the appropriate

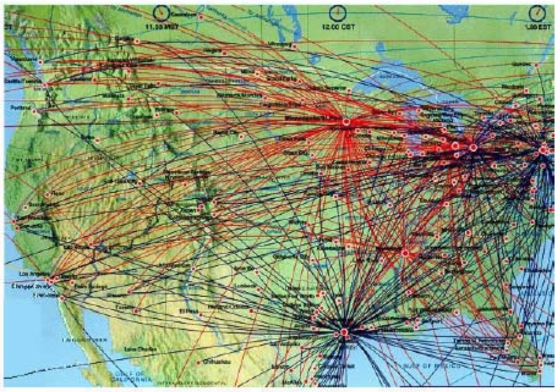


Fig. 4. Free-scale structure of the american air transportation network.

theoretical tools. Among theories using a *holistic* (or *systemic*) approach to explain the passage from the individual to the collective, from the micro to the macro, percolation theory seems to be an adapted theoretical framework in the ATM context : it studies the deterministic propagation of a fluid (or an information) on a random medium (or structure).

This theory was successfully applied to model complex systems in Statistical Physics, Economy and recently in social models studies. We can compute the critical thresholds and study the evolution of the system related to the variation of characteristic quantities such as clustering coefficient, the average size of the clusters and its distribution using the mathematical methods developed in percolation theory. Some interesting methods used in *diffusion theory* may be associated to percolation based modeling to enhance the quality of this holistic approach in complex system behavior analysis. Topological properties of the network could be determined using *random graph theory*.

IV. Conclusion

From a mathematical point of view, percolation theory is interesting because it exhibits relations with random graph theory dealing structural properties of graphs. This makes easier the understanding of the intricate relation between the structure and dynamics in ATM. However, percolation theory is not enough to express particular aspects in ATM (for example the airport effects) because it is based on the notion of neighborhood which is not well defined and very restrictive in expressing qualitative or heterogeneous connections between the different elements.

In order to solve this problem, we have the intention to generalize percolation processes using *pretopology theory* [2]. It is a general formalism that expresses different types of connections that may exist between the components of a system. This step allows us to formalize the concept of neighborhood and represent more realistic relations between the ATM components.

References

[1] Alam S., Abbass, H., Barlow, M., Lindsay P. Mapping lessons from ants to free flight: An ant-based weather avoidance algorithm in free flight airspace The international Society for Optical Engineering (SPIE) Microelectronics, MEMS and Nanotechnology Symposium, 2005, 9.

[2] Z. Belmandt. *Manuel de prétopologie et ses applications*. Hermès, Paris, 1993.

[3] S. Conway. Systemic Analysis Approaches for Air Transportation. *In Proceedings CSER 2005*, Hoboken, NJ, USA.

[4] D. Delahaye and S. Puechmorel. Air Traffic Complexity Map based on Non Linear Dynamical Systems. *In* 4^{th} *EUROCONTROL Innovative Research Workshop & Exhibition*, December 2005, Brétigny sur Orge.

[5] G. Grimmett. *Percolation*. Springer-Verlag, Berlin, 1999.

[6] C. P. Gwiggner. Implicit Structures in ATFM : Thesis overview 2005. *In* 4^{th} *EUROCONTROL Innovative Research Workshop & Exhibition*, December 2005, Brétigny sur Orge.

[7] K. E. Haynes , R. R. Stough, R. G. Kulkarni. Towards a percolation model of accessibility. *In Proceedings of GeoComputation*, Jul. 1999.

[8] J. M. Histon and R. J. Hansman *The Impact of Structure on Cognitive Complexity in Air Traffic Control*. Report No. ICAT-2002-4June 2002, MIT International Center for Air Transportation, Cambridge, United States.

[9] J. Spencer. *Ten Lectures on the Probabilistic Method*. SIAM, Philadelphia, 1987.

[10] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. Taylor & Francis, London, 1992.

[11] B. P. Zeigler, H. Praehofer, T. G. Kim, *Theory of Modeling and Simulation*, 2nd Edition, Academic Press, 2000.

A Fuzzy Logic Modeling Approach for the Assessment of Aircraft Evacuation Procedures: A Proposal

Minesh Poudel, Félix Mora-Camino, Marc.de Coligny

Abstract— **Aircraft evacuation effectiveness is a critical issue for passenger's safety with large aircraft. New aircraft must comply with maximum exit time limits. To validate the aircraft evacuation system design and procedures, demonstration according to many different scenarios have been performed in the past. But more recently numerical simulation tools have been developed to avoid the limitations inherent to real demonstrations. One of the main difficulties encountered by numerical simulation of evacuation operation is related with the modeling of passenger's behavior, which is influenced by a complex mixture of socio-psychological and physical factors. This is the main point discussed in this paper.**

Index Terms— **aircraft evacuation, fuzzy logic, human factors, simulation, stochastic modeling..**

I. INTRODUCTION

Aircraft evacuation effectiveness is a critical issue for passengers safety with large aircraft. New aircraft must comply with maximum exit time limits. To validate the aircraft evacuation system design and procedures, demonstration according to many different scenarios have been performed in the past. But more recently numerical simulation tools have been developed to avoid the limitations inherent to real demonstrations. One of the main difficulties encountered by numerical simulation of evacuation operation is related with the modelling of passengers behaviour which is influenced by a complex mixture of socio-psychological and physical factors.

In this paper an approach based on fuzzy modelling of passengers behaviour is proposed to tackle this problem. Decision points are defined for each evacuating passenger according to the position of the available exit doors and to the location of danger (fire in general). Then at each of these points the passenger is supposed to perform a fuzzy evaluation of his remaining egress opportunities so as to modify if necessary his choice and behaviour. The proposed fuzzy evaluation model is discussed and a preliminary evacuation model is built up to validate the proposed simulation approach. The performed case study considers a double aisle aircraft.

II. THE NEED FOR AIRCRAFT EVACUATION MODELLING

Public demand for air travel has increased steadily over the last two decades. Industry expects further substantial growth into the next century. Forecasts for the doubling of air traffic over the next decade have led airframe manufacturers to start designs for airframes carrying as many as 800 or 1000 passengers, which has already realized. While passenger safety in air travel has always been a high priority within the industry these developments demand that new attention is given to the safety of air travel. Up until 1970, the fatal air accident rate worldwide had been falling dramatically. While the rate, since then, has decreased slightly, this has not been accompanied by an equivalent reduction in the fatality rate of those onboard aircraft. For this reason not only the issues concerned with the prevention of the occurrence of accidents, but also issues related to improving the survival rate in the event of an accident, will have major importance in the years to come. Survivability of passengers and crew in aircraft accident depends upon the global improvements on following three major aspects: impact protection, fire survivability, and aircraft evacuation.

In this study we consider one of the important aspects related to the survivability of the passengers in case of accident: How fast, quick, efficient and effective evacuation can be carried out plays a vital role for those who survive the initial crash.Current international regulations require that an aircraft must be evacuated within 90 seconds with half of the exits operational and this is tested during aircraft certification. However, in real life crashes evacuations often do not run as smoothly as necessary and require more time. There are many factors, which influence evacuation time and survivability in aircraft accidents. These factors can be broadly classified into four groups: *configuration, environment, procedures and Behavior.* Of the numerous elements of an aircraft's emergency evacuation system, exit and slide design, flight attendant training, and full-scale evacuation demonstrations required for type certification have engendered the most attention and public debate. And the major difficulties with conducting research into the survival of aircraft accidents is the introduction of sufficient realism in the process, without putting participants at serious risk of physical or psychological harm. Understanding behavior of the passenger in highly stressful and disorienting conditions (Actual accident) is one of the important factors to improve the probability of a successful evacuation from an aircraft. Unfortunately, as yet, very limited research effort has centered on the impact of passenger behavior on aircraft emergencies and evacuation. However, the information from other disaster situations like building or ship evacuations and along with report from survivor from the aircraft accidents has been developed to build up representation of the behaviors within cabin during the emergency situations. Significant differences are found on the behavior of the passenger during different emergency situations, sometimes they tries to flee out of the plane in an orderly manner whereas in other cases the orderly manner breaks down and confusion in the cabin leads to blockages in the aisles and at exits. In some accidents, all of the passengers assume that the objective is to get everyone out of the aircraft as quickly as possible, and they therefore all work collaboratively. In other emergencies, however, where an immediate threat to life is perceived, the main objective will be survival for themselves, and in some instances, members of their family instead of all passengers being motivated to help each other, In this situation when the primary survival instinct takes over, people do not work collaboratively. The evacuation can become very disorganized, with some individuals competing to get through the exits. Different ways individual passenger respond to emergency which may be fear, anxiety, disorientation, depersonalisation, panic, behavioural inaction and affiliative behaviour.

III. COMPUTER MODELING ISSUES FOR AIRCRAFT EVACUATION

Proposals are been made using of computer programs to simulate the dynamics of emergency evacuations as a method of satisfying evacuation demonstration requirements. The development of aircraft evacuation models is to augment or replace the current certification process. Computer modeling attempts to integrate the complicated interactions of passengers and their individual behaviors with the physical attributes of the airplane cabin. Sets of algorithms are used to impose "characteristics" (age, mobility, gender, and personality) which affect their movement within the cabin. Physical attributes of the cabin such as seat pitch, aisle width, exit size and availability, smoke, fire, and other characteristics that influence the passengers' movements can be included. Any or all of these variables, if data are available, can be varied to examine their effects on the evacuation. The introduction of the computer-based techniques to simulate evacuations offers several advantages over full-scale demonstrations: design and development of safer aircraft bringing safety matters to the design phase while the proposed aircraft is still on the drawing board, implementation of safer and more rigorous certification criteria, development of improved and more efficient crew procedures, improved cabin crew training, accident reconstitution.

However, computer modeling is not recognized by the regulatory body as an allowable method of demonstrating evacuation capability of airplanes. Although it is generally accepted by industry that computer modeling will have a role in evacuation certification in the future, more traditional methods will continue to be used until the models are validated. There are currently two main areas of application for aircraft evacuation models, respectively for either design/certification and for accident reconstruction.

The performance based requirements known as 90 seconds certification test, and the compliance with this rule is done by performing full-scale evacuation demonstration. This test has to be done with representative of all genders, in darkness and utilizing only half of the normally available exit. Crewmembers don't know the exact exit, which will be functional during the test. The test involves evacuating all the passengers within 90 seconds to pass the performance standard, though this performance test is not intended to predict the performance of the aircraft under realistic accident scenario.

Evacuation modelling for accident reconstruction is another major areas where very little research has been carried out. Few of the models have been developed in an attempt to simulate real emergency evacuation scenarios. Modelling real emergency evacuation is much more difficult and complex than certification modelling for the number of reasons. Firstly, the real case emergencies leads to myriad of different possible evacuation scenarios thus the range of human behaviour that needs to be modelled is far more extensive then that found in the certification scenarios. Secondly, reliable data on human behaviour and performance under these realistic accident scenarios is more difficult to obtain, where there exits very little sources of accurate quantitative infomation on human performance in emergency evacuation situations. This information is limited to the testimonies of surviving passengers, crew, and rescue workers and data from contrived experimental trials.

IV. AIRCRAFT EVACUATION MODELS

Few of the aircraft evacuation models are developed in the past: (1970 to 1980) General Purpose Simulation System (GPSS) developed by the FAA, (1987 to 1992) Gourary Associates (GA) model developed by Gourary Associates, (1990 to 1994) AIREVAC/ARCEVAC developed by Aviation Research Co-operation, (1994 to 1996) Macey's Risk Assessment Model developed by Cranfield University, (1996 to 1996) The Oklahoma Object Oriented (OOO) model, (1989 to now) EXODUS developed by the fire safety engineering group of the university of Greenwich,(2001 to now) Robbin's Discrete Element Method (DEM) developed by Department of Mathematics at the University of Strathclyde.

The simulation of the evacuation fall into two categories of model, those that only consider human movement and those which attempt to link movement with behaviour. The first category of model concentrates solely on the carrying capacity of the structure and its various components. Here, each individual are treated as unthinking objects that automatically respond to external stimuli and people are assumed to evaluate the structure, immediately ceasing any other activity. The direction and speed of egress is determined by physical considerations only (e.g. population densities, exit capacities, etc.). The second category of models takes into account not only the physical characteristics of the enclosure but treats the individual as an active agent taking into consideration his response to stimuli such as various fire hazards and individual behaviour such as personal reaction times, exit performance etc. Within the modelling methodologies adopted there are also a number of ways to represent the enclosure, population and the behaviour of the population. Models that simulate evacuation tackle broadly by three fundamentally different manners: Optimisation, simulation, and risk assessment. Models which assume that the occupants evaluate in as efficient a manner as possible, ignoring peripheral and non-evacuation activities, where evacuation paths taken are considered optima, as are the flow characteristics of people and exits are termed as Optimisation model. These tend to be models which cater for a large number of people or who treat the occupants as a homogenous ensemble, therefore not referring individual behaviour. Some models on the other hand tries to represents behaviour and movement observed in evacuations, not only to achieve accurate results, but to realistically represent the paths and decisions taken during an evacuation, which are termed as Simulation models. The behavioural sophistication employed by these models varies greatly, as does the accuracy of their results. These types of models may be used to predict and / or reconstruct realistic evacuation scenarios. They can also be used to determine layouts of configurations that are most conducive for rapid evacuation or be used to determine optimal evacuation procedures. The third category of models is Risk Assessment models, which attempt to identify hazards associated with evacuation resulting from a fire or related incident and attempt to quantify risk. By performing many repeated runs, statically significant variations associated with changes to the compartment designs or fires protection measures, could be assessed. But in all these approaches, it is essential that the enclosure geometry, population and population behaviour be represented, which can be modelled several approaches.

V. EXISTING MODELING APPROACHES

To represent the decision-making process employed by the occupants, the model must incorporate an appropriate method for determining occupant behavior. The behavioral perspective adopted is influenced by the population and geometry approaches taken, and as such is the most complex of all the defining aspects. Using current modeling techniques there are five commonly used approaches to represent behavior within evacuation models; Functional Analogy Behavior, Implicit Behavior, Rule based behavior, Artificial Intelligence based behavior and no behavioral component. Functional analogy models apply an equation, or set of equations, derived from a non-evacuation related discipline, to the entire population, which then completely governs the population's response. Although it is possible for the population to be defined individually in these models, all the individuals will be affected in the same way by this function, and therefore will react in a deterministic manner to its influences, undermining individual behavior. The function used to describe the occupant behavior is not necessarily derived from real-life occupant behavior, but is instead taken from another field of study that is assumed to be analogous to human behavior (e.g. the functions which drive the Magnetic model were taken from Physics). Occupant movement and behavior is then completely determined by this function, which may or may not have been previously calibrated with human movement. Some evacuation models have No behavioral rules at all. These models simply simulate the physical movement of the people. Thus, peoples' decisions are formed on the basis of physical influences, rather than through simulating more complex human decision processes. They do not represent the behavior explicitly but make use of secondary data to represent their affects.These types of models are highly dependant upon the availability, reliability and validity of the data used. Models, which explicitly recognize the behavioral traits of individual occupants, usually make use of a Rule based system.These models have a set of rules, or heuristics, that govern the behavior of simulated people within the model. This allows for decisions to be taken by occupants, according to pre-defined sets of rules. These rules can be triggered in specific circumstances, and in such circumstances, have an effect. A problem with this style of decision-making process is that in simplistic methods the same decisions are taken under the same circumstances, in a deterministic fashion. Deterministic models are open to criticism, as they do not represent natural variation inherent within a scenario. This has the disadvantage of denying the possibility of natural variations in outcomes through repetition. Most of the rule based models overcome this problem by introducing a stochastic component to the decision making process. Artificial intelligence based models utilize methods from artificial intelligence to mimic human intelligence in simulated people. Whilst this approach can yield realistic behavior the level of user control is somewhat reduced.

VI. THE PROPOSED APPROACH

The proposed modeling approach treats passengers at the microscopic level and describes individually their egress behaviour. Considering that the egress behaviour of a passenger is continuous in time between discrete decision-making instants, which occur asynchronously, a discrete time approach with a very short period (about one second) is chosen to process time. Then according to the current situation of the passenger, a decision situation can be identified at which he has to reconsider his personal egress plan.

According to a probabilistic scenario, behavioural parameters are attached to each passenger. These parameters are mainly: situation awareness capability, physical state (dead, wounded, shocked, sane), mean response time to a new situation, free individual egress speed, group links.

A dynamic disaster scenario is proposed, it involves the initial state of the cabin (nature of the disaster, extension of destruction, spatial distribution of surviving passengers and crew, available exits) and the evolution of its state (fire and smoke propagation, etc) as well as the availability of the remaining flight crew staff. Then, according with the current disaster situation which is a result of the chosen scenario, to the cabin nominal organization is associated a background capacitated circulation network representing all remaining available ways for passengers and crews.

Then, based on this background circulation network, for each passenger or crew is built a perceived egress network towards the different available exits. To each path linking the current passenger position to an available exit, is attached a fuzzy impedance which is the result of a fuzzy inference evaluation process which takes into account: Passenger parameters such as situation awareness capability, physical state and group effect, the current state of the cabin with the danger areas, the effect of crew egress directives, the traffic situation including size of queues and current egress speed along that path, the situation of the other members of the passenger's group.

Among all these possible paths, the passenger has to choose a provisory egress path that he will follow during the next time interval. A next decision point is defined along that path at which the passenger can make a new path choice.

According to the nature and the extension of the disaster, at a given point, a passenger can be injured or killed.

A dynamic scheduler defines in which order each passenger or crew will be processed at a given instant. This scheduler is highly reactive since it takes into account the position of each passenger in his current egress path.

Then, different performance indexes, designed according to the regulations constraints, are computed on line during the simulation runs and will be a basis for model calibration and validation.

VII. CONCLUSION

A In this paper an approach based on fuzzy modelling of passengers behaviour is proposed to tackle this problem. Decision points are defined for each evacuating passenger according to the position of the available exit doors and to the location of danger (fire in general). Then at each of these points the passenger is supposed to perform a fuzzy evaluation of his remaining egress opportunities so as to modify if necessary his choice and behaviour. The proposed fuzzy evaluation model has been discussed. A preliminary evacuation model is actually being built while an entire new validation methodology must be designed to tackle efficiently the behavioural aspects of the proposed simulation model.

REFERENCES

[1] Cherry R.G.W. and Associates Ltd; Analysis of factors influencing the survivability of passengers in aircraft accidents, Vol. I-III. EU DG VII contract, 1995.

[2] Prof. Helen Muir, Dr Michel Piers, Mr Frank Taylor; Increasing the survival rate in aircraft accidents. European Transport Safety Council,1996.

[3] R.L Wilson, L.J Thomas, H.C Muir, E.R.Galea, S.J.Blake, A.J.P. Dixon, S.Gwynne, S. Joseph, L. Moulin and E. Andlauer; VLTA Emergency Requirements research evacuation study;2003.

[4] Aircraft Evacuation Testing: Research and Technology Issues; September OTA-BP-SET-121 NTIS Order#PB94-107620; 1993

[5] Emergency Evacuation of Commercial Airplanes; NTSB Safety Study 7266 PB2000-917002 NTSB/SS-00/01.

[6] Survivability of Accidents Involving Part 121 U.S. Air Carrier Operations; NTSB/SR-01/01 March 2001.

[7] Snow C., Carrole M.J. and Allgood M.A); *Survival in emergency escape from passenger aircraft.* Report No. FAA AM 7016. Washington: Civil Aeromedical Institute;1970. 1988.

An Initial European Small Aircraft Prediction Model for 2020

Daniel Rohacs, PhD student
Eurocontrol Experimental Centre &
Ecole Pratique des Hautes Etudes &
Budapest University of Technology and Economics

Abstract— **The further globalization and mobility freedom indicate rapid increase in needs of personal used air transport systems (see NASA SATS program). Such new and further general aviation (GA) requires totally new, innovative ideas and the use of the newest technological achievements to ensure a new, inexpensive, and secure control system.**

The goal of this paper is the development of a European small aircraft prediction model that could be applied for the analysis of the impact of small aircraft (SA) activity on the current ATM parameters. Therefore, this paper analysis the current European small aircraft's traffic (using EUROCONTROL database) and different aspects (development of the economy, changes in market needs, application of new technologies, etc.) having influences on air transportation system in general and especially on the European ATM/ATC. It also deals with the possible model structures, model classes and finally gives a relatively simple model based on Markov chain that has been chosen for testing it in simulation. The simulation results show, that the developed model could be used in preliminary prediction of the European small aircraft activity and in the definition of required changes in regulation, management and service providers to ensure the further development of SA.

Index terms— **general aviation, prediction methods, small aircraft, traffic characteristics**

I. INTRODUCTION

Today's air traffic volume is projected to be doubled by 2020 [1,2]. On the other hand, today there are approximately [3],[4] 300 000 private and small aircraft pilots in Europe who fly more then 60 000 small aircraft. As the technology is already available to establish a safe, economical and environment friendly new air transportation system based on personal air transportation purposes, their traveling habits might change. Maximum satisfaction of requirements could take over the lead, such as traveling on demand, point-to-point and in a more flexible way [6]. Additionally, from our perspective, small aircraft (SA) has to be designed to be affordable to anybody, without any special or extra knowledge and abilities (similarly to the NASA SATS project [2],[6]). As this future state could enhance people's quality of life and ensures their freedom, aviation should use radically new, innovative ideas in order to break down the currently existing limits and constraints in air transportation.

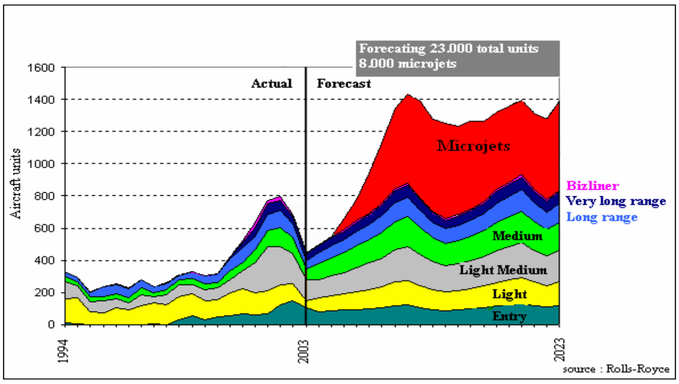


Fig. 1.1. Increasing number of microjets in the near future.

With accordance to our investigation, this market - in new democratic countries and in new members of European Union (EU) - will develop rapidly, with a significant increase in terms of small aircraft. This development will continuously be very high over the next 20 years [5], even higher, than conventional air traffic as we can see it in the figure 1.1. . However here microjet means an aircraft with a maximum take-off weight between 5000 and 10000 lbs (such as Honda Jet, Avocet, Safire, etc.), from our perspective it could also be understood as a small aircraft (coming from the practice of NASA SATS program [2],[6]), a personal aircraft (for personal air transportation purposes [4],[7]) an air taxi [8], or even a UAV.

In the face of this increase, the current system is already reaching its limits, thus most probably it will not be able to meet that future need of tomorrow's capacity (in several domains such as environmental consideration, security, safety, etc. [2]). Moreover, the saturated fact of today's big airports could oblige personal aircraft flights to use regional airports, or other underutilized landing possibilities, which in most of the cases are not equipped with modern radio-location systems for controlling the air traffic. Thus, in the aim of decreasing such and other negative effects - caused by the increased number of small aircraft - current ATC/ATM, airport and airline operators might call for a change. Therefore we have to develop radically new concepts to be able to accomplish future tasks (such as pilot workload monitoring, simplified control [4], automatic control system, etc.), and to minimize the interaction between SA and the conventional air traffic. Hopefully, such change in air transportation of the future has a perfect timing [1], because we still have a little time to prepare ourselves for the work ahead of us, before air transportation will reach the record levels.

For a better understanding of the work to be accomplished in the domain of SA activity, initially we discovered the following related works. First of all, the NASA's Small Aircraft Transportation System (SATS) [2], [6] that desire to expand the use of small airports and small aircraft for public transportation, through cockpit development and some advanced operational concepts in non-radar airspace at non-towered airports. Or the Personal Air Transportation System (PATS) [4], [7] project that is very similar to SATS, however here, investigations are done in the domain of aircraft maneuverability simplification via automation, without focusing on cockpit development. Another thought provoking idea is the SkyCar [13], a vehicle capable of vertical take-off and landing that looks like a cross between a sports car and a tiny jet aircraft. And finally the UK JETPOD [8], which is a European pre-designed study for advanced personal air transportation.

Generally, the purpose of these programs is to make small aircraft as easy and safe to operate as cars, with a cost the same as a mid-range car. However these have been very useful to name the currently available small aircraft, cockpit development benefits and requirements vis-à-vis pilot experiments, their limitation (expect the UK JETPOD) is the focus on the American market. As Europe consists of several countries, with different social and economical characteristics, the importance of our study is to adopt their results and predictions on the European market attributes. Additionally, expect the last few years of development in this field, small aircraft development has not been in focus for over 40 years, since the World War II. The companies producing these aircraft have partly changed their activity, or jointed to a bigger civil and military aircraft development. Due to this past, we do not have enough statistical data and experience for describing the future growth of small aircraft activity, especially for the European market.

In order to deal with this past and to ensure future needs described above, the impact of small aircraft on the European air transportation system has to be analyzed. Among several areas, our research deals with air traffic control, air traffic management domain. Therefore the analysis of the European ATM parameters, from a small aircraft point of view.

As the analysis of small aircraft activity is a pioneer task in Europe, the initial objective is to obtain relevant information and a wide knowledge on its development. In order to define its main characteristics and trends, we firstly should analyze the current air traffic situation, by taking into account the flights that are the most close to a small aircraft of our objective. This approach might restrict the area to be focused on and give us a first idea, how SA might interact with conventional air traffic in Europe.

Using the results of the traffic analysis, our objective is also to forecast/analyze the impact of small aircraft on different fields/parameters of the European ATM. That demands a SA prediction model, which could forecast any SA traffic load on the current system parameters. Using the model and by running a simulation, the definition of the bottleneck - among the ATM parameters that might call for further investigation - could be named.

Once these objectives have been accomplished, further investigations might focus on tasks, to minimize the impact of small aircraft on the air transportation system. This could be imagined by giving proposals, solutions for a particular domain, and shifting the ATM attributes towards a system that could give solutions for future requirements of small aircraft purposes.

II. CURRENT EUROPEAN SMALL AIRCRAFT RESEMBLING AIR TRAFFIC ANALYSIS

In Europe, the air traffic analyze with small aircraft in head is a pioneer task. Thus, our investigation only covers a nine-month period (in 2004) by taking into consideration one day from the weekends and weekdays, from EUROCONTROL's CFMU database. Using our SA definition, this allowed us to analyze more than 50.000 European flights; to recognize the impact of seasons; and finally to define, whether business or leisure flights are more often to happen.

The result and the distribution of the number of flights for the whole examination period is shown in the figure 2.1. . It could be concluded, that the number of flights that take place during the weekdays is nearly twice more important, that is for the weekends, even so the average number of flights for the whole period is 1429. This could mean that the distribution of flights represents more a business market segment, where the

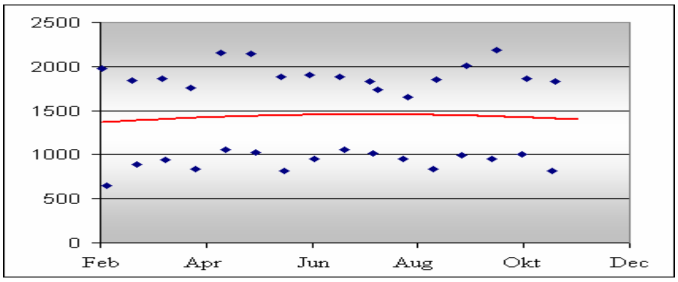


Fig. 2.1. Small aircraft flights distribution for weekdays (upper line), and weekends (lower line). Red – polynomial trend line

flexibility of passengers is underlined.

In order to be able to place, and to assess how these flights are going to penetrate to the airspace, the evaluation of the flight level density is crucial. The study takes into consideration the cruising path that is visible in the figure 2.2. . As we might observe, most of the flights take place around FL 100, and just a very few percent of them (18%) are flying at least up to FL 200. The reason for such a low altitude could be due to propulsion technology, and / or flight distance

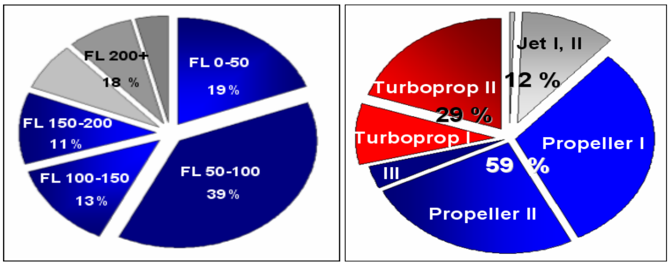


Fig. 2.2. Small aircraft flight altitude (on the left) and propulsion technology preference (on the right).

preference, which calls for further investigations. Firstly, we analyzed the current small aircraft propulsion technology preference. The figure 2.2. shows, that propellered technology is in majority with nearly 59% of the flights, which is followed by turboprop (29%) and jets (12%). We ended up with nearly the same result by analyzing the 15 mostly preferred small aircraft, where propellered technology preference becomes even more clear.

After the analysis of the flight altitude and propulsion technology preference, we already had an idea about short flight distances. To demonstrate it, we accomplish the flight distance analysis such as follows.

At the same time, as CFMU database does not contain flight distance data, we used GPS coordinates and great circle distance calculation to support the computations. Finally, the figure 2.3. shows that the most frequent flight distance

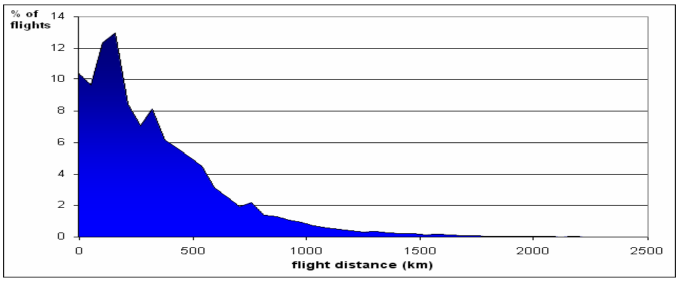


Fig. 2.3. Flight distance distribution of small aircraft flights.

belongs to 150 km, which is already 12% of the total small aircraft flights. Then 85% of them are not longer than 500 km, and just 3% (!) are more than 1000 km, which gives an average of 310 km. Thus we could conclude, that currently short flights at low altitudes are more often to happen, with the preference of a propellered small aircraft. To understand what could be behind such a distribution, and how that traffic could be look like, we represented one typical day of small aircraft flights in Europe, using COSAAC Software (see figure 2.4.). There, the flights are very different one from the other, and it would be difficult to define preferable routes, or city pairs as it might be possible for conventional air traffic. One exception is the traffic between England and some of its islands (such as Alderney, Jersey, Guernsey) where even scheduled flights could be found several times a day, which shifts them to the top of the most frequent city pairs list. As for city pairs, in total we found more than 12.000 for the whole examination period, where just half of them occurred at least once, and just 4,3% took place at least once in a week. Moreover, some of these city pairs have the same origin and

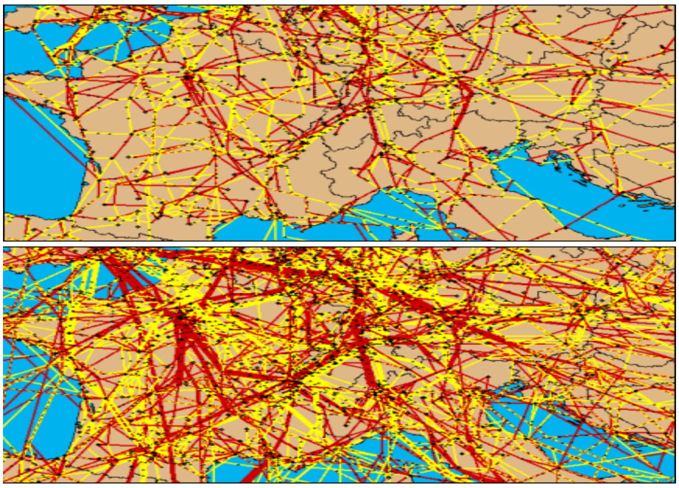


Fig. 2.4. Comparison of traditional, and small aircraft flights distribution. (presented with COSAAC software)

destination, which could be due to flight plan errors, or even training flights.

Hence SAs are flying at low altitudes, for short distances, with a very low number of scheduled flights. Their distribution is very high, thus it would be difficult to identify preferable destinations, or city pairs, as it might be done for the conventional air traffic (see figure 2.4.).

III. LESSONS LEARNED FROM AIR TRAFFIC ANALYSIS

Even so cruising altitude distribution of small aircraft flights has been mentioned, their impact on the airspace density, and on other flights requires further investigations.

Here, both small aircraft and commercial flights have been examined in the same time, which allows us to assess the most preferable altitudes for both cases, and to evaluate whether any kind of impact exists, or could exist between them. The figure 3.1. shows that nowadays the above mentioned flights are quite separated by more than 20.000 feet between their most crowded areas. But if we would like to trace an optimal limit between them – which means, that both impacts the other in a minimal way – that would be around FL 190. In that manner, 93% of the commercial flights will take place above that altitude, and they will be impacted by 18% of small

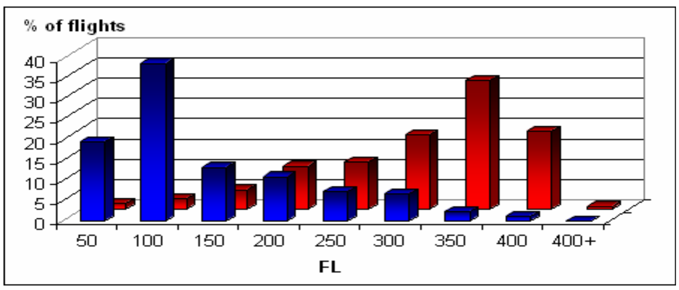


Fig. 3.1. Commercial (red) and small aircraft (blue) flights RFL distribution.

aircraft flights.

Anyhow, even if that trend in small aircraft altitude preference remains the same, a raised number of flights could shift the previously mentioned 18% to higher values, which might seriously impact the commercial air traffic. Moreover the same problem should be faced, if any SA performance or flight distance value changes, which calls for further ATM studies (such as the assessment of the number of conflicts, sectorization constraints, etc.) in order to meet target objectives and make small aircraft transportation relevant.

An additional study to gain more accurate information on the current state of aviation could be the 3 dimensional altitude distribution within a given sector (including incoming and outgoing flights) at airport vicinity.

The figure 3.2. shows the complexity of traffic situation,

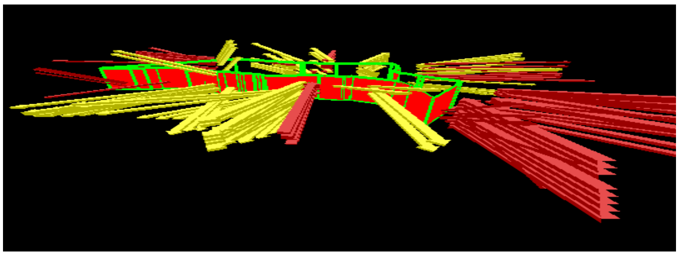


Fig. 3.2. Complex traffic situation at airport surroundings (presented with COSAAC software)

which is no more a clearly separated cruising altitude distribution, but much more a mixture of small aircraft and commercial flights, with some descending, climbing and cruising airplanes at airport surroundings. From that point of view, commercial flights should face with serious problems, while they desire to pass through low level airspaces crowded with small aircraft, especially in case of an increased traffic volume. Certainly, many solutions could exist to solve this problem, while minimizing the impact of small aircraft. One of them lies in the idea to oblige small aircraft flights to make a deviation at airport vicinity, which might enable the rest of airspace users with the same service, as today.

From a small aircraft point of view, airport surrounding means the areas with commercial flights lower than FL 190, which as previously mentioned is the optimal altitude where the impact of both small aircraft and commercial flights is minimal. In the aim, to imagine how seriously such a procedure could limit the freedom of small aircraft users, we calculated the average distance (depending on aircraft performance and airport SIDs: Standard Instrumental Departure routes, STARs: Standard Terminal Arrival Routes [9]) that might be required for a commercial flight to reach FL 190. That gave us 130 km [9]. Finally, by tracing circles with this radius around big European airports we had the figure 3.3. . Surprisingly, these circles are covering an exceptionally large geographic area, which can not be deviated without the serious impacts on small aircraft's flight plan and contradictions with target objectives such as freedom.

Thus other solutions have to take into consideration, such as defining corridors within the airspace that might enable small aircraft to pass by crowded areas. Anyhow, a major

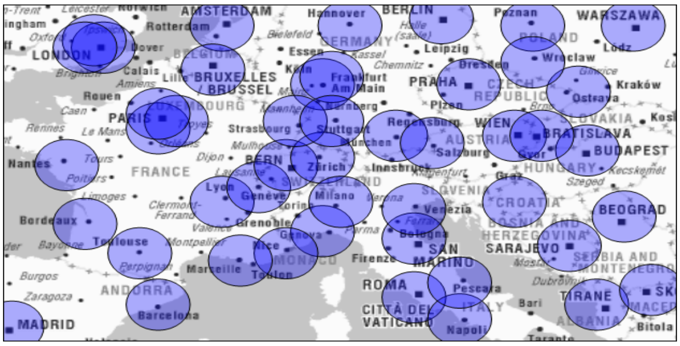


Fig. 3.3. Low altitudes with high aircraft density around big European airports: a constraint on freedom for small aircraft.

conclusion is that a flexible or optimal usage of airspace capabilities that replies to dynamic traffic requirements could become important. Nevertheless that could create an increased pilot/controller workload and high requirements on cockpit instrumentation, which might call for an increased level of automation.

IV. INITIAL SMALL AIRCRAFT PREDICTION MODEL - A GENERAL OVERVIEW

The input data analysis clarified, that the initial problem with SA prediction modeling is, that the generally used solutions can not give a suitable result, since small aircraft transportation does not exist today in Europe, hence the relevant information which might be used is more than limited. This lack of input data, and the aim to have an advanced model that could reply to our specific requirements, forced us to come up with the (already discussed) air traffic analysis. Using its results and arguments on the number of small aircraft, flight level, and other characteristics, the elements and the initial data of a model started to be available. However, we already had something in mid about the key elements, their role and weight remained unclear. Thus, firstly we decided to investigate an initial prediction model that could finally answer the remained uncertainties, and which could serve as a good starting point to represent the relationships, and to understand their functionalities. Moreover, the right balance between the major factors should be found and applied, in order to eliminate unforeseen complexities and side effects (such as increased costs due to a high level of cockpit instruments), while trying to bring into play the most possible benefits. This balance also calls for an initial SA prediction model.

As preliminary analysis show [10],[11],[12] GDP, costs, population density, and other socioeconomic characteristics could seriously influence the demand of small aircraft, thus it should make a part of our investigation. Even so, our approach to initial prediction modeling is, that the GDP might totally describe the mobility of passengers, hence our input data – that we call here market attributes – mainly consists of GDP, and other SA relevant elements, like technological development and regulation. Naturally other fields could also be added; nevertheless even this simplified input data should fulfill our requirements: to obtain an initial prediction model.

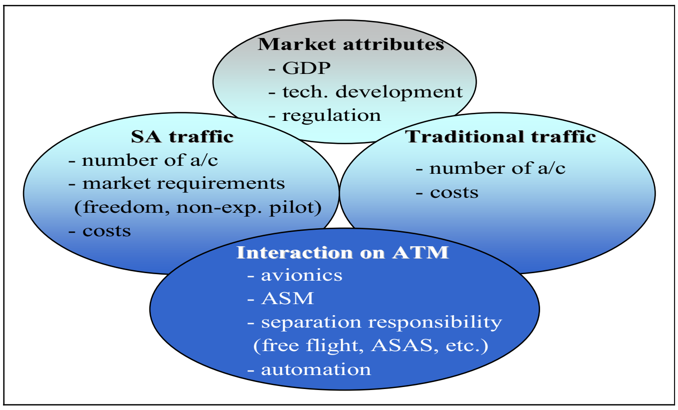


Fig. 4.1. The initial prediction model elements.

As for the rest of our input data, we do believe that the demand mainly depends on the availability of SA, which might be driven by the technological development. Anyhow, the effect of regulation has been added as well, since it might seriously influence the SA activity, in any future scenario (such as noise restriction at airport vicinities).

Then, as the figure 4.1. shows, that market attributes drives the characteristics of small aircraft (SA) and traditional air traffic as it comes for economical theories [11] [12]. Thus the small aircraft group (figure 6.1.: left, in the middle) consists of the number of aircraft (SA need) and its cost (assuming, that it could be defined by the technological development and regulation). Additionally, this same group contains a supplementary domain called flexibility of market requirements. This element takes into consideration the flexibility of the market to SA cost and to the fulfill of its requirements. Thus a low market flexibility would describe the possibility where SA need is highly dependent on its cost and the fulfill of market requirements, even thought any GDP and technological development increase. With other words, this means, that future SA users will not be able to accept higher costs, or a lower satisfaction of their requirements (such as the non-availability of SA for pilots with limited experience).

In order to have the total amount of traffic, which interaction on ATM might be analyzed, the traditional traffic characteristics has been added as well (figure 4.1.: right in the middle).

The final element of the model – interaction on ATM - is made up of some of the air traffic management domains from a SA point of view. These elements are mainly the outcome of the SA traffic, current ATM constraints and future perspectives analysis [10]. For instance as chapter V showed the difficulty of small aircraft flights at airport surroundings, some revolutionary solutions might require a new ASM, or an enhanced automation, therefore these two domains could make a part of our investigations. Following this logic, avionics – for example - means cockpit instruments and navigation tolls (such as TCAS, GPS, ADS-B, and others). Separation responsibility is defined by its importance, without taking into consideration whether pilots or controllers should deal with it. Similarly, the domain of "automation" means more its significance, without underlining, that it might range from automation of controllers' routine tasks to autonomous operation, with advanced airborne system application (such as Airborne Separation Assurance Systems) and even free flight.

V. PREDICTION MODEL STRUCTURES

Generally, the transportation systems are the dynamic systems. The future of the dynamics systems can be defined with the following general models written in continuous/discrete form such as follows:

$$\mathbf{x}(t_0) = \mathbf{x}_0 \,, \tag{1a}$$

$$\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}(\zeta)\big) + \mathbf{F}(\mathbf{p}(\zeta))\mathbf{n}(t) \,, \tag{1b}$$

$$\mathbf{y}(t_i) = \mathbf{g}\big(\mathbf{x}(t_i), \mathbf{u}(t_i), \mathbf{p}(\zeta)\big) + \mathbf{G}(\mathbf{p}(\zeta))\boldsymbol{\eta}_i \,. \tag{1c}$$

where **x**, **u**, **p**, **n**, **y**, $\boldsymbol{\eta}$ are the state, control (input or regulatory) and parameter (system structural and operational characteristics), state noise, observation (output), and measurement noise vectors, f and g are system state and observation functions, **F** and **G** are system matrices, t is time and ξ is a random value.

However in general – and with accordance to control theory - **u** is a known vector, in our case it is defined by regulatory aspects, like changes in requirements generated by safety reasons, or changes in taxation systems, application of radically new technologies, etc. The system characteristics, **p** are depend on the development of SA, economical characteristics, and on the competitor aspects (such as growth of the high speed trains, road traffic problems, etc.). Hence, **p** is an unknown vector, which can be changed randomly and non-continuously that is described by a random value, ζ depicting real position of the parameter vector, **p** in its possible space, Ω_p. As usually, the state noise vector is assumed to be zero-mean; and the measurement noise vector is assumed to be a sequence of independent Gaussian random variables with zero mean and identity covariance.

Finally, the model (1) can be given in linearised form (as is the case of stability and control derivatives of an aircraft):

$$\begin{aligned} &\mathbf{x}(t_0) = \mathbf{x}_0 \,, \\ &\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{F}\mathbf{n}(t) \,, \\ &\mathbf{y}(t_i) = \mathbf{C}\mathbf{x}(t_i) + \mathbf{D}\mathbf{u}(t_i) + \mathbf{G}\boldsymbol{\eta}_i \,. \end{aligned} \tag{2}$$

On the other hand, the model (1) represents the following stochastic (random) differential equations in system of equations (1a):

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \boldsymbol{\sigma}(\mathbf{x}, t)\boldsymbol{\xi}(t) \,, \tag{3}$$

which is called as diffusion process. The first, deterministic part at the right side of the equation describes the direction of the changes of the stochastic process passing through the $x(t) = X$ at the moment *t*, while the second part shows the scattering the random process, where, $\xi(t)$ is the white noise.

In prediction and forecast technology several models based on the use of diffusion models were developed [14]. From one hand the innovation diffusion theory has applied the S models for getting information about the introducing of a new product/service into the market that aims to describe the market share changes [14].

On the other hand the model classes defined by (3) can apply to the description of the system dynamics rewriting the model (1b) into the following form of controlled diffusion process:

$$\dot{\mathbf{x}} = \mathbf{\Phi}(\mathbf{x},t) + \mathbf{b}(t) + \mathbf{\sigma}(\mathbf{x},t)\xi(t) \, , \qquad (4)$$

where $\mathbf{\Phi}$ is the deterministic vector function describing the rate of changes in state vector, as the products of the functions of the state and the time increment; **b** is the vector of control effects and finally $\mathbf{\sigma}$ is the transfer matrix giving the effects of the noise disturbance on the state vector.

Replacing the state (or phase) vector **x** by $\mathbf{x} = \mathbf{m_x} + \Delta\mathbf{x}$ the equations (4) can be statistically linearised in the area closed to $\mathbf{x} = \mathbf{m_x}$:

$$\dot{\mathbf{x}} = \frac{d}{dt}(\mathbf{m_x} + \Delta\mathbf{x}) = \mathbf{F}(\mathbf{m_x},t) + \mathbf{U}(\mathbf{m_x},t)\Delta\mathbf{x} + \mathbf{b}(t) + \\ + \mathbf{\sigma}(\mathbf{x},t)\xi(t) \qquad (5)$$

where $\mathbf{U}(\mathbf{m_x},t)$ is the sensitivity matrix, i.e. matrix of partial derivatives of vector function $\mathbf{F}(\mathbf{x},t)$ respectively to state vector **x** determined at $\mathbf{x} = \mathbf{m_x}$.

In case of stationer white noise the equation (5) represents the well known linearised model of aircraft motion:

$$\dot{\mathbf{m}}_\mathbf{x} = \mathbf{F}(\mathbf{m_x},t) + \mathbf{b}(t) \, . \qquad (6)$$

However, this class of models could be applied for the prediction of the small aircraft activity, the relevant preliminary information – in Europe - that might be required for model estimation, is more than limited. Moreover, according to our analysis, the prediction model can not be given in a generalized form, since the system should include the major effects influencing on SA growth, which defines a large and very complex system with internal coupling and discrete (step) changes - depending on the regulational aspects or the application of the new technological achievements.

Thus, the prediction model - in a general form - is the result of the superpositions of the general growth (exponential), periodical changes (in requirements) and the discrete changes (in characteristics of general growth). Hence, the model (1b) should be rewritten in the form of stochastic equations, such as follows:

$$\dot{\mathbf{x}} = \mathbf{f}_x(\mathbf{x},\mathbf{u},t) \qquad (7)$$

In order to be able to use this model, the SA traffic characteristics should be analyzed at first, as it was already mentioned at the previous chapters.

VI. INITIAL SMALL AIRCRAFT PREDICTION MODEL - MATHEMATICAL DESCRIPTION

In order to set up a mathematical model – that replies to our complex interactions – we assume that the equation (7) hold the form, and where **x** is the vector of the dependent variables like:

$$\mathbf{x} = \begin{bmatrix} SAneed \\ SAmark\_req \\ SAneed \\ Tneed \\ T\cos t \\ avionics \\ ASM \\ sep\_resp \\ automation \end{bmatrix}$$

u is the input vector,

$$\mathbf{u}^T = [GDP \quad regulation \quad techn\_development]$$

and t is the time.

The equation (7) is now a non-linear differential equation. With its linearization we receive (8) such as follows:

$$\dot{\mathbf{x}}(t) = \mathbf{A}^* \cdot \mathbf{x}(t) + \mathbf{B}^* \cdot \mathbf{u}(t) \qquad (8)$$

Where

$$\mathbf{A}^* = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & a_{ij} & \dots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}, \text{and} \quad \mathbf{B}^* = \begin{bmatrix} b_{11} & \dots & b_{1m} \\ \dots & b_{ke} & \dots \\ b_{n1} & \dots & b_{nm} \end{bmatrix}$$

with

$$\mathrm{a_{ij}} = \frac{\partial f_i(x,u,t)}{\partial x_j} \, , \text{and} \quad \mathrm{b_{ke}} = \frac{\partial f_k(x,u,t)}{\partial u_e}$$

The equation (8) could be discretized such as follows:

If T is the disretization time, we could define:

t=k*T , where $k \in N$ is the time in years.

Thus the equation (8.) takes the following form:

$$\frac{\mathbf{x}[k+1]-\mathbf{x}[k]}{T} = \mathbf{A}^* \cdot \mathbf{x}[k] + \mathbf{B}^* \cdot \mathbf{u}[k] \quad (9)$$

knowing that $\Delta t \Rightarrow T$.

With the rearrangement of (9), the prediction of the elements of the vector **x** could be done with the following equation:

$$\mathbf{x}[k+1] = \mathbf{A}\mathbf{x}[k] + \mathbf{B}\mathbf{u}[k] \quad (10)$$

Where, the matrix **A** describes the relationships between the

$$\mathbf{A} = \begin{bmatrix} SAneed & a_{12} & \dots & a_{17} & a_{18} & a_{19} \\ a_{21} & SAmarkreq. & \dots & a_{27} & a_{28} & a_{a29} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{71} & a_{72} & \dots & ASM & a_{78} & a_{79} \\ a_{81} & a_{82} & \dots & a_{87} & sep.resp. & a_{89} \\ a_{91} & a_{92} & \dots & a_{97} & a_{98} & automation \end{bmatrix}$$

elements of the vector x, therefore a_{ij} (where $i \neq j$) are coefficients, that express the connection between a_{ii} and a_{jj} (such as the connection between automation and SA need). **B** similarly to **A** describes the relationships between the input elements (vector u).

The advantage of such an approach is that it could be used for any interaction modeling even with different small aircraft characteristics.

However, the challenge is to find the coefficients for both **A** and **B** matrixes. For our small aircraft purposes, they are partially based on statistics, and on estimations.

Finally, the outcome of the equation (10.) could give us an initial prediction of each elements of the vector **x**, that might help us to understand their role, and to foresee, what could be a relevant domain to focused on.

VII. SIMULATION

To have the prediction of the key elements (vector *x*), in this paper, only a example of a simulation is presented, where the input data is defined with five scenarios, that ranges from optimal (*4*: large number of SA) to catastrophic (*5*: limited number of SA) and some between such as follows: scenario *1* is defined by a moderate GDP, technological development growth with a non-flexible market (see chapter IV). Scenario *2* is the opposite of the previous, thus here, the market is flexible, and it can accept higher cost, and / or lower level of satisfaction of their requirements. Additionally, for both scenario *1* and *2*, a regulation has been added, since it might be interesting to define, and to analyze its impact on the key elements. Note that in this model, the effect of regulation is considered as a decreasing factor on the “flexibility of market requirements”, and an increasing cause on “SA costs”. Finally, scenarios *3*, *4* and *5* are the ones with a flexible market, and without regulation, where the difference between them lies in the definition of a low (*4*), moderate (*3*), and rapid (*5*) GDP, technological development growth. As a final point, note, that all the input data are based on assumptions.

By running a simulation with these scenarios, the model gives the outcome of each key elements (vector *x*), however, this paper only presents three of them (SA need, fulfill of market requirements, and the role of automation), as it is presented in the figure 6.1. .

By focusing on the “SA need” outcome, it predicts nearly six times (see figure 7.1.) more small aircraft (that is today) for the most optimal scenario (*4*). This number becomes a bit smaller with scenario *3* and *5*, which is caused by the decreased GDP and technological development growth. Finally, scenario *1* and *2* are the situations, where the outcome is impacted by a regulation. Its effect is clearly visible on both curves, causing a wave part that lasts for several years. Logically, for scenario *1* it is easier to track, due to the non-flexible market situation, which is more sensitive to any

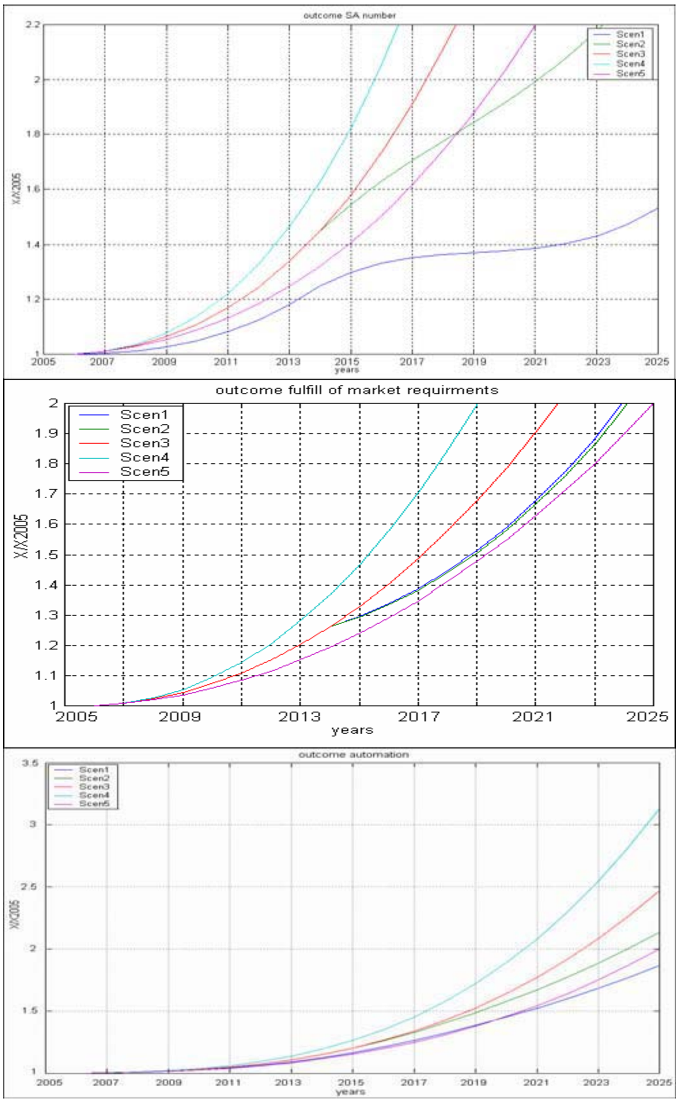


Fig. 7.1. The output data of the model. On the top: SA need, middle: fulfill of market requirements, bottom: role of automation.

“cost” or “flexibility of market requirements” change.

As for the effect of regulation, it could be also observed in the figures of ”flexibility of market requirements” outcome (see figure 7.1.). Here again, only scenarios *1* and *2* are impacted, otherwise the curves are slightly different due to the variation in SA number.

As for the curve automation, it has to be underlined again, that this analyze does not aim to point out one of the automation capabilities, since it only represents its importance to achieve the future air transportation of small aircraft. Anyhow, in the figure 7.1., its increasing role might be observed, that could be even three times more important, that is today. This result might correspond to any future scenario, since automation capabilities should become a major factor in the future, where aviation should face with an increased number of small aircraft that might significantly impact the conventional air traffic (as it was already foreseen in the small aircraft resembling traffic analysis). Moreover, if these small aircraft are handled by pilots with limited experiments who would not be able – or just do not wiling to – deal with high separation responsibilities and difficult procedures, automation might become even more important.

VIII. CONCLUSION

The increasing economy and air traffic volume might allow the future of small aircraft transportation. In that scenario, traveling habits will change. Maximum satisfaction of market requirements will take over the lead, which could enhance people’s quality of life and ensures their freedom.

Such a shift in air traffic attributes, and the increased number of flights could call for a change or demand enhancement is several areas of the current ATM, thus it has to be analyzed. The identification of the bottleneck among these ATM domains could be imagined by the results coming from a mathematical model that describes the impact of small aircraft on ATM.

In the face of this model calls for relevant information and the knowledge on small aircraft transportation attributes, in Europe, the availability of these data is extremely limited. This forced us to accomplish the initial objectives, namely the analysis of small aircraft flights. This partial result allowed us to place small aircraft flights in the airspace, and to use arguments on flight level, flight distance, propulsion technology preference, that is mainly impacts the conventional air traffic at airport surroundings.

With these clarified main characteristics, and the restricted area that should be focused on, the definition of the mathematical model became possible. The initial prediction model showed the complexity of the work ahead of us, and gave some preliminary results on the number of small aircraft, and the characteristic of some of the ATM fields that might be focused on.

Anyhow, the prediction model could enable the definition of the bottleneck among the ATM parameters, that might be focused for any further investigations, where new technologies and ideas - enabling such modifications in air transportation – could be applied to the entire system intelligently and in an integrated fashion, in the aim of having a new, effective, more safe and accessible way of transportation to any users.

IX. FUTURE WORKS

After the initial prediction model, and once the relationship between some of the most important elements of the European small aircraft transportation has been clarified, further investigations should focus on an advanced mathematical model that already replies to a complex system approach. A general system overview of such a small aircraft prediction modeling could be described such as follows: the ATM could be considered from a complex system modeling view at the macroscopic level, where the whole system is decomposed into the following three subsystems:

- Society Subsystem: that contains the society constraints and drives for growth including passengers and airlines,
- Technical Subsystem: that includes all technical infrastructures supporting the functioning of the ATM system,
- Human Subsystem that takes into consideration all human components of the whole system, from flow managers to supervisors, pilots and controllers.

From such a point of view, the impact of small aircraft on ATM could be modeled in terms of sets of components of different subsystems and sets of interactivities between them, to accomplish the mission of simultaneously maintaining safety and sustaining growth.

This model will also rely on a more advanced small aircraft demand model, using theories of macroeconomics in air transportation. This challenging task could be accomplished using similar activities that might be adapted to European small aircraft and market requirements.

Naturally, due to the lack of preliminary information and dependence on the applied scenarios, the model must be tested and enhanced in accuracy for further use.

Finally, once the model is already available, with the use of scenarios, we could provide simulation based recommendations and proposals for further Air Traffic Management research areas, from a small aircraft point of view.

REFERENCES

[1] Airbus (2003): Global Market Forecast 2003-2022, December 2003 http://www.airbus.com/pdf/media/gmf2003.pdf

[2] Dr. Bruce J. Holmes, NASA; John Scott, Icosystems : Transportation Network Topolo-gies, April 27, 2003 http://spacecom.grc.nasa.gov/icnsconf/docs/2004/01_plenary/PS-06-Holmes.pdf

[3] Gundlach, M., Rohacs, J.: Regional Flight 2000, Hungary, reports I - III, BUTE - Budapest, Dornier - München, RHTW - Aachen, 1991-93.

[4] Rohacs, D. Diploma thesis: Nouveau systeme de controle automatique pour de petits avions, 2004 July, INSA de Lyon & BUTE.

[5] Rolls-Royce : Business Jet review and forecast, NBAA Las Vegas, October 2004
Available:
http://www.rollsroyce.com/civil_aerospace/overview/market/outlook/downloads/busjet04.pdf
(10/11/2004)

[6] Dr. Bruce J. Holmes, NASA : Small Community Air Transportation, Oklahoma – CASI, November 14, 2003 OK_CASI_11-14-03.ppt

[7] Rohacs, J.: PATS, personal Air Transportation System, ICAS Congress, Toronto, Canada, CD-ROM, 2002, ICAS. 2002.7.7.4.1 -7. 7.4.11.

[8] AVCEN LIMITED: Maker of JETPOD aircraft.
Available: http://www.avcen.com/index3.php
(05/04/2005)

[9] Italian vACC.
Available:
http://www.vatita.net/?dir=download&pagina=planning/charts
(13/08/2005)

[10] Rohacs, D. Analysis of the impact of small aircraft on ATM in Europe. In Proceedings of the 9th Air Transport Research Society world conference (ATRS), Rio de Janeiro, Brazil, June 2005.

[11] Stouffer V., Design document for European Air Carrier Investment model, Working document, Logistic Management Institutes, EU201L2, May 2002.

[12] Wingrove E.R., Gaier E.M., Santmire T.E., The ASAC Air Carrier Investment Model (Third generation), Logictic Management Institutes, NASA/CR-1998-207656, April 1998

[13] Peakin, W.: "Flying will be as easy as driving a car" *The Observer*, February 25, 2001
Available:
http://www.observer.co.uk/u_news/story/0,6903,442773,00.html
(22/04/2006)

[14] Murrae J. Bowden: "Moore's Law and the Technology S-Curve"
Available:
http://howe.stevens.edu/SATM/archive/v8i1/bowden.pdf
(10/04/2006)

Uncertainty Decision Making in Unmanned Aerial Vehicles

Ricardo A. V. Gimenes and Jorge R. Almeida, Jr.

Abstract—**Due to a rising interest on civilian use of Unmanned Aerial Vehicles (UAV), developed countries have been making significant efforts to integrate UAV operations in controlled aerial space. Albeit consensus is not yet established on which criteria should be put to use, it is commonsense in the aviation community that at least current civilian aviation safety standards should be upheld. For this purpose, this PhD research proposes a hazard analysis, initially focused on two critical situations in UAV operations. Studied cases are "on-ground" collisions and "midair" collisions. In several cases, an on-ground impact model must be used in order to calculate UAV system's reliability, which is a pre-requisite to achieve desired safety level in "densely populated" area for different UAV categories. Midair collision risk must be estimated in the vicinity of controlled "flight corridors", according to an aircraft collision model based on air traffic density for such areas.**

Index Terms—**Hazard, Reliability, Safety, Unmanned Aerial Vehicle.**

I. INTRODUCTION

THE concept of Unmanned Aerial Vehicle (UAV) is not a particularly new one. As a matter of fact it first appeared on early World War I. Notwithstanding, US only took a first series of meaningful experiments on it in the 1950s.

The idea of accomplishing espionage missions or delivering ammo for troops beyond the enemy lines, among other risky assignments, not endangering a human crew, has been tempting military strategists all over the world.

Initial enterprises on UAV usage were unsuccessful and yet, war in Vietnam and the Cold War instigated several R&D programs, producing assorted "search and recognition" UAV, such as "Fire bee" and "Lightning Bug" [1].

Although the first generation of UAV was hard to maneuver and demanding on maintenance, American Air Force deployed them on a wide range of tasks, such as acquiring images at low altitudes both by day and by night.

By that time, the need for unmanned aerial vehicles was over simultaneously with Vietnam War. Even then US Armed Forces were still interested in exploring potential usage for this kind of aircraft. Military events in the Mid East from the 1990s revived interest on UAV. As a result, good performances by "Predator" and "Global Hawk" on those further foster interest on civilian use of UAV [2].

On civilian use, NASA run some programs in the 1960s, such as Pa-30 in 1969, resembling a remote control operated aircraft, controlled by a pilot from an on-ground station [3].

Currently, UAV's suitability for boring and/or risky assignments ("dull, dirty and dangerous", in military jargon) has been boosting the emerging civilian market. A Dull mission can expose the crew to a high fatigue as a dominant concern. That fatigue can seriously compromise the operational proficiency and expertise of the crew. A Dirt mission is a typical mission where the crew could be exposed to hazards that can compromise their lives. An example has happened with a crew who was responsible to collect samples after a nuclear explosion test. A Dangerous mission includes those that expose a crew to real threat. The use of UAV in military missions can reduce the risk to loss of human life in high threatening environments.

Many magazines specialized on prognostics for technological applications mention UAV deployment on missions for which operational costs are, notwithstanding, too high.

Although development of UAV civilian use is still at its initial phase, it has been looking quite promising when it comes to spawning new navigational control technologies, as well as more efficient propulsion systems and others, such as sensors and other avionics devices.

And yet, there are still some barriers to overcome in order to enable civilian usage of UAVs, like:

--Lack of airspace regulation, so every UAV model may properly "fit in";

--Lack of a safe and reliable protocol and "Data link";

In order to develop a civilian market, cost-benefit analysis should support UAV use to replace everyday procedures, in addition to the accomplishment of currently impracticable missions. Anyway, safety factor must be assessed and accounted for.

Lack of clear regulations is a major barrier against civilian and commercial operations. This paper aims at sorting out its potential while fulfilling the according safety requirements.

The rationale for military usage of UAVs is easily understandable, in view of the concept of Dull, Dirty and Dangerous missions [4]. This concept is fully justified in cases of crew endangerment situations. For civilians, the major

Manuscript received April 21, 2006.

R. A. V. Gimenes is with the Electrical Engineering Department as PhD Student, School of Engineering, University of São Paulo, Brazil. (e-mail: ricardo.gimenes@poli.usp.br).

J. R. Almeida, Jr, is with the Electrical Engineering Department as Associated professor, School of Engineering, University of São Paulo, Brazil. (e-mail: jorge.almeida@poli.usp.br).

motivation has been the constant search, by aircraft manufacturers, for cost reductions, implying in pressing demands for use of new technologies that should be, on the grounds of aeronautics principles, safe and reliable [5]. That is rationale behind studies like this one, even if they come to demonstrate limitations to the use of UAVs.

II. UAV CATEGORIZATION

Unmanned Aerial Vehicle may be described and categorized according to many criteria, such as flight altitude, endurance, observability and size, among others. For aeronautics, endurance is a jargon that conveys a compound meaning including the ideas of ruggedness, reliability and availability for a given period.

There are some international attempts on grouping them in "Tiers" [6], but there is such a wide assortment of them that category overlapping is nearly unavoidable.

Intending on minimizing such an inconvenience, international Forum on UAV [7] set a set of descriptors based on flight characteristics, size/weight and functionality.

There already are some civilian UAV usage either operating or in developmental stage, namely those related to surveillance and access to risky areas such as nuclear facilities, power transmission line maintenance and others [8]. Those UAV do not have significant influence on the air traffic system, because they fly over clearly defined areas at low altitudes and low speed.

But, in order to the aeronautics industry to integrate civilian UAV usage, some issues must be dealt with: technical (aircraft control technology), political (border control issues) or legal (liabilities in case of accidents).

One can see that categories still cling to a military heritage, because the military are still involved in most major UAV development. The main military feature is avoiding human crews at all, in spite of severe critics against armed military operation, such as what criteria should an algorithm assess in order to open fire [9].

Opposing to this, crewless, civilian passenger aircraft raise a completely diverse paradigm in terms of reliability, compared to military usage. Because of this, new hurdles come in the way of a future categorization of passenger carrying UAV, as well as of non-passenger carrying UAV sharing airspace with conventional passenger carrying aircrafts.

Then, operating UAV on international airspace must imply two closely related conditions [10]: the UAV must be safe and reliable enough to fly over densely populated areas and must be safely operated through airspace. Although both requirements are defined generically, they should account for the widest possible range of unsafe situations both for the airplane itself and its surrounding environs.

From such general safety requirements, arise research subjects such as risks on collision to the ground and midair collision. Both relate primordially to integrating UAV operation to the current system, which demand appropriate rules that enforce fulfillment of those requisites. In this paper we assume that UAVs are fit for use in the traffic system, and building on this hypothesis present ground collision and midair collision related safety issues.

III. SAFETY ANALYSIS METHODOLOGY

In order to analyze safety issues arising from operating the different types of UAV in densely populated areas, one should run a "Preliminary Hazard Analysis” - “PHA" -, according to the safety rules by the International Civilian Aviation Organization (ICAO) [11].

An preliminary analysis will identify critical risks associated to UAV operation, in order to model predictable values for each risk and assess their implications on UAV systems design and operation. Our methodology is consistent with current literature on systems safety [12]. To keep it applicable to a wide variety of UAV systems, risk identification is done by means of an approximation based on event consequences. This approximation focuses on events classified as extremely harsh. Due to reliability demands for aviation systems, events with most grievous consequences should be the guide for the project of critical systems. Two UAV operation related events are of major interest for public safety: ground collisions and midair collisions (the latter involving other aircraft).

Albeit there are other potentially catastrophic risks, the aforementioned critical events have direct influence on integration of UAV to the current controlled airspace.

Finally, perceptions of risks and benefits of civilian operation of UAV by the general public, together with interest on the system by investors, will have a deep influence on air traffic safety and management politics.

A. Ground Collisions

Statistics on aircraft accidents show that 15% of all accidents happen on take-off and initial climb and other 50% on descent (near landing) and landing (go-around included) [13]. It is quite interesting to realize that main reasons behind such mishaps are lack of positional awareness in the air and inappropriate decision making. For UAV advocates, such failures may have a lesser incidence on use of automated crewless aircraft control. Such an example is CAT-IIIc landing, for which few airports have qualified operators authorized to land in zero/zero conditions.

On the other hand, one should care about problems arising from the very use of avionics replacing human operators aboard an aircraft. Namely, criteria definition are needed to achieve acceptable safety and reliability levels for a totally automated aircraft control system.

Automation and avionics became unavoidable in every field of aviation, be it Fly By Wire flight control module, or be it a set of navigation aiding systems. In the realm of in-flight airplane, any maintenance operation becomes staggeringly complex, and components must follow the ‘safe mission’ criterion, for they do not allow for in-flight maintenance

(MTBF is made equal to MTTF). This immediately implies that any project in avionics has special requirements on tolerance to failure [14].

As an example on such concerns, accidents such as that involving a (currently) brand new Airbus A320, that was first effectively operating on Fly By Wire. It happened in 1988 under suspicion of Fly By Wire system failure, according to the airplane captain. Official version states that the aircraft operated properly, but since May 1998, it is proven that the Flight Data Recorder was switched after the accident [15].

From the moment a given aircraft is defined as UAV, all of its control systems are run by electronic components that are held liable for all decision making required on flight. In such a case, redundant analogical equipment are not necessary, once there is no crew to use it.

On this subject, one must also account for the different types of landing procedures, including emergency landing, both on airport runways and on inadequate terrain.

As examples to unpredicted situations, one may quote cases such as the landing gear problems of a Jet Blue Flight 292 or that Piedmont Airlines' (USAir) airplane, which was forced to make an emergency landing with one of its main landing gears completely retracted.

In extreme cases, use of algorithms for decision making is still perceived as unreliable, since their capabilities are very limited as yet.

Another complex situation for decision making by an algorithm is high speed approaching for engines-out emergency landing, as happened to a Boeing 767 that run out of fuel at 41,000 feet. Twelve kilometers above Manitoba countryside, the unthinkable happens: a brand new Air Canada Boeing 767 runs out of fuel dropping at over 600 meters per minute. Despite of the critical situation, the powerless plane makes a successful emergency landing on an abandoned airbase. This accident is understandable, if one keeps in mind that when it happened (1983), pilots as a whole didn't trust onboard computers and that airplane's crew didn't take heed of the alarms [16][20].

And yet a very similar accident happened to an Airbus 330 in 2001, also due to fuel leak [17]. The aircraft glided in to the airport, carried out an engines-out visual approach at night, in good weather conditions and good visibility. The aircraft landed fast on the runway, with reduced braking possibilities due to lack of some electrical systems, and came to a halt [18].

This kind of incidents must be thoroughly analyzed in order to overcome the challenges inherent to UAV control systems.

B. *Midair Collision*

As well described in [19], the main task of air traffic controllers managing arrival traffic is to sequence, group and individuate aircraft for landing. Unfortunately, the airborne counterpart of the arrival manager, which could help the flight crew to converge aircraft towards a fixed distance according to a constrained sequence, is not yet available. This kind of maneuver seems difficult to perform by humans, and may induce excessive increase in flight crew workload, thus requiring new on-board automated functions.

One may easily see the search for automation of several systems, both avionics and air traffic control related. This leads industry to search for aerial systems automation, envisaging optimizing maximum use of airspace. In this sense, a trend in new projects from big companies is research on UAVs that manage flying in the airspace.

But to make UAV operation in the airspace truly safe, it should minimize the unsafe states in take-offs and landings, in order to avoid midair collisions an UAV should take heed of mobile obstacles (such as adverse weather and other aircrafts) that may or may not be another UAV.

Answers involving artificial intelligence software or multi-agent systems have been researched on, in the search for decision-making systems capable of avoiding aircraft encounters in a given airspace [20].

To make such automated systems safe and reliable, their implementation should be designed and defined for the entire current air traffic control scenario. And that because UAVs demand new communication methods and equipments for the air traffic control system, since voice command by ATC to a crewless aircraft is still a problem.

In such a scenario, CNS/ATM enables technological progress, and communication may happen by means of a data link. And yet, one must also think about critical situations arising from signal losses, be it due to a general failure by ATC (Air Traffic Control), weather interference or a failure by avionics in the UAV.

IV. UNCERTAINTY DECISION MAKING

The belief that all airspace systems will be automatic controlled generates very hard challenges by controlling elements. Sabatini, as associate administrator for Aviation Safety in FAA, believes that UAV "is the next great step forward in the evolution of aviation" [21]. This declaration was made in March 29, 2006, what proves UAV is a state of the art research which offers a huge challenge to all airspace research community.

We should consider an environment of uncertainty that all UAV and others systems involved have to make decisions. It is a consensus today that fuzzy logic and similar theory are not recommended to be used in safety systems, but the demand for more and more aircrafts flying at the same time exposes a new scenario. This scenario obliges that the technology should be improved and new premises of safety have to be implemented.

A. *The use of Uncertainty Decision Making Methods*

There are many methods for automatic decision-making, including uncertainty domain. The first step of our research is approaching to investigate a set of methods that are appropriate to an airspace domain. It is possible that the new decision making systems will use a set of techniques like Bayesian Belief Nets, Expert Systems or Fuzzy Logic, integrated each other to minimize problems like safety and to obtain the "best" decision at a situation.

It is well known that this research do not have the ambition to derive all possible paths to describes all critical scenarios in a unmanned airspace traffic control. Then, some goals are realistic like the use of a specific method approach to an unmanned aircraft vehicle considering all boundaries defined by the research.

B. Uncertainty Decision Making in UAV Systems

The objective of this research is to map what a decision making system have to do considering the typical phases of a flight. The list of all decisions in a typical flight offer the possibility to describe safety states that can be shown in a fault tree, for instance.

After that, it should present a specific set of methods, that are being studied, that offers feasible solutions in an environment of uncertainty. The reference [22] presents the details of the UA platform, system identification, reconfigurable controller design, development, and implementation on the UA to analyze the performance metrics. An online learning neural network is augmented to form an outer loop to reconfigure and supplement the optimal controller to guarantee a practical stability for the airplane. This paper also presents some simulations from the hardware-in-the-loop testing and concludes with an analysis of the flight performance metrics for the controller under investigation. It is very clear the limitations that are imposed by the study to obtain an unmanned aircraft control.

V. NEXT STEPS…AND FIRST CONCLUSIONS

One may see that UAV has been used for new military applications and civilian research, such as surveillance. Aviation is directly related to the amount of risk that society accepts taking. Then, this preliminary research shows necessary concerns and new fields for study. Research on impacts to the ground and midair collisions defines two main concerns of the air traffic control system.

Building on studies about aforementioned critical risks, research presented in this paper is a first step in air traffic safety and reliability related research.

Further studies may then be conducted and expanded to diagnose and identify critical issues on UAV use and all its implications on safety in a environment of uncertainty.

REFERENCES

[1] Major Robert C. Nolan II, The Pilotless Air Force? A Look at Replacing Human Operators with Advanced Technology The Research Department Air Command and Staff March 1997

[2] Global Hawk Completes First Flight March 2, 1998 - Global Hawk, The Department of Defense's Newest Reconnaissance Aircraft, Flew For The First Time At Edwards Air Force Base, California, On Saturday, February 28th.

[3] NASA Technical Memorandum 4396, National Aeronautics and Space Administration Office of Management Scientific and Technical Information Program, 1992

[4] Office of the Secretary of Defence. Department of Defense. USA. Unmanned Aircraft Systems Roadmap 2005-2030. 2005.

[5] Unmanned Aerial Vehicles Spur Composites Use. Reinforced Plastics, ISSN: 0034-2617/04. Elsevier Ltd. 2004.

[6] Lax, M. & Sutherland, B., An Extended Role for Unmanned Aerial Vehicles in the Royal Australian Air Force, Air Power Studies Centre Paper Number 46, July 1996, p11.

[7] International UAV Forum, www.uavforum.com

[8] Dossier. Civilian Applications: the Challenges Facing the UAV Industry. Uninhabited Aerial Vehicles (UAVs). Air & Space Europe. Vol. I. No. 5/6. 1999.

[9] Sullivan, J. M. "Revolution or Evolution? The Rise of the UAVs. IEEE ISBN: 0-7803-9284-1/05. 2005.

[10] Allouche, M. The Integration of UAVs in Airspace. Air & Space Europe – Operations and Safety. Vol. 2, No. 1. 2000.

[11] International Civil Aviation Organization - http://www.icao.int/

[12] Leveson, N., Safeware: System Safety and Computers, 1st ed., Addison-Wesley, Boston, 1995, Chaps. 13-14.

[13] Jump, M. Padfield, G. D. Progress in the Development of Guidance Strategies for the Landing Flare Manoeuvre Using Tau-based Parameters . Aircraft Engineering and Aerospace Technology 78/1 (2006) 4–12, Emerald Group Publishing Ltd. [ISSN 1748-8842. 2006.

[14] Boléat, C; Colas, G. "Overview of Soft Errors Issues in Aerospace Systems" in 11th IEEE International On-Line Testing Symposium, Proceedings, 2005.

[15] Kilroy , C. Air Disaster. Last access: 20-02-2006. http://www.airdisaster.com/investigations/af296/af296.shtml.

[16] Nelson., Wade H. The Gimli Glider. Copyright WHN 1997 All Rights Reserved. Soaring Magazine. 1997.

[17] Aviation Accidents Prevention and Investigation Department. Accident Investigation Final Report: All Engines-out Landing Due to Fuel Exhaustion Air Transat Airbus A330-243 marks C-GITS Lajes, Azores, Portugal. 22 / ACCID /GPIAA / 2001

[18] Air Transat Flight 236: The Azores Glider Peter B. Ladkin RVS Group, University of Bielefeld, November 18, 2004.

[19] Miquel, T., Loscos, J.M., Casaux, F. Statistical Performance Evaluation Between Linear and Nonlinear Designs for Aircraft Relative Guidance. Toulouse, France, Félix Mora-Camino, Ecole Nationale de l'Aviation Civile.

[20] Lian, Z. Deshmukh, A. Performance Prediction of an Unmanned Airborne Vehicle Multi-Agent System. European Journal of Operational Research. Elsevier. ISSN: 0377-2217. 2004.

[21] Sabatini, N. A. Testimony, Federal Aviation Administration, March 29, 2006.

[22] Janardhan, V. Schmitz, D. Balakrishnan, S.N. Nonlinear control concepts for a UA, Aerospace and Electronic Systems Magazine, IEEE, Publication Date Feb. 2006, ISSN 0885-8985

Author Index