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.

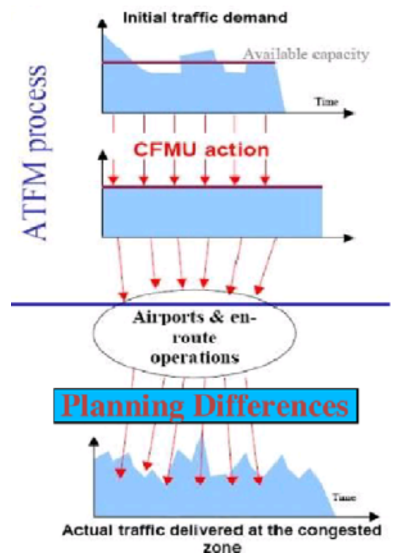

Fig. 1. Current ATFM Procedure

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

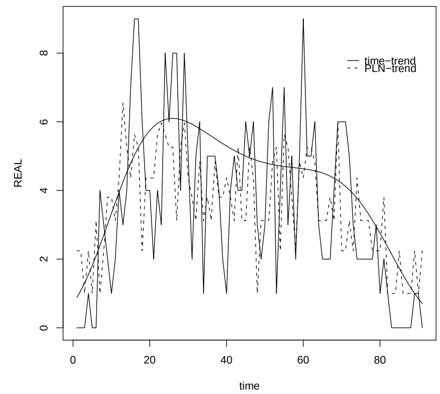

Fig. 2. Arrivals in fifteen minutes intervals with t^4 time-trend (bold) and logarithmic PLN trend (dotted)

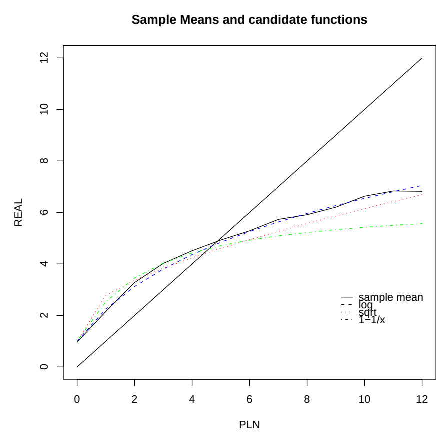

Fig. 3. Relationship between number of planned arrivals PLN (x-axis) and average number of real arrivals $REAL$ (y axis) in 15 minutes time intervals. The bold line corresponds to the sample means. The dotted functions are candidates for modeling this trend: logarithmic, square-root and 1-1/PLN.

$Pr(REAL_t^S | PLN_t^S)$ 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,

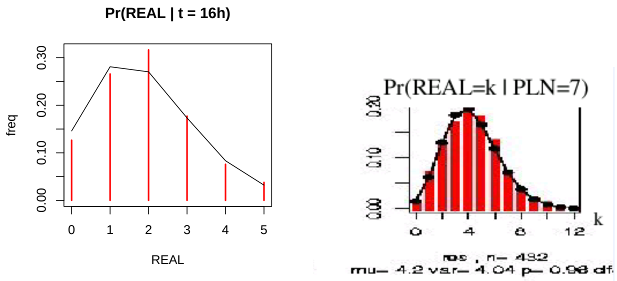

Fig. 4. Examples for distributions of number of real arrivals ($REAL_t$) dependent on time of day (left) and on number of planned entries (PLN_t) (right). Superposed are Poisson distributions. Their parameter is the sample mean.

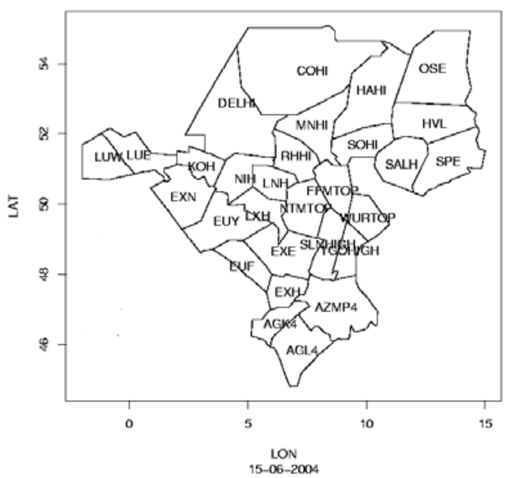

Fig. 5. Central European Upper Airspace. 31 sectors covering the area above London, Zurich and Berlin

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

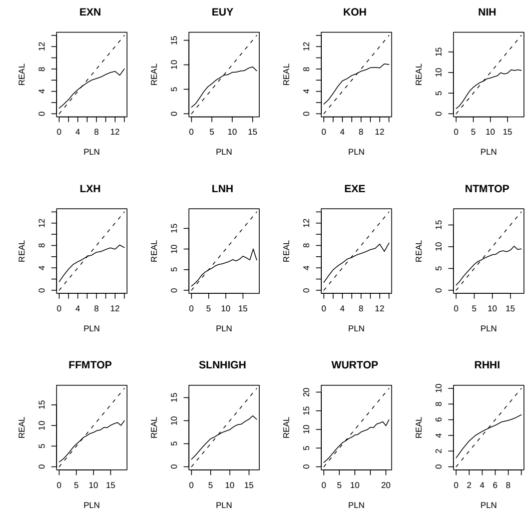

Fig. 6. Same shape of trend for all sectors: plots of the sample means

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.

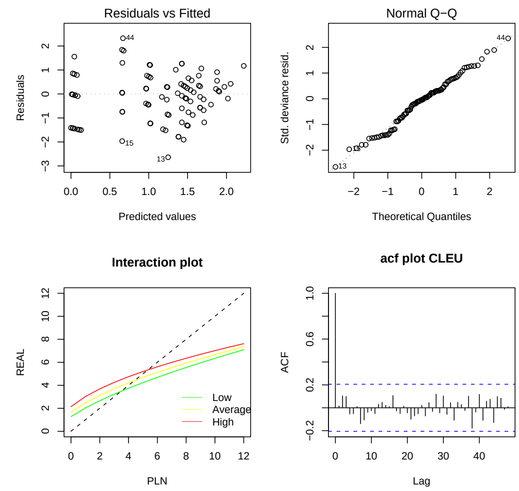

Fig. 7. Typical instance for validation of residuals

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,

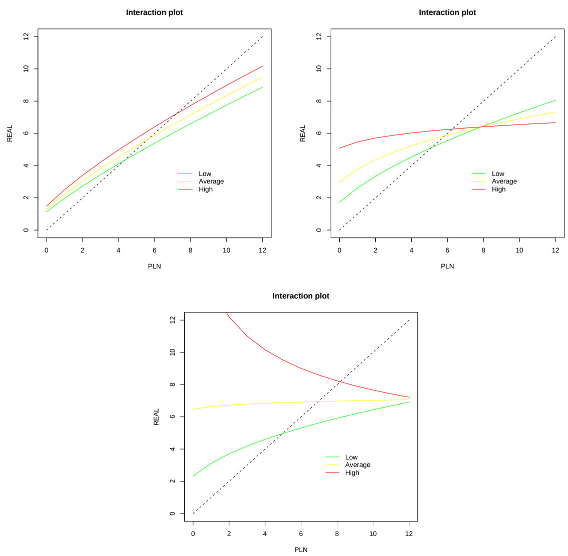

Fig. 8. Three types of predictions. Same (left), with changing effects (middle) and diverging starting values (right)

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.