HAL
open science

An ordered logit model of air traffic controllers conflict risk judgment

Philippe Averty, Kevin Guittet, Pascal Lezaud

To cite this version:

Philippe Averty, Kevin Guittet, Pascal Lezaud. An ordered logit model of air traffic controllers conflict risk judgment. Air traffic control quarterly : an international journal of engineering and operations, 2008, 16 (2), pp 101-132. hal-00940904

HAL Id: hal-00940904

https://enac.hal.science/hal-00940904v1

Submitted on 1 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Ordered Logit Model of Air Traffic Controllers' Conflict Risk Judgment

Philippe Averty
averty@cena.fr

Kévin Guittet
guittet@recherche.enac.fr

Pascal Lezaud
lezaud@recherche.enac.fr

Abstract

Though there is a considerable agreement amongst past studies about the great variability in conflict judgments by Air Traffic Controllers (ATCos), certain behaviors observable in control rooms speak in favor of the existence of a common core they would that controllers share regarding conflict risk perception. The study presented in this paper began with the construction (from real recordings) of traffic scenarios showing two converging aircraft in approach. Three variables characterized these traffic scenarios, quantifying respectively horizontal separation, vertical separation and the momentum of the formation of the judgment (prediction time span). The conditions created by factorial manipulation of three variables led to the design of short scenarios (about 1min) involving two aircraft upon which 161 controllers gave their judgments about possible occurrence of a separation loss. A first descriptive analysis of the data, conducted in Averty (2005), confirmed the large variability of the experts' judgments but also clearly indicated the global consistency of the results. The data thus called for a deeper statistical analysis, the results of which will be presented in the article. In a first step, particular models have been constructed for each value of the prediction time span. The comparison of the model's parameters allows evaluating the influence of the time span on the conflict perception. It appears for example that the horizontal dimension has more "separating power" that the vertical dimension far from the conflict location, but that its relative importance diminishes (along with uncertainty) as the conflict resolves. Individual models are then nonlinearly aggregated into an "integrated model" by maximum likelihood estimation on the whole dataset. Finally, the relevance of this model to individual models is statistically validated, indicating that very few information has been lost in the aggregation process.

Introduction

Air Traffic Control (ATC) is composed of several tasks that generally overlap over time. The conjunction of these, especially in heavy traffic, may result in the controller's processing capacities being exceeded. Global agreement exists for aiding air traffic controllers (ATCos) by providing them with automated tools that would perform some of the most demanding subtasks. Among these, the Conflict Detection and Resolution (CD&R) processes are traditionally singled out. This sharing of tasks in real time between operator and system requires a sufficient knowledge of experts' key processing mechanisms, so as to comply with safety and capacity target levels. In this study, we address the conflict detection process and provide a predictive model of ATCos conflict judgments.

Relevant research in the domain of conflict detection can be categorized into two fields that include automatic detection schemes based on engineering intuition (CD&R systems) as reviewed in Kuchar and Yang (2000), as well as controller centered literature that aims at understanding the cognitive processes involved in the detection task and favors the investigation of factors that affects controllers performance (Leroux, 1999) (Nunes and Mogford, 2003). The CREED project (Conflict Risk Evaluation based on Expert Detection) undertaken at the DSNA stems from these two fields as its objective is to provide an expert model of Conflict Detection, that is, a model that will have as output the expected conflict risk perceived by controller rather than a conflict/non conflict diagnosis based on objective threshold metrics as in usual CD&R systems.

To this end, a large experiment was conducted in which experienced controllers, confronted with pairs of converging aircraft in an approach environment, were asked to provide judgments about possible occurrence of a separation loss. In this article we analyze the collected data and present a statistical model that predicts the controllers' judgments for the conflict geometries considered in the experiment.

Our contribution is twofold. First, we show how information gathered from the cognitive literature on conflict detection can be exploited to describe the conflict geometry in a way that remains close to the mental representation of ATCos. Then, we show that this description can be used further to build a statistical model that performs well in predicting ATCo conflict risk judgments for the class of conflicts considered in the CREED experiment.

The paper is organized as follows. The first sections present a brief review of the existing research on conflict detection. We then detail the experimental plan and the assumptions which underpin it. The derivation of the statistical model is then presented and discussed.

Need to complete CD&R schemes

The difficulty in predicting an insufficient separation between aircraft stems from the fact that a significant amount of air traffic cannot be sorted out in advance into the two cases – conflict and non conflict – as long as the strict separation minima remain the criteria (Alliot, Durand and Granger, 2001). Indeed, a fundamental feature of ATC is that uncertainty vitiates the data required for making conflict diagnoses. This uncertainty includes a large range of different aspects: variability, ambiguity, incomplete or missing data, etc. These come from various (and varying) flight environment factors – wind shift, engine parameters setting by pilots, etc. – which result in inaccurate or even inappropriate conflict prediction. Management of this uncertainty is therefore central to the conflict detection process and appears to differ when it is processed by CD&R systems rather than by using ATCo expertise. Thresholds metrics used in CD&R systems have indeed failed until now to adequately meet results of experts' heuristics, especially as they cannot easily take into account important cost aspects such as controller's workload (Kuchar and Yang, 2000). Consequently, the subset of conflicting aircraft coming from ATCo expertise and that coming from system algorithms are not the same in general, and can not be identified with the subset of factual "losses of separation" that would have resulted from the initial conflict contingencies. There is nothing inherently surprising in this difference as, in most cases, algorithms used in "tactical collision alerting systems" ultimately aim to automate the conflict detection and resolution tasks. As such, they are only marginally interested in complying with controllers' processes.

Xu and Rantanen (2003) acknowledge, however, that data on human performance could provide "a guideline for designers to develop and improve automated conflict detection that can off-load the controller's spatial temporal cognition". So, our guess is that a step forward could be made if a conflict detection model that reproduces the judgments of controllers could be proposed. From the workload point of view, this makes perfect sense as conflict-related measures of complexity have persistently been found to predict the controller's workload or assessment of traffic complexity in a number of studies, most of which were reviewed in Hilburn (2004) (see also Kopardekar and Magyarits (2003) for a review of these variables and their integration in a "unified dynamic density metric"). If these variables, built on "objective" (as in Kuchar and Yang, 2000) conflict detections models, do well in

predicting workload, why shouldn't analogous variables built upon a "subjective" model of conflict detection perform even better[1]? A more ambitious application of such a model is provided by the ongoing European Community-funded project ERASMUS: the basic idea of ERASMUS (and so-called *subliminal control*) is that imperceptible alterations of an aircraft's speed or climb rate could be used to provide ATCos with "lucky traffic", where conflict contingencies are uncannily but "naturally" rare. Accordingly, the objective of the "subliminal control problem" as presented in Crück and Lygeros (2007) is to "minimize the level of risk *that will be perceived* by the ATCo, using only imperceptible maneuvers". This requires that the controllers' judgments for any potential conflict situation can be estimated. Resource-consuming conflict situations could therefore be identified and avoided whenever possible.

Human-centered analysis of Conflict Detection

To build a predictive model of controllers' judgments on conflict detection, it seems necessary to understand how this task is performed in real field operations and to focus on the human centered aspect of conflict detection. Various cognitive processing and strategies are in fact likely to be used by ATCos but their connections with traffic configurations to which they are applied are not clearly established (Xu and Rantanen, 2003) (Loft et al., 2007).The fundamental result on the subject is that experience acquisition by ATCos particularly results in becoming long in heuristics and short in computing within the conflict detection process (Bisseret, 1981).

This particular feature of ATC practice is responsible for the fact that a large part of conflict detection is automated. Indeed, perceptual processes - underlying detection processes - both involve heuristics and visually available data. As airspace structure and flow distribution in each sector determine specific locations and events from which conflicts are more likely to occur (Loft et al., 2007), these processes are basically interpretation ones, automatic, influenced by context and expectation (Willems et al., 1999) (Landry, 1999). A daily repeated practice and the rule-based environment of ATC (published or usual trajectories) therefore allows controllers to substantially automating the detection process (Leroux, 1999).

Also, this prevalent use of heuristics is likely to favor the development of inter-individual differences that may well have existed *ab initio*: Law et al., (1993), for example, show that men and

[1] Note that, following this idea, conflict detection measures built around a preliminary version of this work have been included in the set complexity indicators used in Gianazza and Guittet (2006a, 2006b) to model the link between workload and dynamic sectorization in the French airspace.

women differ in their ability to process velocities and distances concurrently. This explains why, when asked to estimate conflict risk, controllers put forward partly different judgments (Averty, 2005), which could, *a priori*, appear unexpected from full performance-level experts. For all that, this variability does not systematically prevent consistent judgments from emerging, since judgments also largely depend on objective features of the configuration of aircraft trajectories (Averty, 2005) (Nunes and Kirlik 2006).

Typically, separation assessment on the horizontal plane involves visual heuristics and generally calls for pattern processing (Enard, 1974) (Nunes and Mogford, 2003) (Xu and Rantanen, 2003). Since conflict detection requires controllers to foresee future behavior and positions of aircraft, a projection process (extrapolation) is assumed using both the perceptible current flight data and the individual experiences of operators (Endsley, 1995) (Xu and Rantanen 2003) (Davison 2006) (Boag 2006). Focusing on the former, the separation assessment in the horizontal plane requires the controller to evaluate the "time to conflict". While the time to the closest point of approach (CPA) is the exact information from which a possible miss distance could be inferred, Xu and al. (2006) sensibly put forward that an intermediate step could exist in such a task, leaning on the 2-D trajectory intersection point (IP), easier to perceptually extrapolate than CPA. Although no evidence exists yet to explain the method they use, it is conceivable that ATCos initially grade conflict contingencies with regard to the mere distance of each target to the IP. This point has been reported to be needed – i.e. extrapolated – by controllers (Enard, 1974) (Bisseret, 1981) (Davison, 2006). The distance of each aircraft to this point is the salient information, biasing judgments (Law et al., 1993). The closest aircraft is indeed preferentially estimated as arriving first over the IP, even for cases where the velocity ratio actually makes the furthest away come first ("distance-over-speed" bias). In fact, the speed ratio has contradictory effects, according to literature (Enard, 1975) (Xu and Rantanen, 2003) and it can be used in ATC for reinforcing a judgment that has already begun to emerge (Bouju, 1978). Furthermore, when present, acceleration is a parameter that controllers perceive with difficulty and do not frequently integrate into projected positions (Davison, 2006).

Despite being established from numerical values, and therefore needing some computation, altitude is often found to be the privileged parameter to establish conflict diagnosis (Rantanen and Nunes, 2005) (Loft et al., 2007). This diagnosis is magnified by the fact that numerous studies only involve aircraft flying levels. When vertical separation is not granted, aircraft attitude – i.e. the fact they are level, climbing or descending – seems to have strong effects (Lafon, 1978) (Lamoureux, 1999):

except for non-radar separated aircraft, i.e. flying either at different altitudes or on clearly distinct routes, conflict judgment needs information on the four dimensions (spatial and temporal). As soon as one of the involved aircraft is climbing or descending (or supposed to get this attitude in the short or medium term), judgment will depend on whether the aircraft will be found to reach the same altitude at the moment of the position overlap on the horizontal plane (Boag et al., 2006). Rates of climb/descent by themselves are seldom used by controllers, bringing about inaccurate assessment of vertical positions (Michard, 1976), except probably for aircraft of identical attitude (Bouju, 1978). Boag et al. (2006) underlined the link between complexity of conflict detection and the "number of relations" between variables from a pair of aircraft that needed to be processed to perform the diagnosis. They proposed metrics to categorize conflicts according to this relational complexity.

Other geometrical features of trajectories also have an impact on judgment accuracy (Remington et al., 2000). An increase in convergence angles generally makes judgments more inaccurate, reaching a maximum effect near 90° (Bouju, 1978) (Nunes and Scholl, 2004). According to Law et al. (1993), an interaction between aircraft speed and convergence angle also seems to exist, impacting the perceived complexity and the accuracy of judgment.

Finally, conflict processing and workload have the closest relationship in ATC (Lamoureux, 1999) (Remington et al., 2000) (Averty et al., 2004) (Loft et al., 2007). In particular, conflict classification is radically impacted by contextual workload, since it has been shown that ATCos may sort out a part of the traffic into either conflict or non conflict with the aim of managing their own workload – a phenomenon of workload homeostasis by strategy change (Sperandio, 1978). Actually, maximizing the relevance of diagnoses is time/resources consuming (Leroux, 1999) and leads to early defusing of the current conflict possibilities in order to regulate workload. This basically means increasing the number of conflict diagnoses. Thus, "diagnosing conflicts" does not uniquely depend on the configuration of the involved aircraft and their intrinsic separation probability.

Integrating doubt management in Conflict Detection

A large part of research on human-centered conflict detection focuses on the performance of the subjects, either through the accuracy of the diagnosis or by the time needed to formulate them. Some authors, however, acknowledge that conflict judgment inherently includes a doubt dimension (Enard,

1974) (Bisseret, 1981) (Leroux, 1999). Bisseret (1981) even showed that the precision of the calculations is not of prime importance for the controller. It is indeed far less detrimental for him to accept a certain degree of imprecision if it gives him better guarantees against omissions of conflict diagnosis, than to use calculations in order to reduce uncertainty to a maximum, and be "surprised" from time to time by a diagnostic error. As experience is acquired, computation processes step aside in favor of more holistic but robust processes, capable of preventing the cognitively disruptive effects of certain events.

Consequently, maintaining the doubt in his/her mental representation is very profitable to the ATCo. By deferring the decision to classify a certain traffic configuration into conflict or no conflict for as long as is reasonable, the controller minimizes the costs of an error of appreciation. Furthermore, she can more easily approach the resolution that is necessary and sufficient. This phase of doubt is fundamental in ATC – air traffic control is the “art of doubting” (Leroux, 1999). In other words, it is the management of this doubt, i.e. the decision to prolong it or, on the contrary, to suppress it, which makes up the heart of know-how in ATC. In line with Bisseret (1981) and Leroux (1999), the acquisition of expertise by controllers leads them to substitute computing and accuracy in conflict detection in favor of doubt, resulting in a more robust diagnosis, which is globally equivalent in accuracy since false responses decrease.

Our purpose was then to draw an accurate picture of the detection capacity of ATCos, in order to model their real conflict judgments. Consequently, we made the following experiment comply with four main features:

- To keep only the number of task components that could be experimentally handled. In practical terms, only realistic uncertain data were kept (from the §b above) and a single fixed geometrical configuration was considered.
- To consider a simple (pair-wise) configuration of aircraft.
- To present traffic scenarios dynamically, using devices (radar screen and maps, flight strips) peculiar to each center and flights coming from real local recordings.
- To integrate doubt expression in responses – doubt being viewed as self-assessed reliability on each individual judgment e.g. as a part of ATC expertise.

To sum up, our objective may be restated in the taxonomy proposed by Kuchar and Yang (2000): we want to build a three dimensional Conflict Detection model where state propagation is nominal (aircraft are assumed to follow their "maximum likelihood trajectory") and management of uncertainty is left to the controller. The literature on controller centered issues therefore provides a guideline for modeling in that it indicates which factors are most likely to impact judgments and how they should be modeled.

Experimentation

In keeping with the above, a set of experiments (CREED project: Conflict Risk Evaluation based on Expert Detection) was designed to evaluate the controllers' real detection capacity. It was designed so as to allow them to express their opinions on a wide range of conflicts, in order to identify the values of the variables (or sources of variations) associated with them: when does it "work" and when does it "not work", simply from the controller's point of view? The statistical aspect of the study was taken into consideration when drawing up the experimental plan, allowing the data gathered to be processed appropriately.

Construction of Conflict Scenarios

The first step was the design of scenarios for the tests. We call "scenario" a traffic sequence involving two converging aircraft, dynamically displayed on a radar screen for about one minute and fifteen seconds. Aircraft attitudes and conflict geometry were considered as variables to be controlled for and were consequently set once for all scenarios: one aircraft was stable while the other one was descending (attitudes) and we imposed a 90° crossing (geometry).

In this study we focused on conflict configurations for which the descending aircraft crossed the altitude of the stable one before overflying their IP. Situations that did not comply with this assumption were excluded from our experimental test: even if these cases may perfectly well occur in real operations, ATCos handle them by delivering up to the descending aircraft an initial safety clearance, e.g. a leveling of one thousand feet higher than the altitude of the level aircraft. This clearance is usually imperative until one of the aircraft has overflown the IP. While doing this, controllers show they systematically process them as 'certain' conflicts. By no means would responses from controllers upon these geometrical configurations have represented ATC expertise, and they were not included in our scenarios for this reason.

Once these restrictions on the conflict configurations under examination had been set, three variables (sketched in Figure 1) appeared to be necessary to characterize the scenarios. The first two variables relate to the horizontal and vertical separation, while the third one describes the temporal dimension of the conflict and is considered to be its most significant extrinsic variation.

Figure 1

1. *Ed*: the horizontal miss distance between the two considered aircraft, irrespective of their altitudes. Accordingly, *Ed* differs from the horizontal miss distance at the 3-D closest point of approach (CPA), and is generally lower unless aircraft are flying level. Therefore, a valid conflict diagnosis for low values of *Ed* requires the additional processing of the vertical dimension. We chose five levels for the construction of our traffic scenarios: 0, 3, 6, 9, and 12 nmi.

2. *Efl*: this variable was set to describe the vertical separation of aircraft. Hypothesizing that this would be closer to the actual mental representation of controllers (built from a radar display), we defined it as the longitudinal distance (in nmi) existing when the altitude difference between the two aircraft regain the separation minimum. Following Bisseret (1981), this variable attempts to reproduce the fact that conflict detection involves visual heuristics more than computing. Regardless of the strategies they actually use (Boag et al., 2006), controllers need to assess and combine vertical and horizontal projections of aircraft, as soon as one of them is climbing or descending. Quantifying what can be understood from visual effects of vertical dimension on conflict, *Efl* aims to remain close to those elements supposed to be handled mentally by controllers. As a result, it should reduce noise among the responses, therefore giving a better robustness to the mathematical model we aim to infer from them. We also selected five levels for this variable: 0, 5, 10, 15 and 20 nmi.

3. *DA*: (stands for Degree of Anticipation). This variable aims at characterizing conflict development and was defined as the longer distance to the trajectory intersection. Though both time ('prediction time span') and distance ('visual span') measures could have been selected, the distance choice reflects the authors' own experience that, as main information support (radar) shows plane or spatial information, controllers spontaneously translate times into distances: to the question "when do you think the problem will disappear?", an ATCo will typically respond "when aircraft A arrives here" and show a location on the radar screen. This choice of a distance formulation for the conflict development variable also finds supports in the literature in conflict detection as it relates to the "distance-over-speed bias" phenomenon. Of particular relevance are the experiments on arrival time judgments performed in Law et al. (1993), where a systematic "over-reliance on relative distance

information" was observed. There are six levels for this variable: 6, 10, 15, 21, 28 and 36 nmi. These values were computed at the end of the scenario display.

For each TRACON (Terminal Radar Approach Control) in which we planned to perform the experiment and for each triplet value (*DA*, *Ed*, *Efl*), we built 3 similar basic scenarios, i.e. using a different pair of aircraft each time. This way, we were able to make the callsigns, the aircraft types, the location of the IP, the magnetic orientation of trajectories, the vertical speed and the relative speeds of the involved aircraft vary - all other things being equal. The chosen number of the modalities of the three explanatory variables resulted from a trade-off between the experimental combinatory - keeping it manageable - and the search for accuracy that would benefit the mathematical model. Concrete values were inferred from recorded traffic in past study investigating conflicts from real ATC operations (Averty, 2004). Moreover, we had to take account of the fact that relationships between these variables are constrained. Indeed, since *DA* represents the distance of one of the aircraft to the IP, for low value levels of this variable (say 6 nmi), scenarios with high values of *Ed* and *Efl* could not exist. Consequently, feasible combinations allowed about 75 scenarios to be generated for each of the three pairs of aircraft, in each TRACON (Averty, 2005).

The construction of scenarios specific to each facility was motivated by our desire to reproduce the real expertise of controllers. It was therefore important for us to provide the subjects with conflict situations that could have been extracted from real local traffic and would not be perceived as artificial. Trajectories of aircraft were actually copied from files of real recorded traffic, so that aircraft did not fly with artificially constant speeds.

Procedure

Every participant in each experimental site was presented with a random set of scenarios built in the following way: for any of the 75 classes of basic scenarios (as characterized by the triplet (*Ed*, *Efl*, *DA*), one of the three basic scenarios was selected. The presentation order was itself random. For each scenario, the subject was asked to give an opinion in terms of the existence of the separation minima (3 nmi or 1000 ft). Overall duration for each participant was less than one hour, as judgments were often formulated before the end of the available 1'15" of each scenario.

During the instructions given before the experiment started, it was stressed that the aim was to express simply "what your eyes would perceive", referring only to the separation minima. The aim, as indicated above, was to minimize the systematic and variable addition of margins linked to setting up a resolution. The effect of the "natural" coupling of the two processes - detection and resolution - is to produce "noise" in the replies, and to produce judgments that dissimulate the detection process model itself. In order to achieve this, it was suggested to the controller, while she was being given the test instructions, that she should not project herself into the role of the "controlling controller", but rather into that of a tester, or a controller in a mirror position. So the controller was invited to be concerned about the existence (or otherwise) of the separation minima, but without being able to act on it, or to resolve it.

Figure 2

Controllers' judgments had to be expressed on an 8-degrees scale on a tactile display (see Figure 2). From the analysis of conflicts in a past study (Averty, 2004), we postulated that a 8-degrees scale should reflect the highest differentiation an expert could express in a conflict diagnosis. Possible responses were ordered from "non-conflict sure" to "conflict sure" and presented in two groups: the four on the left stood for separation-loss judgments (conflict) whereas the four on the right meant that sufficient separation was the most probable outcome (non-conflict). Thus, each group conveyed four levels of certainty so that subjects could self- assess while selecting each one (the lower the degree of doubt, the darker the key).

The traffic scenarios were submitted to 161 controllers coming from four different TRACONs, each time using the usual local devices and radar maps. The choice of TRACON controllers was only governed by opportunities for experiment deployment at that time. The overall sample was composed of 75.2% male (mean age 44.8) and 24.8% female (mean age 40.6). All of them had to be fully qualified for at least two years, but the mean ages (above) show that most were considerably more experienced (generally more than 10 years).

Statistical Models

Introduction

In Averty (2005), a descriptive statistical analysis of the collected dataset was presented. Though judgments formulated for given conflict configurations were sometimes found to be very heterogeneous across controllers, the global consistency of responses appeared remarkably high, as increases in the horizontal or vertical separation of aircraft systematically resulted in a safer "median" judgment. Moreover, the graphical analysis of the data also showed that controllers managed uncertainty by expressing doubtful responses and thus confirmed the importance of doubt in the conflict detection process. The data therefore supported Leroux (1999) in its statement that air traffic control is the "art of doubting". These results suggested that the sources of variations identified at this stage of the study (*Ed, Efl* and *DA*) could be used further to develop a predictive model of conflict detection and that the collected data could actually be used to calibrate it.

The following section presents the derivation of the statistical model and discusses the estimated parameters.

The Ordered Logit Model

The air traffic controllers judged the risk of collision on an ordered scale of eight levels, going from "airprox certain" (separation between aircraft will be lost), through "caution", to "no problem". The Ordered Logit Model provides a common and convenient framework for analyzing such data in which the dependent variable is both discrete and ordered (Greene, 2000). It is based on a natural way of representing the decision-making process, consisting of defining a continuous "latent variable", U, whose level determines the subject's final reply. If the value of U is less than a certain threshold, noted α_1, the reply will be "airprox certain"; if the value of U is higher than α_1, but less than another threshold α_2, the reply will be the next one up on the scale of possible replies, and so forth.

As the latent variable *U* is not observable, it must be reconstructed from the observable variables *Ed, Efl* and *DA*. In the specific context of the CREED experiment, these variables are supposed to give the best explanation of the controller's reply. However, due to its particular nature, the variable *DA* will be treated differently from the others and we will build a model for each value of *DA*. An "integrated" model will be obtained later through the interpolation of these models. Other sources of variations may affect the controller's judgment, but though they cannot be or were not observed in the experiment, they must nevertheless be incorporated into *U*, which must be broken down into an observed component and a non-observed component:

$$U = \beta' x + \varepsilon.$$

Usually, the non-observed component ε is considered as a noise, thus as a random variable, whose law determines the probability of the replies; here we used a logistical law[2]. The observed component is generally modeled by a linear combination of the explanatory variables; for a fixed value of *DA*, it will therefore be formalized by the linear relation:

$$\beta' x = \beta_1 \times Ed + \beta_2 \times Efl + \beta_3 \times Ed \times Efl$$

where the unknown coefficients β_1, β_2, and β_3, are called regressors; note the presence of the term *Ed x Efl* which captures the crossed influences between the two variables *Ed* and *Efl*.

Ordered Logistic models for each *DA*

For each given *DA*, the model is determined by ten coefficients that need to be estimated. The method used for the estimation is based on the maximization of the log-likelihood, which depends on the data observed and the coefficients; an optimization algorithm is then used. The estimated parameters are shown in Table 1.

---

[2] The distribution function of a logistical law is $F(t) = \Pr\text{ob}(\varepsilon \le t) = \exp(t)/(1+\exp(t))$. Therefore, the probability that the reply is the discrete value according to the interval] α_i, α_{i+1}] is

$$\Pr\text{ob}(\alpha_i \le U \le \alpha_{i+1}) = \frac{\exp(\alpha_{i+1} - \beta' x)}{1+\exp(\alpha_{i+1} - \beta' x)} - \frac{\exp(\alpha_i - \beta' x)}{1+\exp(\alpha_i - \beta' x)}.$$

	DA=6	DA=10	DA=15	DA=21	DA=28	DA=36
α_1	1,506	0,862	-0,040	-0,704	-1,376	-1,793
α_2	2,399	1,898	1,124	0,525	-0,084	-0,237
α_3	2,840	2,316	1,668	1,078	0,589	0,413
α_4	3,181	2,653	2,155	1,564	1,239	1,157
α_5	3,733	3,185	2,751	2,215	1,850	1,804
α_6	4,050	3,603	3,255	2,826	2,346	2,376
α_7	4,747	4,299	4,143	3,595	3,137	3,178
β_1	0,232 ✶✶	0,405 ✶✶✶	0,315 ✶✶✶	0,281 ✶✶✶	0,188 ✶✶✶	0,231 ✶✶✶
β_2	0,660 ✶✶✶	0,400 ✶✶✶	0,236 ✶✶✶	0,158 ✶✶✶	0,097 ✶✶✶	0,077 ✶✶✶
β_3	-0,036 ✶✶✶	-0,026 ✶✶✶	-0,006	-0,003	0,001	-0,002

Table 1: Estimation of the ordered logistic model parameters:
(Significance codes: 0 '✶✶✶' 0.001 '✶✶' 0.01 '✶' 0,05 '●' 0,1 ' ' 1)

The Ordered Logit Model belongs to the class of Generalized Linear Models: a linear combination of explanatory variables (the latent variable *U* in our setting) is nonlinearly linked to the dependent variable. Interpretation of the parameters of the model is therefore analogous to the interpretation in linear models, provided the link between the "latent variable" and the dependent variable is well understood. In our model, an increase of *U* implies that more risk is perceived by the controller.

First, one may observe that β_1 and β_2 are always positive and significant. This supports a very natural intuition: the direct effect of an increase in the separation minima translates into a decrease in the perceived risk. However, one needs to consider the impact of the cross-product parameter β_3: for values of *DA* greater or equal to 15, β_3 is not significant and its point estimate is very low. It is thus unsurprising that we actually observe a positive direct effect of *Ed* and *Efl* on the latent variable for reasonable conflict geometries. Moreover, the fact that the cross influence of *Ed* and *Efl* for these conflict geometries is low supports the idea that, when the time pressure is low, controllers have a rough estimate of the conflict geometry: any increase in vertical or horizontal separation directly translates into a decrease in the perceived risk.

When *DA* equals 10, β_3 is negative and significant but remains small enough for the direct effect of *Ed* and *Efl* on the latent variable to stay positive: the perceived risk still decreases when the conflict geometry becomes less dangerous. However, the negative β_3 indicates that actual separation somehow begins to matter. For example, the perceived risk decreases less for a given increase in *Ed* (resp. *Efl*) when *Efl* (resp. *Ed*) is higher. That is, when the prediction time span is short and separation

of aircraft is ensured vertically (resp. horizontally), an increase in horizontal (resp. vertical) separation does not reduce the perceived as much as it would have in the case of non-separated aircraft. This, again, is a very intuitive result.

Finally, when *DA* equals 6, β_3 is negative and significant but becomes large enough to induce a negative direct effect of *Ed* and *Efl* on the perceived risk. This happens for *Ed* greater than 6.44 or *Efl* greater than 18, that is once separation, given the very short prediction time span, is obvious: indeed, we observe that for these values of (*Ed*, *Efl*), the median response regarding the perceived risk is the safest.

Figure 3

Another interesting fact about estimated parameters is that the relation β_1/β_2 varies linearly as a function of *DA* as shown in Figure 3. This relation provides information on the variable that is the "most used" during the "conflict/no conflict" diagnosis: in an extreme case, if it was equal to 0 (respectively infinity), the controller would only base his judgment on *Efl* (respectively *Ed*). As these two variables are expressed in the same unit and describe horizontal distances, it can be estimated that when the relation β_1/β_2 equals 1, *Ed* and *Efl* are evenly discriminating in the controller's judgments. This fact tends to prove that, in our setting, ATCos privilege *Efl* for *DA* less than 12nmi whereas *Ed* becomes dominant for *DA* over 12nmi. This finding is consistent with previous results on conflict detection: according to Loft et al. (2007) and the studies quoted therein, ATCos indeed “appear to prefer using altitude to heading and speed information when determining the likelihood of aircraft conflict” and seem to use “lateral separation only in circumstances where vertical separation is questionable”. The scenarios proposed in this study display conflict situations in which one aircraft is descending and crossing the other's altitude. Since separation is first granted vertically (as explained in the experiment design section, the other types of conflict are avoided as much as possible in real life operations), it is natural to recover the dominance of the vertical dimension for small *DA*, that is when aircraft have similar altitudes and the rate of descent may actually be used (Bouju, 1978). For high values of *DA*, the vertical separation is instead questionable since vertical positions of the descending aircraft are

inaccurately projected (Michard, 1976). ATCos have therefore to rely primarily on planar separation for these large *DA*.

Figure 4

Using these coefficients, we can then calculate for each value of *Ed* and *Efl* the probability of observing a given reply and compare these predictions with the observations. Such a comparison is provided in Figure 4, where probabilities for each response are presented for the pair *Ed* = 0 and *Efl* = 0. Answers are ordered by perceived risk and with black (resp. dark gray) standing for high (resp. low) perceived risk. The agreement between the model and the data is very good for each value of *DA*.

Integrated model

So far we have constructed logit models for each of the values of explicative variable *DA*. In order to be able to interpolate, it becomes necessary to integrate this variable into a more general model, which we will call "integrated model". We opted for the approach depicted in Figure 5: first, the variable *DA* being given, the regressors and thresholds have to be estimated. This will in turn allow the observed part of *U* to be computed for fixed values of *Ed* and *Efl*. The expression of the noise distribution function will then be used to compute the probabilities for each reply.

Figure 5

We therefore need to evaluate the seven thresholds and the three regressors, for *DA* values that are different from those already considered. An appropriate technique to handle this problem is to use non linear regression: after having established as many ordered logistical models as values of *DA* (*DA*=6, 10, 15, 21, 28, 36), we have a sample of six values for each parameter. Non-linear regression would then allow each of the parameters to be expressed as a function of variable *DA*. These functions have to be chosen *a priori* in a certain family: the better the family choice, the closer the fit of the parameters. So, the first step needed to build the integrated model consists in determining the form of the 10 parameters as a function of *DA*. To this end, we used the samples for each parameter

we obtained by computing the six ordered logit models, one for each *DA*. We selected, *a priori*, a common model for the parameters:

$$
\begin{aligned}
&\alpha_i(DA) = \psi(DA, A_i^{\alpha}, B_i^{\alpha}, C_i^{\alpha}),\\
&\beta_2(DA) = -\psi(DA, A_2^{\beta}, B_2^{\beta}, C_2^{\beta}),\\
&\beta_3(DA) = \psi(DA, A_3^{\beta}, B_3^{\beta}, C_3^{\beta}),\\
&\beta_1(DA) = \left[A_1^{\beta} + B_1^{\beta} \times DA\right] \times \beta_2(DA)\\
&with\\
&\psi(DA, A, B, C) = A + (B - A)\exp\left(-\exp(C) \times DA\right)
\end{aligned}
$$

This choice for the family of non-linear functions was made because they provide a parsimonious way to model monotonous responses that have a finite value in *DA=0* and an horizontal asymptote at *DA*□∞, that is properties that are satisfied by the seven thresholds and two of the three regressors (β_2, β_3). As the last regressor (β_1) could be easily parameterized using the linear relationship described in the previous section, this specification of the model was expected to perform well on our data.

Non-linear regressions gave a first numerical estimation for each of the 29 parameters of the model and a graphical inspection of the results confirmed the relevance of these choices. It also gave some further information on the form of the model: in particular, the coefficients C_i^{α} were close enough to be constrained to be equal, as could the coefficients C_i^{β}. These restrictions imply that thresholds on the one side and β_2 and β_3 on the other side converge at the same speed to their respective asymptotes. Lastly, the coefficient A_3^{β} was nearly equal to zero and we decided to assume it was. This imposes that β_3 is asymptotically null for large values of *DA*. Given that β_3 was found to be statistically significant for *DA* greater or equal 15 nmi, this assumption is actually very natural. Under these assumptions, the integrated model depends on 21 parameters.

However, though it gives a good insight into the relevance of the family functions, performing independent regressions is best considered as a first step in the modeling process. Indeed, the initial dataset is only used through the estimated parameters. Then if a single parameter is estimated using 6 observations rather than the 10385 that could be potentially used, it is reasonable to think that some

information has been lost in process. More importantly, the impacts on the latent variable of the departure from point estimations (for the available *DA* values) are not likely to compensate for one parameter to another and may lead to inconsistent estimate. A way to circumvent these problems could be to perform a global weighted regression of the parameters (using the computed covariance of the parameters as weights), but it might be preferable for simplicity to maximize a global log-likelihood constructed on the whole data sample. Note that the latter method presents the great advantage to work with non-concentrated values of *DA* (and thus would extend to other variables standing for the prediction time span, one being suggested [Stankovic and al., 2006]).

Surely, even if they are estimated on a large dataset, 21 parameters remains a large number. However, it should be acknowledged that a large part of these parameters relate to the thresholds, the number of which was "imposed" by the number of response categories that were offered to the controllers during the experiment. The latent variable, which formally describes how the perception of the geometry evolves with the prediction time span, needs only 6 parameters to be described. 3 more parameters would be needed to calibrate the model if we were only interested in the binary choice "conflict/no-conflict". The supplementary parameters describe the data with more richness as they allow the whole set of judgments to be recovered and therefore adequately reproduce doubt.

In this paper, the final model was obtained by expressing the integrated log-likelihood and numerically maximizing it. To facilitate convergence, we used the parameters obtained by non-linear regression as starting values for the optimization algorithm. The resulting curves are displayed in Figure 6.

Figure 6

Finally, it is of interest to investigate the loss of accuracy of the integrated model for the 6 values of *DA* for which we know the optimal parameters. This can be assessed by testing for each *DA* the null hypothesis that the local parameters (the parameters obtained by maximizing the log-likelihood for a given *DA*) are equal to the integrated parameters (the parameters obtained though the integrated model). We therefore ran a series of 6 likelihood ratio tests (see e.g. [Greene, 2000]). The results are presented in Table 2, where LVG (resp. LV) stand for the integrated (resp. local) log-likelihood.

DA	*6*	*10*	*15*	*21*	*28*	*36*
LVG	-828.08	-1308.74	-2374.87	-3369.98	-3906.24	-4002.14
LV	-823.71	-1305.09	-2370.92	-3364.29	-3901.13	-4001.57
p-value	0.56	0.7	0.64	0.33	0.42	1

Table 2: Comparison between the model LVG with the model LV by way of the log-likelihood test.

These p-values[3] indicate that the null hypothesis may never be rejected at any usual significance level. This means that the loss of accuracy induced by the interpolation of the parameters is very low. Hence, there is nothing in the data that would cause rejection of the integrated model. This robustness supports the idea that the model could be used as a building block to design ATC support tools.

Conclusion

In this article, we have presented the results of an experiment conducted on a large number of experienced controllers that would make it possible to derive a “controller-like” model of conflict detection.

With regard to the experimental plan, a particular effort was put into the design of traffic scenarios and the definition of the explanatory variables. As these were supposed to reproduce the perceptual heuristics used by controllers in their actual detection process and therefore often rely on a planar representation of the conflict geometry, they differ from the ones that would typically be used in engineering approaches, such as those reviewed in Kuchar and Yang (2000). However, they are, as explained, relatively well supported by the cognitive literature on conflict detection. As for the statistical model to be constructed, a key requirement was that it should properly account for doubt. Doubt, indeed, is central in the work of ATCos and a better understanding of traffic configurations where it occurs could therefore be of some interest in the design of operational tools. In this aspect, the European project ERASMUS provides a concrete example of potential application.

[3] The null hypothesis is not rejected as soon as the p-value is greater than the significance level α, e.g. α=0.05.

A graphical analysis of the data, presented in Averty (2005), showed that judgments exhibited a high degree of variability but were “globally” consistent, in the sense that any increase in the horizontal or vertical separation of aircraft systematically resulted in a safer “median” judgment. As variability was mainly observed for conflict situations for which judgments were not definite, the author concluded that the data adequately reproduced doubt and could be used to build a controller-centered model of conflict detection. The derivation of such a model is presented in some detail in the present article. In the first part of the statistical analysis presented here, we estimated Ordered Logit Models for given values of *DA*, our variable describing the temporal dimension of the conflict situation. The use of this type of model is very natural given the discrete and ordered nature of the expressed judgments, but also presents the great advantage of inherently accounting for doubt. Comparison of the model parameters across *DA* indicated that the relative importance of the horizontal and vertical separations evolved with conflict development, with the horizontal separation dominating away from the conflict location, and the vertical dominating close to it. In the second part of the statistical analysis of the data, an “integrated” Ordered Logit Model was derived using maximum likelihood estimation and a proper family of functions. The integrated model was then statistically validated on the “fixed DA models” and we proved that no significant loss of accuracy had occurred in the derivation process.

As of now, the model presented in this article may be used to predict controllers’ judgments, including doubt, for the particular class of conflict geometries considered in the experiment. It is a very supportive result in our research effort to provide a general model of controllers’ judgments, but much work remains to be done if one wants to generalize the model to other conflict geometries. Some factors (convergence angle, contextual workload, for example) that impact the conflict/non conflict diagnosis have been kept constant in the experiment and the robustness of our model should be tested against these sources of variations. We can expect, for example, our model to be “conservative” for non-perpendicular crossings, as it is often shown that controllers’ judgment is the least precise when the convergence angle is 90° (e.g. in Law and al. (1993)). However, if drastic changes rather than small variations were to be imposed on the conflict geometry, it is possible that other explanatory variables would need to be considered to describe the conflict. Indeed, the ones defined in this study remain close to the perception of the controller for the range of approach conflict geometries considered and it is unclear whether they generalize well or not. In this direction, the

recent findings of Nunes and Kirlik (2006) suggest that good agreement between controllers' judgments and "objective" conflict probabilities may exist. The judgments collected for this study could therefore be used to calibrate more general models of conflict detection, such as the one introduced in Cruck and Lygeros (2007). Note that these potential extensions require new data to be gathered and a new experiment, which we are currently designing.

Finally, from a human-factor perspective, it would also be of great interest to evaluate the heterogeneity of controllers. Variability of judgments was broadly discussed in Averty (2005) and is partly accounted for in the Ordered Logit Model presented in this paper, but could be further investigated. This is the purpose of ongoing research in the CREED project.

Acknowledgements: We are deeply indebted to the editor and the anonymous referees for their very helpful comments and suggestions. We are also grateful to Esa Rantanen for having contributed to a greatly improved presentation of this paper. All the statistical computing and graphics have been obtained with the free software R (R Development Core Team, 2006).

References

Averty P., Collet C., Dittmar A., Athenes S., Vernet-Maury E. (2004) "Mental workload in air traffic control: an index constructed from field tests" *Aviation, Space and Environmental Medicine*, Vol.75 (4), pp. 333-341.

Averty, P. (2005), "Conflict perception by ATCS admits doubt but not inconsistency", *Proceedings of the 6th Air Traffic Management Research & Development Seminar*, Baltimore, USA.

Alliot J.M., Durand N., Granger G. (2001), "A statistical analysis of the influence of vertical and ground speed errors on conflict probe", *Proceedings of the 4th Air Traffic Management Research & Development Seminar*, Santa Fe (USA).

Bisseret, A., (1981) "Application of signal detection theory to decision making in supervisory control: the effect of the operator's experience", *Ergonomics*, 24, (2), pp. 81-94.

Bisseret, A., Girard Y. (1973) *Une expérience sur le "choix des solutions" dans le contrôle du trafic aérien*, Rapport IRIA C.O.7301 R35.

Boag C., Neal A., Loft S., Halford G.S. (2006), "An analysis of relational complexity in air traffic control conflict detection task", *Ergonomics*, Vol. 49, No. 14, pp. 1508-1526.

Bouju, F. (1978) *Analyse de la prévision de la séparation entre deux avions évolutifs convergents*, Rapport IRIA C.O.7802 R54.

Crück, E., Lygeros, J., (2007) "Subliminal air traffic control: Human friendly control of a multi-agent system", *Proceedings of the American Control Conference, 2007 - ACC '07*, pp.462-467.

Davison, H.J. (2006), *Modeling the Air Traffic Controller's Cognitive Projection Process*, Report No. ICAT-2006-1, Massachusetts Institute of Technology, Cambridge, USA.

Enard, C.J. (1974), *Analyse de la prévision de séparation entre deux avions convergents sur l'écran radar,* Rapport IRIA C.O.7405 R44.

Enard, C.J. (1975), *Le degré de certitude dans l'estimation de la séparation entre deux avions convergents*, Rapport IRIA C.O.7507 R46.

Gianazza, D., Guittet, K. (2006a), "Evaluation of air traffic complexity metrics using neural networks and sector status", *2nd International Conference on Research in Air Transportation*, Belgrade, Serbia.

Gianazza, D., Guittet, K. (2006b), "Selection and evaluation of air traffic complexity metrics", *Proceedings of the 25th Digital Avionics Systems Conference*, Portland, USA.

Greene, W.H. (2000), *Econometric Analysis* (4th edition), Prentice-Hall International.

Hilburn, B. (2004), *Cognitive Complexity In Air Traffic Control – A Literature Review*, Note No. 04/04, Eurocontrol Experimental Centre, Brétigny, France.

Kopardekar, P., Magyarits, S., (2003) Measurement and prediction of dynamic density, *5th USA/Europe Air Traffic Management R&D Seminar,* Budapest, Hungary

Kurchar, J.K., Yang. L.C.,(2000), A Review of Conflict Detection and Resolution Modeling Methods, *IEEE Transactions on Intelligent Transportation Systems*, Vol.1, No. 4, pp.179-189.

Lafon, M.T. (1978), *Observations en trafic réel de la résolution des conflits entre avions évolutifs*, Rapport IRIA C.O.7802 R55.

Lamoureux, T. (1999), "The influence of aircraft proximity data on the subjective mental workload of controllers on the air traffic control task", *Ergonomics*, 42, 1482–1491.

Landry, S.J. (1999), *Perceptual grouping in Air Traffic Control*, Thesis for Master of Science in Aeronautics and Astronautics, Massachusetts Institute of Technology, Boston, USA.

Law, D. J, Pellegrino, J. W., Mitchell, S. R., Fischer, S. C., McDonald, T. P., and Hunt, E. B. (1993), "Perceptual and cognitive factors governing performance in comparative arrival-time judgments", *Journal of Experimental Psychology: Human Perception and Performance,* 19(6), 1183-1199.

Leroux, M., (1999), "Cognitive Aspects and Automation", *Proceedings of the first USA/Europe Air Traffic Management R&D Seminar*, Orsay (France).

Loft S., Sanderson P., Neal A., Mooij M. (2007), "Modeling and Predicting Mental Workload in En Route Air Traffic Control: Critical Review and Broader Implications", *Human Factors*, Vol. 49, No. 3, pp.376-399.

Michard, A. (1976), *Analyse de la représentation des trajectoires d'évolution chez les Premiers Contrôleurs*, Rapport IRIA C.O.7611 R50.

Nunes, A., Kirlik A., (2005), "An Empirical Study of Calibration in Air Traffic Control Expert Judgment", *Proceedings of the 49th* Human Factors and Ergonomics Society, pp. 422-426(5).

Nunes, A., Mogford R.H., (2003), "Identifying Controller Strategies that Support the Picture", *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society,* Santa Monica, CA, USA.

Nunes, A., Scholl, B.J., (2004), "Interactions between Convergence Angle, Traffic Load, and Altitude Distribution in Air Traffic Control", *Proceedings of the 3rd Annual Conference on Human Performance, Situation Awareness and Automation Technology*, Daytona Beach, USA, March.

R Development Core Team (2006). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rantanen, E.M., Nunes, A., (2005), "Hierarchical Conflict Detection in Air Traffic Control", *International Journal of Aviation Psychology*, Vol. 15, No. 4, Pages 339-362

Remington, R. W., Johnston, J. C., Ruthruff, E., Romera, M., Gold, M., (2000), "Visual Search in Complex Displays: Factors Affecting Conflict Detection by Air Traffic Controllers", *Human Factors*, Vol. 42(3), pp 349-366.

Sperandio, J. C. (1978), "The regulation of working methods as a function of workload among air traffic controllers", *Ergonomics*, 21(3), 195-202.

Stankovic, S., Raufaste, E., Averty, P. (2007) "Determinants of Conflict Detection: A Model of Risk Judgments in Air Traffic Control", submitted to *Human Factors*, March 2007.

Willems, B., Allen, R.C., Stein E.S., (1999) *Air Traffic Control Specialist Visual Sanning II: Task Load, Visual Noise, and Intrusions into Controlled Airspace*. Report DOT/FAA/CT-TN99/23; W.J. Hughes Technical Center, Atlantic City, USA.

Xu, X. and Rantanen, E.M. (2003), "Conflict detection in air traffic control: a task analysis, a literature review, and a need for further research", *Proceedings of the 12th International Symposium on Aviation Psychology*, Dayton, USA.

Xu, X. Rantanen, E.M., Huo, Y., (2006), *Relative Importance of Conflict Geometry Variables in Influencing Pilots Conflict Detection Using a Cockpit Display of Traffic Information*, Technical Report HFD-06-02, University of Illinois at Urbana-Champaign, USA.

Biographies

Philippe Averty is Confirmed Expert (PhD in Cognitive Psychology) with the DSNA/DTI/R&D. He first worked as an Air Traffic Controller and Instructor at Lyon TRACON and Paris ACC. He is now involved in different projects including workload assessment and decision modeling and collaborates to the European project ERASMUS.

Kévin Guittet is researcher in the DSNA/DTI/R&D. He holds a PhD in Applied Mathematics and a Master Degree in Economics. His research interests include air transport economics and statistical modeling applied to airspace complexity.

Pascal Lezaud received his Diploma of Electronics Engineer from l'Ecole Nationale de l'Aviation Civile, Toulouse France, in 1982. He joined the DSNA/DTI/R&D in 1992, in which he was involved in dependability studies for ATM systems. He obtained his PhD in Probability in 1998. Now, his main interests include collision risk models, probability, Markov processes and simulation of rare events. He collaborated to the European Research project "Hybridge" and participates to the new project "iFLY".

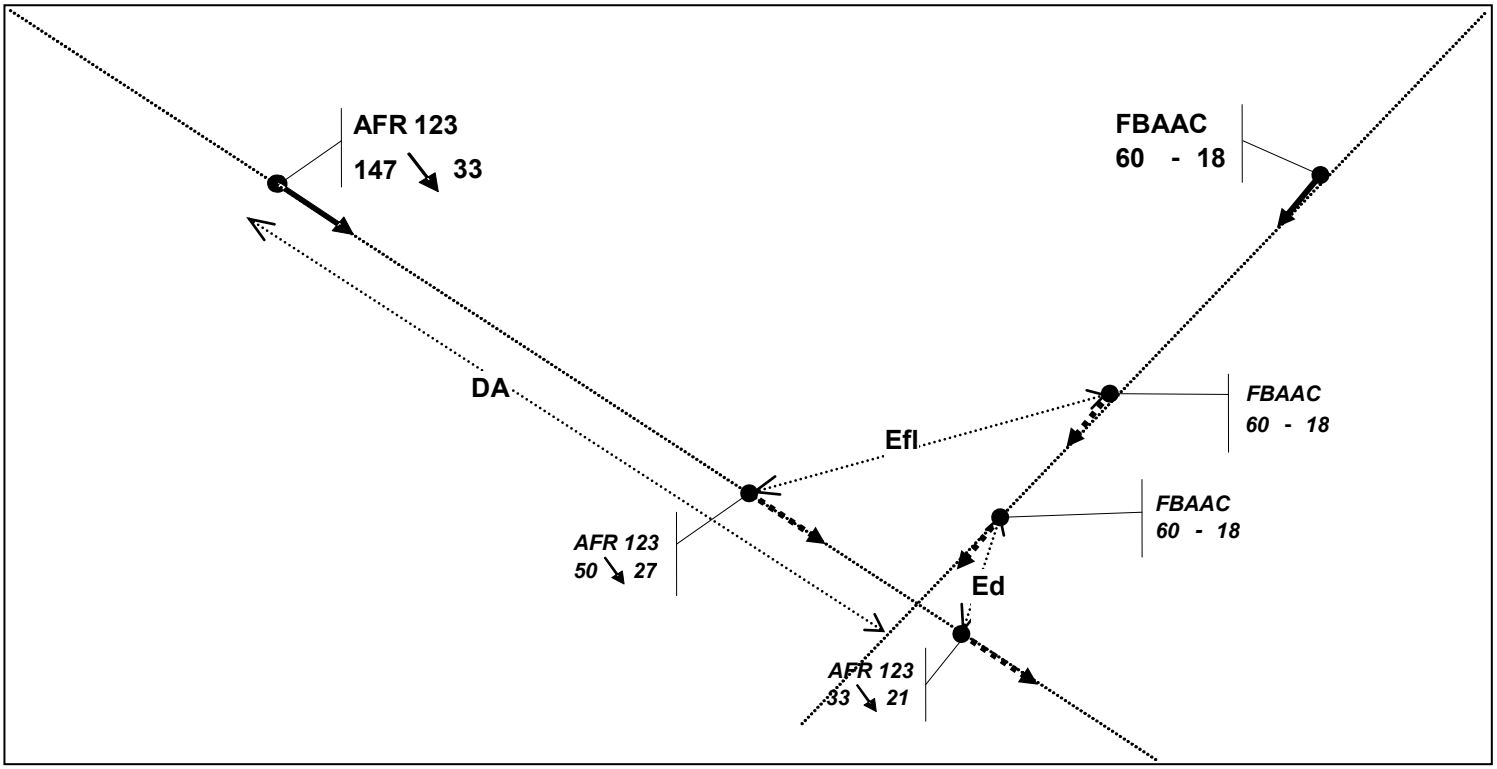

Figure1: Illustration of the three explanatory variables as they were used in the traffic scenarios. Two aircraft (here: AFR123, passing through 14700ft and FBAAC flying level at 6000ft) were dynamically shown on radar during 1 minute and 15 seconds. Before the end of this period, i.e. when the value of DA was computed, the subjects had to express their judgments about the conflict risk. These judgments also related to the values of *Ed* and *Efl*. The positions of the aircraft, where *Ed* and *Efl* are computed, are only displayed here for clarity purpose.

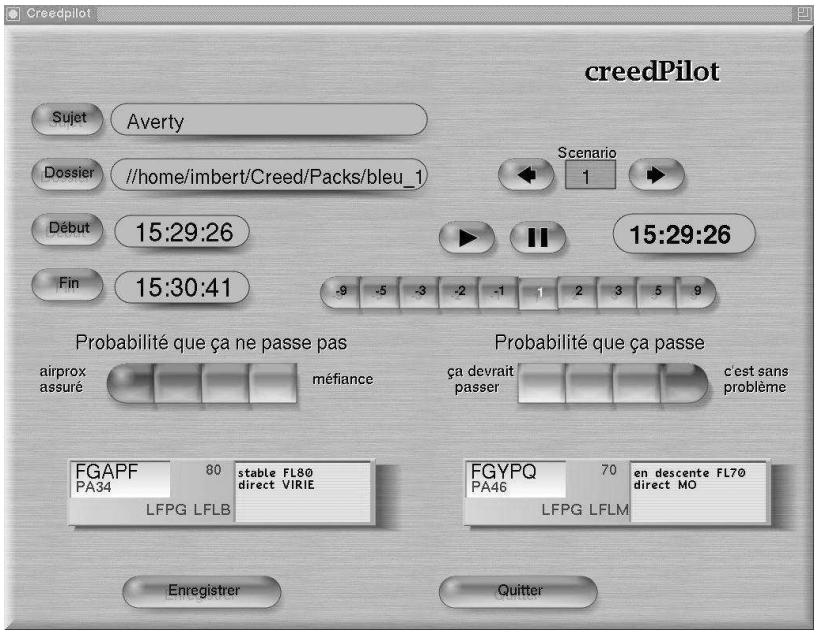

Figure 2: The tactical display used to collect the responses. Two groups of four keys were available for the subjects on the interface, ranging from “conflict sure” on the left to “non conflict sure” on the right. For each scenario, ATCos had to select the key that reflected at best their judgment.

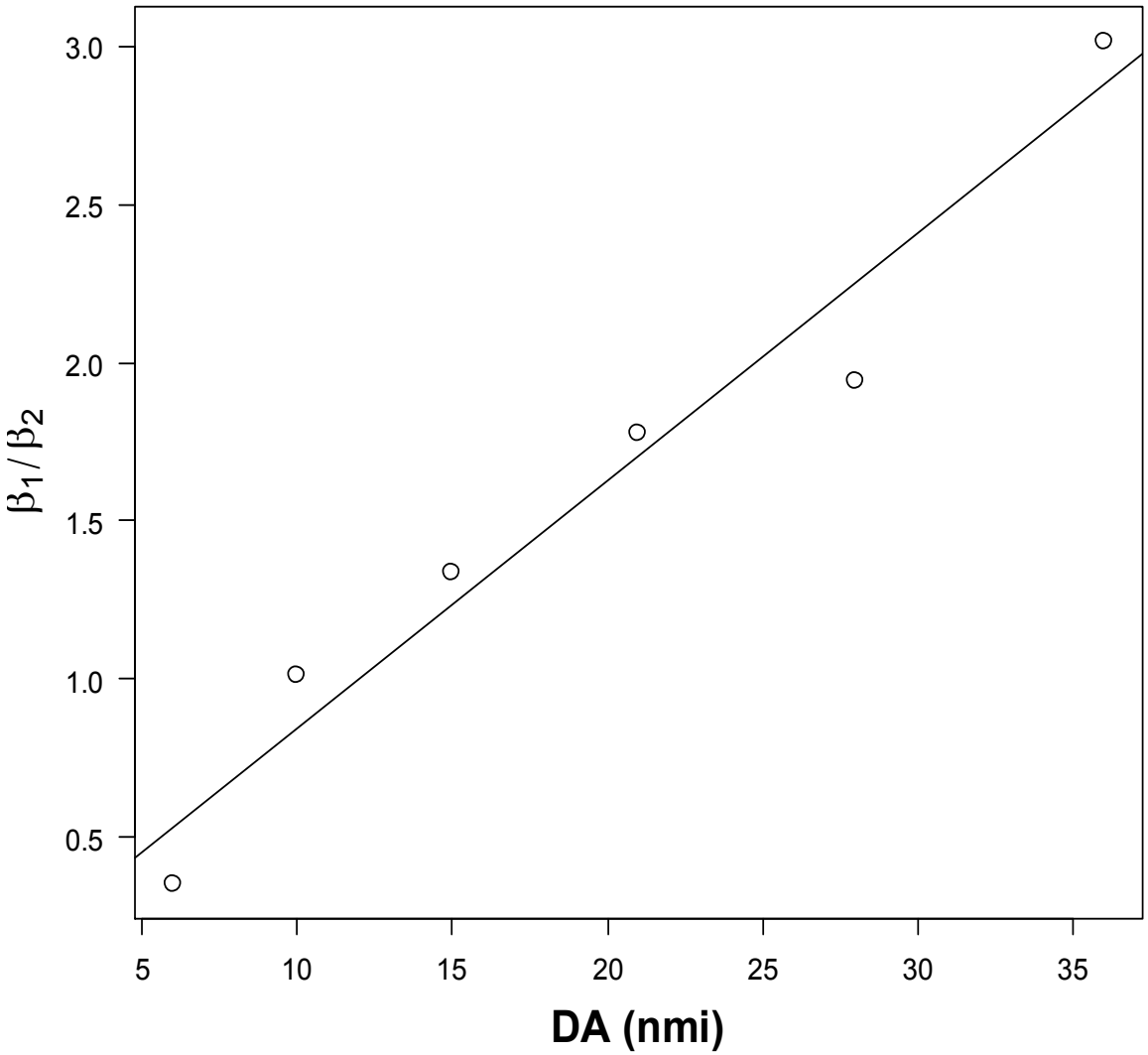

Figure 3: Variation of the ratio β_1 / β_2 versus *DA* (nmi): estimated ratios and fitted linear relationship.

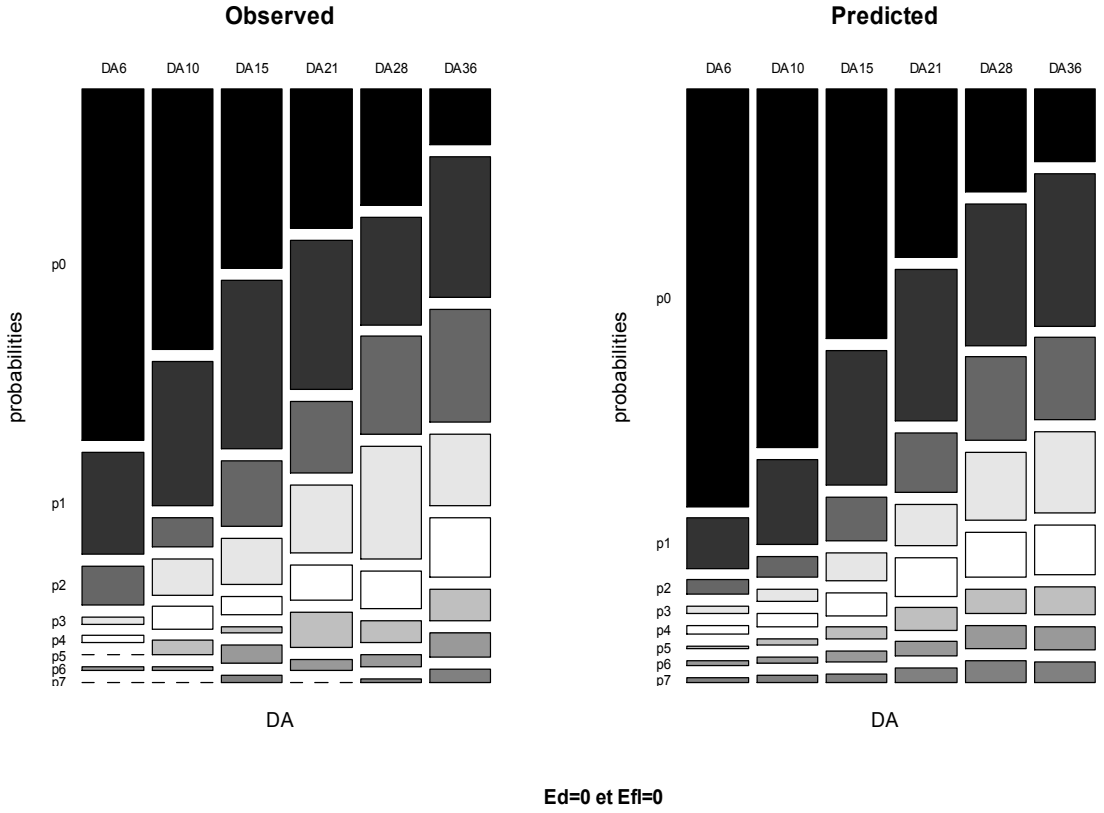

Figure 4: Comparison between observed data and predicted data for *Ed=0* and *Efl=0*.

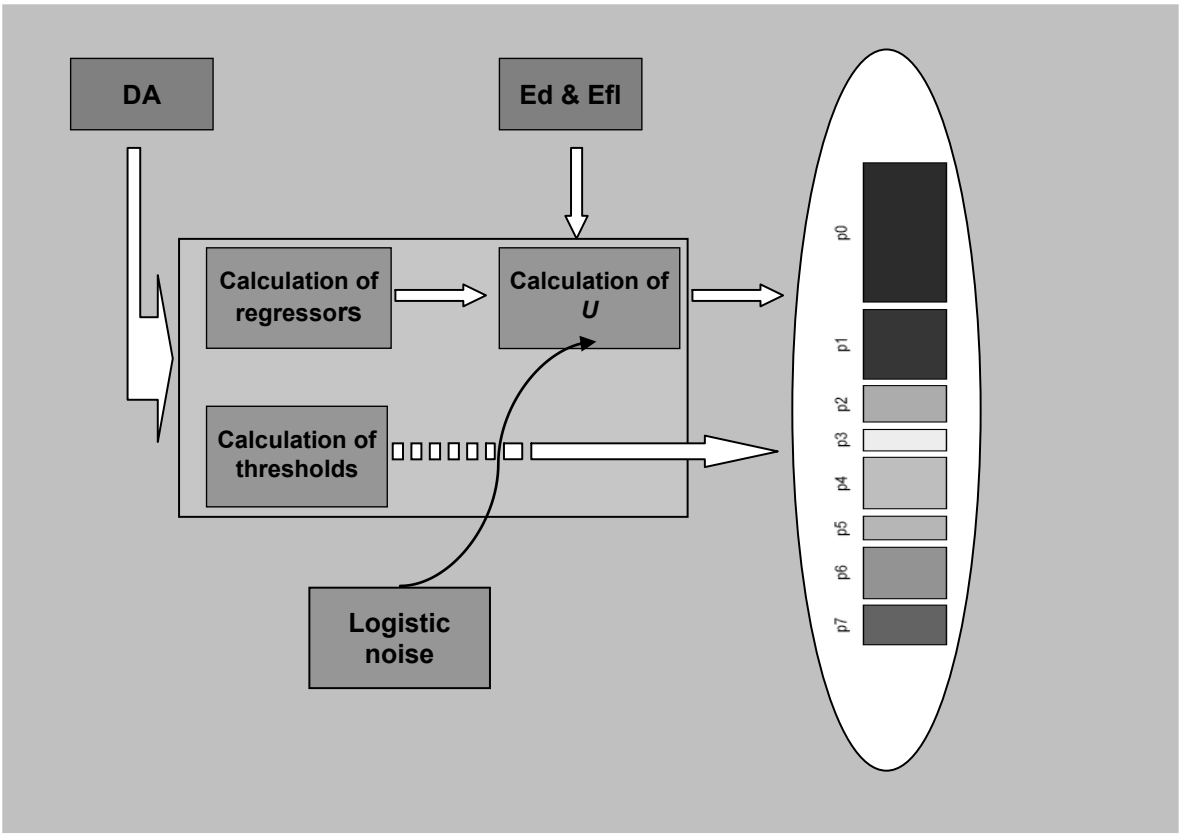

Figure 5: Integrated approach: as *DA, Ed* and *Efl* are fixed, the regressors and thresholds will be calculated, hence the value of the observed part of *U*. The noise distribution function allows calculation of the probabilities for each reply.

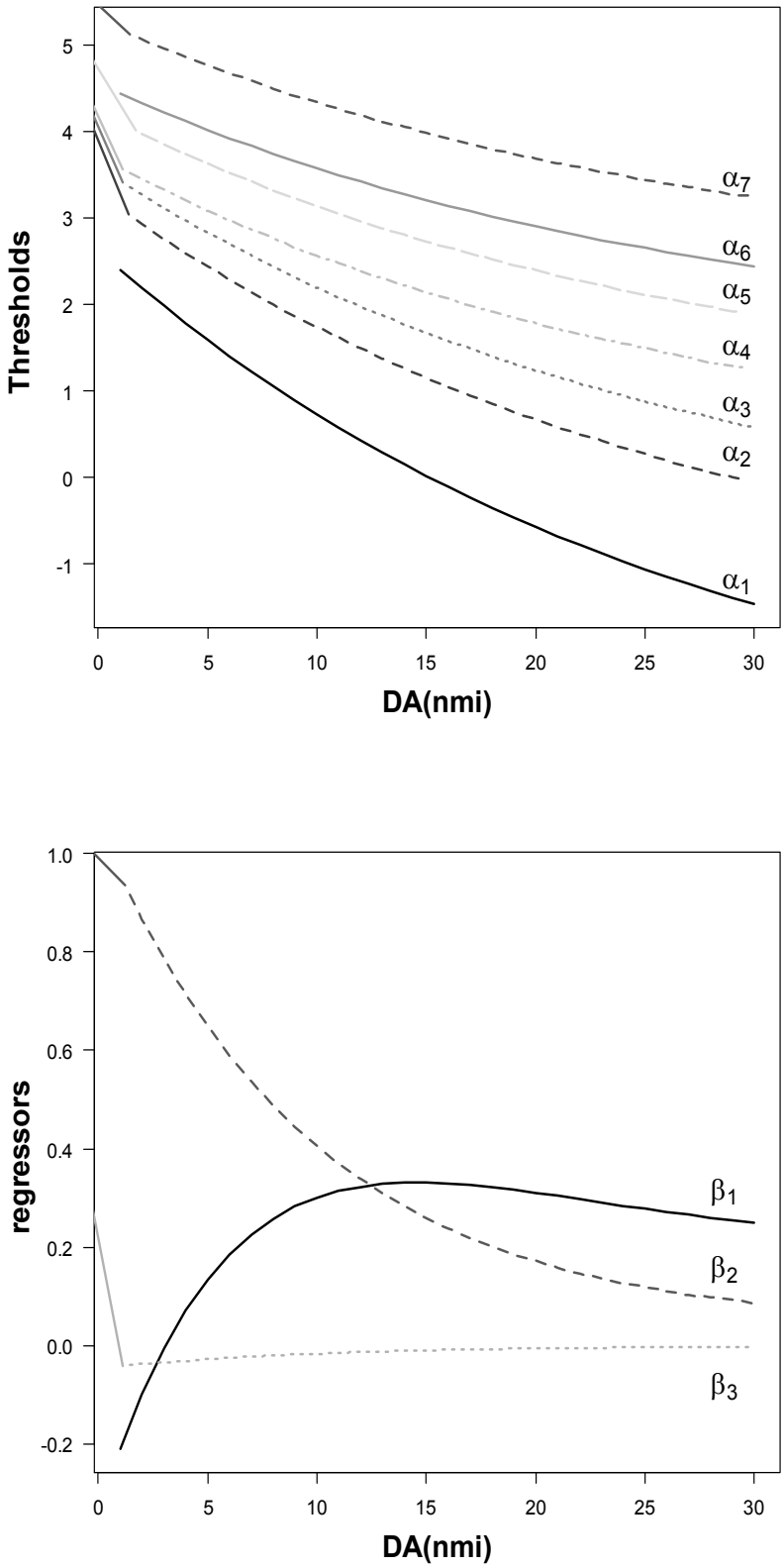

Figure 6: Integrated model: variation of the estimated thresholds and regressors versus DA