3rd USA/Europe Air Traffic Management R&D Seminar Napoli, 13-16 June 2000

Human cognition modelling in ATM safety assessment

Henk A.P. Blom, Jasper Daams and Herman B. Nijhuis

National Aerospace Laboratory NLR, PO Box 90502, 1006 BM Amsterdam, The Netherlands
e-mail: blom@nlr.nl

ABSTRACT

This paper develops a mathematical model for cognitive performance of a tactical air traffic controller in an en-route ATC context. The aim of this model-based approach is to enable the evaluation of both accident risk and aspects like cognitive workload and effectiveness of ATC in managing air traffic situations safely. Use is made of human error modelling, Hollnagel's cognitive mode model and Wickens Multiple Resources model. The paper describes how these psychological sub-models are combined into a single model of ATCo cognitive performance, and how the interaction of these human sub-models with the technical sub-systems is brought into account.

The model is applied to an exemplary ATM scenario which consists of two parallel lanes of opposite direction traffic at the same flight level, and where the air traffic controller has no automation support tools like Short Term Conflict Alert (STCA) or Flight Path Monitoring. The obtained results for this conventional ATC situation are compared to those previously obtained under the hypothetical assumption that a tactical air traffic controller reacts to aircraft deviations in case of an STCA alert only.

1. INTRODUCTION

1.1 Safety based ATM design

Over decades, the aviation industry has been able to compensate the increase in traffic with a decrease in accident risk per flight hour. In view of the rapid growth of air traffic and the technological and organizational complexity of it, this has been a major accomplishment. Unfortunately, the point has been reached where it is unclear how to continue such compensation. The reason is that in the past the decrease of risk per flight hour has come in large part from technology driven improvements of safety. The effect of this technology-driven approach is shown through the accident statistics; they reveal that the relative share of human related causes is some eighty percent. This means that the historical air traffic safety compensation process can be continued if one learns to understand how the human and procedure related accidents could be reduced. This should be accomplished by learning the principles behind human related accident causes in aviation.

*) This research has been performed at NLR within a series of projects with support from the European Commission and the Netherlands CAA (RLD).

If we would try to understand these principles on the basis of an evaluation of incidents and accidents alone, then several difficulties arise. The number of incidents and accidents is limited, while the situations that caused them are quite complex and reports are not free from discussion. Due to the limited availability of data and the questionable validity of data, statistical analysis alone is not sufficient to model safety in complex situations with multiple human involvement. By now there is a broad consensus that appropriate safety models are needed to understand the mechanisms behind the remaining accident risk in relation to separation criteria and near-misses (Cohen & Hockaday, 1998). It is also recognized that such a safety modelling approach should be useful in optimizing advanced ATM operations (Haraldsdottir et al., 1997), (Odoni et al., 1997), (Wickens et al., 1998).

1.2 ATM safety modelling

Most existing studies on ATM safety either focus on hazard analysis techniques or on collision risk analysis: studies with thorough hazard analysis results generally use simplified collision risk models, advanced studies on collision risk between aircraft usually do not take into account hazards or non-nominal events (except in adapted tails of probability density functions). It appears that most established techniques fall short in integrating hazard analysis techniques with advanced collision risk analysis techniques. In a series of studies, at NLR this problem has been addressed with the development of a accident risk assessment methodology and supporting evaluation tool-set (both named TOPAZ) that takes an integral approach towards ATM safety assessment (Blom et al., 1998). Recently it has also been shown how this approach effectively supports safety management and the building of modern Safety Cases for advanced operations in ATM (Blom et al., 1999).

At the basis of the TOPAZ approach lies the use of a very general class of mathematical models for describing the ATM process. The models used are hybrid state Markov processes, which allow to describe stochastic evolution of both continuous and discrete variables over time. This means that e.g. both aircraft 3D position and velocity and operator states can be described as a function of time, including their interactions and the effects of probabilistic disturbances. To accomplish this, existing and newly developed models, such as Generalised Reich collision risk model (Bakker & Blom, 1993) and high level Petri Net model (Everdij et al., 1997a),

were combined in order to model and evaluate ATM operations on safety.

In parallel with its development the TOPAZ methodology, is applied to a variety of accident risk assessment studies, e.g. converging runways (Everdij et al., 1996), free flight equipped aircraft (Daams et al., 1998a, 1999a) and wake vortex induced accident risk (Kos et al., 2000). Another type of application considered is conventional en-route traffic in a scenario of two parallel opposite direction lanes. In (Everdij et al., 1997b), for this scenario, risk has been evaluated under two operational concepts. In the first concept, named 'No ATC' there is no ATC surveillance of the traffic at all, in the second concept, named 'STCA-only based ATC', the tactical air traffic controller sends deviating aircraft back to their lane if and only if there has been an STCA alert. Although the demonstrated possibility to obtain accident risk results for such complex operations as ATM is quite promising in itself, several operational experts pointed out that an STCA-only based ATC concept is overly conservative as a representation of conventional ATM concepts, since e.g. routine monitoring and anticipation is not incorporated. Therefore the follow-up was to develop an appropriate human performance model for risk assessments of such routine monitoring situations.

1.3 Human performance modelling

A crucial issue in ATM safety assessment is how the human factor is incorporated into the risk model. Hence there is a clear need for a modelling approach to assess and understand accident risk in relation to the performance of the human operators involved. This means that appropriate human performance models are required that describe human cognitive and responsibility principles up to the level of accident risk. This paper aims to present the developments of such a human cognition/performance model for a tactical controller within the context of conventional en-route ATC, and is based on a series of studies (Buck et al., 1996; Biemans and Daams, 1997; Daams et al., 1998b, 1999b). The resulting set of mathematical models is named HOMEROS (Human Operator Model to Evaluate Reliability Organisation and Safety).

At present, the view on human reliability has shifted from a context-free error centred approach, in which unreliability is modeled as failures of human information processing, towards a contextual perspective in which human actions are the product of human internal states, strategies and the environment, (Amalberti & Wioland, 1997) (Hollnagel, 1993), (Bainbridge, 1993). From this viewpoint, safety critical human actions should be modelled in their relation to the other activities of the operator and the environment. Thus for a proper description of human reliability it is necessary to include the cognitive processes that underlie the operator actions. As a result, one obtains a comprehensive model of the operator performing his job.

The main benefits expected from contextual models for safety assessment is that they provide better feedback to designers and that they remove the need to use overly conservative individual sub-models of relevant operator actions which may complicate understanding of how safety is achieved in aviation.

1.4 Organisation of this paper

The paper is organised as follows. Section 2 provides the background of three complementary psychological models on which human cognitive performance modelling in this paper is based. In Section 3 we explain how these three psychological models are jointly used in a mathematical human cognition/performance model for a tactical en-route controller. This is largely done on the basis of human factors ATC expertise. Next, in Section 4 this mathematical model is reduced to a simpler model on the basis of clearly defined model aggregation steps. In Section 5, the reduced human cognition/performance model is used to evaluate a conventional en-route ATC situation w.r.t. accident risk and air traffic controller actions. Finally, in Section 6 we discuss the results obtained.

2. HUMAN MODELLING APPROACHES

The mathematical human performance/reliability model development in this paper is based on the following three complementary psychological models:

- Multiple Resources Model
- Contextual Control Mode Model
- Human Error Modelling

In this section we outline these three psychological models. One should be aware that several other psychological human error type of models exist that have potential application in ATM, see e.g. (Isaac & Ruitenberg, 1999).

2.1 Multiple resources model

The main reference used here is (Wickens, 1992). The multiple resource model reflects the idea that humans have several different mental capacities with resource properties. In this view, task interference depends on the extent to which tasks use the same resources: two difficult tasks may be time-shared easily if they use different types of resources. The multiple resources approach has been well developed both for military applications (AGARD, 1998) and for ATM (e.g. Corker et al., 1997; Kilner et al., 1997). The principal idea behind the model is that human cognitive effort can be divided over several activities. This is called the *resources metaphor*, (Norman & Bobrow, 1975). Since human cognitive effort is limited, the resources metaphor may readily account for failures in time-sharing between competing activities. The underlying assumption of the resources metaphor is that the human is an information processing system with limited processing capacity. The model focusses on how this limited processing capacity can be used to time-share several processing tasks.

When two or more tasks are to be successfully time-shared, the first important aspects are the efficient scheduling and switching between activities. If sufficient time is available, the operator can occupy himself with one task at a time, although this does not necessarily mean that the tasks are performed sequentially. However, if the available time is not sufficient to apply this strategy, *concurrent* task performance becomes necessary. With respect to concurrent task performance, Wickens mentions three performance influencing task characteristics:

Confusion When an element of one task is similar to an element of a concurrently performed other task, the elements may become confused, leading to a decrease in performance.

Cooperation In some cases, the similarity between performance routines for elements of two tasks leads to cooperation between the routines. It is even possible that the two task elements can be merged into one new task.

Difficulty The task difficulty highly influences whether a second task can be performed concurrently.

The confusion and cooperation aspects are closely related: both emerge from the similarity between tasks. However, cooperation is associated with similar processing *routines* whereas confusion emerges from similar input *material*. By taking into account the confusion and cooperation aspects of concurrent task performance leads to multiple resources dimensions.

Multiple resources dimensions

On the basis of a large number of dual-task studies, Wickens proposes a three dimensional resource quantity, with dichotomous dimensions. The dimensions are:

Information processing stages Dimension with early and central processing on the one extreme (sensory processing, encoding and perception of stimuli) and late processing on the other (deciding on the best response and its execution). For example, the requirement for an air traffic controller to give a response to each change in aircraft state (late processing) is predicted not to disrupt the ability to maintain an accurate mental model of the radar display (early processing).

Modalities Input modalities differentiate between the encoding of auditory and visually presented stimuli. It is easier to divide attention between the eye and ear than between two auditory or two visual stimuli. Response modes refer to the choice between a vocal and a manual response. The reason that manual and vocal outputs can be efficiently time-shared is probably due to the separation of spatial and verbal information processing resources (manual responses are spatial in nature, while vocal ones are verbal).

Processing codes Human controllers can rely on two working memory codes, namely a spatial and a verbal one. Each is used to process or retain qualitatively different kinds of information (spatial and visual versus temporal, verbal and phonetic) and each can be disrupted by different concurrent activities. Resources underlying spatial processing and left-hand control reside predominantly in the right hemisphere of the brain. Resources underlying verbal processing, speech-responses and right-hand control reside more in the left hemisphere.

A note should be made about the modality dimension. In (Wickens, 1989) it is pointed out that the resources metaphor does not readily apply to input modality. Instead, pre-emption and attention-switching seem to dominate cross-modal time-sharing. However, these effects are relatively small in comparison to the effort required for the extra scanning activity that is generally involved with intra-modal time-sharing.

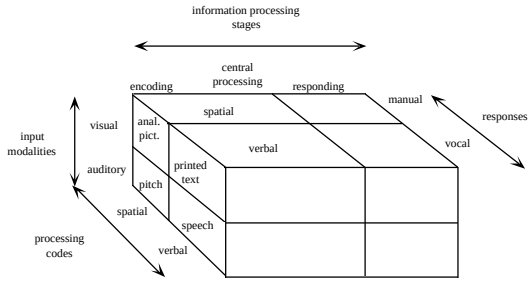

Figure 1: Proposed structure of resource dimensions (free after (Wickens, 1992)).

Figure 1 is a representation of the multiple resources theory. Although the theory does not pretend to account for all influences on multiple-task performance and time-sharing, research showed that the identified dimensions account for a reasonably large proportion of these influences and can be used in predicting task interference. Ideally, the loads on the dimensions must be established a priori. Input modalities are easy to define, as are vocal and manual output modes. Information processing stages and processing codes will cause more problems. Sometimes task analysis can reveal memory requirements.

We can now model the operators ability of time-sharing by evaluating which resources are used during the simultaneous execution of the tasks. Heavy concurrent demands upon the same resource then reduce time-sharing, whereas tasks using different resources can be time-shared easily. However, the Wickens model does not further describe this.

2.2 Contextual control mode model

A major trend in human performance modelling is the *Cognitive Viewpoint*, (Hollnagel, 1993). Within this approach, human behaviour is looked upon as a cyclic process, where human action is determined as much by the context as by inherent traits and mechanisms of human cognition. In this view humans do not passively react to events, they actively look for information and act based on intentions as well as external developments.

This approach in human performance modelling is in accordance with concepts from ecological psychology. Ecological psychology studies the information transaction between living systems and their environments, especially as they pertain to the perceived significance of environmental situations for the planning and execution of purposive behaviour (Gibson, 1986).

Based on the cognitive viewpoint, as stated above, Hollnagel (1993) describes a new approach that focuses on different control modes of the human operator's cognition, which reflect different control strategies in operator behaviour.

Control modes

The specific four control modes that are described by Hollnagel (1993) characterise in more detail regions of the continuum of control and can be specified as follows:

Scrambled Scrambled control denotes the case where the choice of the next action is completely unpredictable or random. The scrambled control mode constitutes the extreme situation of zero control.

Opportunistic Opportunistic control corresponds to the case when the next action is chosen from the current context alone, and mainly based on salient features rather than on more durable intentions or goals. It is opportunistic in the sense that the operator takes a chance, not because he is deliberately exploring an alternative, but because there is no time or possibility to do anything better.

Tactical Tactical control is characteristic for situations where the operator's performance is based on some kind of planning. Hence, the operator more or less follows a known procedure or rule. The planning is limited of scope and/or limited of range, and the needs taken into account may sometimes be ad hoc.

Strategic Strategic control means that the operator is considering the global context, i.e. using a wider event-horizon and looking ahead at higher level goals: either those which have been suspended and have to be resumed or those which, according to experience and expectations, may appear in the near future. This mode should provide a more efficient and robust performance.

An obvious question that arises is what determines the degree of control an operator has of a situation in a particular control mode and how the control mode changes. These topics are discussed next.

Control mode characteristics

In the Hollnagel approach, the control mode model consists of a high-level description of human behaviour, rather than a description of human tasks like planning, monitoring and decision making. To stress the high level character of these activities, they will be called meta-activities. Following Hollnagel (1993) the following four meta-activities are briefly elaborated upon below:

- Number of simultaneous goals
- Availability of plans
- Event horizon
- Mode of execution

Number of simultaneous goals This variable describes whether the operator considers only a single goal at a time or whether possible actions from multiple goals are considered. It is not the same as considering multiple choices or actions that may lead to the same goal (i.e. evaluating the effects of several possible lines of action, such as different ways of avoiding a conflict).

Availability of plans This variable describes whether the operator can refer to predefined or pre-existing plans (action templates) as a basis for choosing the next action. A plan can either be made on the spot, have been learned by experience, or have been defined explicitly in advance, e.g. as a written procedure. In either case, the availability of plans requires that the situation is familiar. A plan can either be followed rigidly or serve as a guideline for actions to be taken.

Event horizon This variable refers to how much of the past and how much of the future are taken into consideration when a choice of action is being made. The event horizon is described in terms of the number of steps, moves or items that are considered, rather than in subjective or objective time. The extent of the past is referred to as the history size, while the extent of the future is referred to as the prediction length.

Mode of execution The mode of execution can vary from feedforward to feedback driven. In the feedforward driven execution the operator carries out the steps of a chosen plan until either a predefined checkpoint has been reached or until external conditions force an interrupt. The better a feedforward is, the longer the uninterrupted period may continue. In the feedback driven execution each step (or group of steps) is followed by the evaluation of the feedback before the plan is continued. Even with a feedback mode of execution there may be sets of actions that are carried out in a feedforward way. Thus feedforward and feedback modes of execution define two ends of a continuum of possibilities.

Control mode changes

It is reasonable to assume that several factors will determine how and when an operator's control mode changes from one to another. Two of the obvious candidates that are described by Hollnagel are the amount of subjectively available time and the outcome of the previous action in terms of success and failure. These two main control parameters are briefly elaborated upon:

Subjectively available time This involves a consideration of the number of activities that remains to be carried out e.g. suspended actions, the number of simultaneous goals, the predicted changes and developments in the process and the environment (hence 'objective' time), the level of arousal, the level of familiarity of the situation, etc.. The estimation ranges from being quite detailed and precise to guessing and gut feeling.

Determination of outcome (of previous actions) This is not just a matter of ascertaining whether the previous action succeeded or failed. On the contrary, the determination of the outcome is different for each control mode, and may vary between a rudimentary detection of noticeable changes in the scrambled mode to a detailed evaluation of the feedback in the strategic mode. Complicating factors are for example the possible delays in outcome (for systems with large time constants), ambiguous or incomplete state indications and equivocal rules for interpretation.

2.3 Human Error Modelling

The main reference used here is (Kirwan, 1994). For safety-critical operations like nuclear power plants, it has been tried to take into account the human factor in probabilistic risk assessment (PRA) approaches. In general a PRA starts with an identification of hazards that might compromise the plant's safety. Next the propagation of the possible consequences of the identified hazards through the plant to the level of accidents is described. Frequently, this is done by means of fault- and event trees. After quantification of the frequency of occurrence of the identified hazards, plant accident risk are evaluated using the fault and/or event trees. Human error modelling approaches consist of two main elements: human error identification (HEI) and human error probability (HEP)

assessment. The results of these then fit within the fault/event tree analysis framework.

Human error identification

At the basis of human reliability modelling lies the (structured) identification of human operator related hazards. Generally, this involves a task analysis and a human error analysis. During the task analysis, the required operator actions during the various (sub-)processes of the plant are identified. In this stage, the equipment, interfaces, procedures and (trained) skills that are related to these actions are also identified.

After the task analysis, the HEI considers systematically what can go wrong. Commonly, the following types of errors are considered:

Error of omission Failing to carry out a required action.

Error of commission Failing to carry out a required act adequately: insufficient accuracy, wrong timing, actions performed in a wrong sequence.

Extraneous action Unrequired act performed instead of, or in addition to, required act.

Error-recovery opportunities Actions which can recover previous errors.

Underlying the HEI is a taxonomy of human error. As an example of a taxonomy of human error we describe the framework of skill-, rule- and knowledge-based behaviour, (Rasmussen, 1983), in the next subsection.

Skill, rule, knowledge based taxonomy

In (Reason, 1990), human error is related to cognitive processes that underlie human performance. Here, the human operator is looked upon as an information processor. The human information processor receives stimuli from the outside world, then processes these stimuli and finally responds. In this framework, human error emerges when during this scheme a deviation from normal processing routine occurs. One of the more influential models of human (erroneous) performance is the *Step Ladder* model of Rasmussen (1983). This model distinguishes three levels of human information processing: Skill-based level, Rule-based level and Knowledge-based level. These levels induce the following taxonomy of human errors:

Slips and Lapses Slips and lapses are unintended deviations from planned actions due to execution or memory failures.

Rule-based errors These are errors resulting from erroneous intentions, due to the application of bad rules or due to the misapplication of good rules.

Knowledge-based errors These are errors due to wrong reasoning about the to-be-controlled process. These mistakes may emerge from wrong or incomplete knowledge of the process or the bounded rationality of the operator.

Human error probability assessment

The second step in human error modelling is to quantify the probability of occurrence of the identified errors, for which many methods exist. As an example, we briefly describe the quantification part of THERP (Technique for Human Error Rate Prediction, see (Swain & Guttman, 1983)). Other examples are SLIM (Success Likelihood Index Methodology, (Embrey et al., 1984)) which relies mainly on expert judgement or HEART (Human Error Assessment and Reduction Technique, (Williams, 1988)) which focusses on the effects of identified Error-Producing Conditions.

3. MODELLING FOR EN-ROUTE ATC

The three psychological models of Section 2 are now used to develop a single mathematical model of a tactical en-route ATCo performing his job at a high (cognitive) level. Detail is given only when necessary. The model focuses on the following aspects of the interaction between the controller and the ATM process:

- Maintaining situational awareness
- Timely taking of safety critical actions
- Effectiveness of safety critical actions
- Occurrence of hazardous situations that involve the controller

This section is organised as follows. First the tactical controller task is described in terms of a suitable set of subtasks. Subsequently, the performance of the identified subtasks is related to the context in which the tasks are performed. Next, the scheduling of subtask performance is discussed, and it is explained how clearance errors that are initiated by the controller are incorporated. Finally the resulting mathematical model is described.

3.1 Description of controller task

The idea is to decompose the controller's task into several subtasks. This decomposition has been carried out along two dimensions: first a *generic dimension*, where the task is decomposed into cognitive activities at a general level which is independent from the scenario and operational concept. Secondly, the task is decomposed according to a *scenario/concept specific dimension*, where the controller task is described at the level of operational functions in the scenario. This twofold decomposition of the controller task allows flexibility in incorporating detail into the model: in this set-up we can restrict detail in the task description along the scenario/concept specific dimension to subtasks relevant for the problem under consideration, while the over-all interaction between controller and ATM process may still be properly modelled using the task description at the generic dimension.

First, a task decomposition along the generic dimension has been identified from (Buck et al., 1996). The resulting subtasks originate in (Jackson, 1989), however in (Buck et al., 1996) it was merged with several existing task-analyses (Ammerman et al., 1987), (Cox, 1994), (EATCHIP, 1996), (Endsley & Rodgers, 1994). The following subtasks resulted:

1. Sensing (to gather all information which is needed to get an overview over the air traffic situation).
2. Integration (to connect the gathered information thus forming a more global air traffic picture).
3. Prediction (to use the more global picture to anticipate future situations and events).
4. Complementary communication (pass the information to aircraft in order to improve the pilots understanding of the situation).

5. ATC problem solving planning (to use the understanding gained from the more global perspective to plan and prioritise aircraft actions).
6. Executive action (to communicate information and priorities as instructions to the aircraft in the system).
7. Rule monitoring (to ensure that the active components of the system behave in accordance with the 'rules'; monitoring and taking corrective actions for exceptions).
8. Co-ordination (to coordinate laterally with other parts of the ATC organisation).
9. Over-all performance (to ensure that the objectives of the operation are achieved, and that the infrastructure functions correctly).
10. Maintenance and monitoring of non-human part (to ensure that all systems supporting the controller work correctly).

Secondly, subtasks are also defined along the en-route ATC specific dimensions, where attention is focused on safety critical actions in the definition of the subtasks. This leads to the identification of three en-route context specific tasks:
A. Anticipate for aircraft deviating from intentions.
B. React to Automation alerts.
C. Perform other control activities.

We are now in the following position: the ATCo's task has been decomposed into subtasks along two dimensions: one relating the task to generic cognitive activities and the other dimension relating the task to specific situations in the scenario and operational concept considered. We next identified the task overlap *across* the dimensions in Table 1. This leads to 19 combinations across the dimensions, and thus a decomposition into 19 combined ATCo subtasks.

	A. Anticipate	B. Alerts	C. Others
1. Sensing	X		X
2. Integration	X		X
3. Prediction	X		X
4. Complementary communication			X
5. Problem solving /planning	X	X	X
6. Executive action	X	X	X
7. Rule monitoring	X	X	X
8. Coordination			X
9. Overall performance			X
10. Maintenance			X

Table 1: Task overlap across the generic cognitive activities and the en-route ATC specific tasks.

3.2 Task performance and control modes

In modelling the influence of the context on performance we adopt a mathematical model that incorporates two control modes: tactical control and opportunistic control. In Table 2 we identify the characteristic influence of these control modes on the performance of the A and B subtasks.

Since we may look upon subtasks C as representing a range of subtasks other than A and B along the en-route ATC specific dimension, it suffices to describe differences in tactical and opportunistic control mode at a general level only (see Daams et al., 1998b).

A1	Sensing:
	Tactical: Whenever possible the controller scans his display to detect possible deviations from ATC intentions. The controller divides the display into regions of interest and assesses these regions in a particular order. If scanning is interrupted at some time instant, the controller will resume scanning starting at the region that he was scanning when the interruption took place. Further information may also be obtained through R/T communication.
	Opportunistic: Whenever possible the controller scans his display to detect possible deviations. The controller scans in a random fashion.
A2	Integration:
	Tactical: The ATCo systematically integrates the information derived from scanning to improve his mental picture of the traffic situation. When some relevant information is not available, the ATCo may return to sensing to actively seek information to improve his assessment of the situation.
	Opportunistic: The ATCo integrates the randomly obtained information. An incomplete or even distorted mental picture may develop.
A3	Prediction:
	Tactical: The ATCo extrapolates his mental picture to the future traffic situation. On the basis of the assessment of the situation, the ATCo decides whether a problem may occur in the mid-term future.
	Opportunistic: The assessment of the future situation is restricted to a short time horizon and is based on incomplete information. It is assessed whether a problem may be expected in the short-term future.
A5	Problem solving/planning:
	Tactical: On the basis of the assessment of the (future) situation, the ATCo decides a resolution to the expected problem. In principle, the resolution involves replanning the aircraft trajectories in an optimal fashion with respect to safety, efficiency.
	Opportunistic: The resolution is aimed at solving the imminent problem only.
A6	Executive action:
	Tactical: The controller gives a series of R/T instructions to the aircraft involved. He verifies whether the pilot(s) readback these instructions correctly.
	Opportunistic: The verification of correct readback may be omitted.
A7	Rule monitoring:
	Tactical: After the R/T communication the controller verifies whether the aircraft comply to his clearances.
	Opportunistic: This may be omitted or be performed less thoroughly.
B5	Problem solving/planning:
	Tactical: On the basis of the assessment of the situation, the ATCo decides a resolution for the conflict. The resolution may range from vectoring both aircraft to doing nothing.
	Opportunistic: Same as in tactical control mode
B6	Executive action:
	Tactical: The controller gives the necessary R/T instructions to the aircraft involved. He verifies whether the pilots readback these instructions correctly.
	Opportunistic: The verification of correct readback may be omitted.
B7	Rule monitoring:
	Tactical: After the R/T communication the controller verifies whether the aircraft comply to his clearance.
	Opportunistic: Monitoring may be done less thoroughly or even be omitted.

Table 2: Subtasks related to Anticipation and Alerts.

3.3 Scheduling of subtasks

In this subsection the scheduling strategy applied will be defined for the subtasks. The scheduling strategy is expressed in the following (input) task parameters:

Pre-emption For each subtask an assumption is made whether it may pre-empt another subtask.

Concurrency For each subtask it is known whether it may be performed concurrently with another subtask.

Initiation For each subtask the circumstances under which the subtask should be performed are known.

The assumptions concerning Pre-emption and Concurrency are implemented according to Tables 3 and 4. These tables have been identified on the basis of ATC human factors expert knowledge.

	A1	A2	A3	A5	A6	A7	B5	B6	B7
A1	-	y	y	y	n	y	n	n	y
A2	y	-	y	y	n	y	n	n	y
A3	y	y	-	y	n	y	n	n	y
A5	y	y	y	-	n	y	y	n	y
A6	n	n	n	n	-	n	n	n	n
A7	y	y	y	y	n	-	n	n	y
B5	n	n	n	y	n	n	-	n	n
B6	n	n	n	n	n	n	n	-	n
B7	y	y	y	y	n	y	n	n	-
C1	y	y	y	y	n	y	n	n	y
C2	y	y	y	y	n	y	n	n	y
C3	y	y	y	y	n	y	n	n	y
C4	y	y	n	n	n	y	n	n	y
C5	y	y	y	y	n	y	y	n	y
C6	y	y	n	n	n	y	n	n	n
C7	y	y	y	y	n	y	n	n	y
C8	y	y	n	n	n	y	n	n	y
C9	y	y	y	y	n	y	n	n	y
C10	y	y	y	y	n	y	n	n	y

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
A1	y	y	y	y	y	y	y	y	y	y
A2	y	y	y	y	y	y	y	y	y	y
A3	y	y	y	n	y	n	y	n	y	y
A5	y	y	y	n	y	n	y	n	y	y
A6	n	n	n	n	n	n	n	n	n	n
A7	y	y	y	y	y	y	y	y	y	y
B5	n	n	n	n	y	n	n	n	n	n
B6	n	n	n	n	n	n	n	n	n	n
B7	y	y	y	y	y	n	y	y	y	y
C1	-	y	y	y	y	y	y	y	y	y
C2	y	-	y	y	y	y	y	y	y	y
C3	y	y	-	n	y	n	y	n	y	y
C4	y	y	n	-	n	n	y	n	y	y
C5	y	y	y	n	-	n	y	n	y	y
C6	y	y	n	n	n	-	y	n	y	y
C7	y	y	y	y	y	y	-	y	y	y
C8	y	y	n	n	n	n	y	-	y	y
C9	y	y	y	y	y	y	y	y	-	y
C10	y	y	y	y	y	y	y	y	y	-

Table 3: Concurrent performance of subtasks.

Tables 3 and 4 should be read as follows. Consider subtasks C4 (general communication) and A3 (prediction with respect to deviations). It follows from Table 3 that these two subtasks can not be performed concurrently. Next inspect the row C4 in Table 4 at the collum corresponding to A3, we see that C4 pre-empts A3. Thus if A3 is carried out and C4 is initiated, execution of A3 will stop and C4 will be performed first. If concurrent performance were possible (i.e. there would be a 'y' in Table 3), then pre-emption would mean that C4 and A3 are performed concurrently, with C4 as the primary and A3 as the secondary task. In terms of a stack of to-be-performed subtasks this scheduling principle can be formulated generically as the following two rules:

Rule 1: An initiated subtask will be placed in the stack before the subtasks that it may pre-empt.

Rule 2: If the first two subtasks of the stack can be processed concurrently, this will be done (subtask duration will be slightly longer, however).

	A1	A2	A3	A5	A6	A7	B5	B6	B7
A1	-	n	n	n	n	n	n	n	n
A2	n	-	n	n	n	n	n	n	n
A3	n	n	-	n	n	n	n	n	n
A5	A5	A5	A5	-	n	A5	n	n	A5
A6	A6	A6	A6	A6	-	A6	n	n	A6
A7	n	n	n	n	n	-	n	n	n
B5	B5	B5	B5	B5	n	B5	-	n	B5
B6	B6	B6	B6	B6	B6	B6	B6	-	B6
B7	n	n	n	n	n	n	n	n	-
C1	n	n	n	n	n	n	n	n	n
C2	n	n	n	n	n	n	n	n	n
C3	n	n	n	n	n	n	n	n	n
C4	C4	C4	C4	n	n	n	n	n	n
C5	n	n	n	n	n	n	n	n	n
C6	C6	C6	C6	n	n	C6	n	n	C6
C7	n	n	n	n	n	n	n	n	n
C8	C8	C8	C8	n	n	n	n	n	n
C9	n	n	n	n	n	n	n	n	n
C10	n	n	n	n	n	n	n	n	n

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
A1	n	n	n	n	n	n	n	n	n	n
A2	n	n	n	n	n	n	n	n	n	n
A3	n	n	n	n	n	n	n	n	n	n
A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5
A6	A6	A6	A6	A6	A6	A6	A6	A6	A6	A6
A7	n	n	n	n	n	n	n	n	n	n
B5	B5	B5	B5	B5	B5	B5	B5	B5	B5	B5
B6	B6	B6	B6	B6	B6	B6	B6	B6	B6	B6
B7	n	n	n	n	n	n	n	n	n	n
C1	-	n	n	n	n	n	n	n	n	n
C2	n	-	n	n	n	n	n	n	n	n
C3	n	n	-	n	n	n	n	n	n	n
C4	C4	C4	C4	-	C4	n	C4	n	C4	C4
C5	C5	C5	C5	n	-	n	C5	n	C5	C5
C6	C6	C6	C6	C6	C6	-	C6	C6	C6	C6
C7	n	n	n	n	n	n	-	n	n	n
C8	C8	C8	C8	C8	C8	n	C8	-	C8	C8
C9	n	n	n	n	n	n	n	n	-	n
C10	n	n	n	n	n	n	n	n	n	-

Table 4: Pre-emption between subtasks.

3.4 Errors in flightplans and intents

An important safety issue is that for one single aircraft there may be all kind of differences between the flight intents on the ground and in the air, and the ATCo and pilot awareness of those intents, i.e.:

- Tactical ATCo's awareness of the flight intent
- Flightplan in the ATC system
- Pilot's awareness of the flight intent
- Flightplan used by the FMS

To allow for these differences the following mathematical modelling approach is adopted:

ATCo The tactical ATCo's awareness of the flight intent is assumed to be ATC's true reference. The quality of ATC's true reference is in one of the following two discrete modes: i) the true reference provides separation, ii) the true reference does not provide separation. In general the latter mode value may be reached if an ATCo has made a knowledge-based error.

ATC The quality of the flightplan in the ATC system may be in one of the following two discrete modes: i) agrees with ATC's true reference, ii) differs from ATC's true reference. The latter is due to an ATCo input error, or an ATC database error.

Pilot The quality of the pilot's awareness of ATC's true reference is in one of the following two discrete modes: i) agrees with ATC's true reference, ii) differs from ATC's true reference. The latter may happen due to a clearance error. There are two types of clearance errors: 1) intended clearance given to wrong aircraft or 2) wrong clearance given to intended aircraft. The causing factor may be with the ATCo, or the pilot or both, and may be knowledge-based, rule-based or skill-based.

FMS The quality of the flightplan used in the FMS is in one of the following two discrete modes: i) agrees with ATC's true reference, ii) differs from ATC's true reference. The latter happens if pilot awareness differs from ATC's true reference or is due to a pilot input error or an FMS data base error.

In elaborating the above it is assumed that all the ATCo related errors may occur at random during performance of subtasks A6, B6 or C6, (executive action) where the frequency of occurrence depends on the control mode the controller is in. Furthermore, such errors may be detected and corrected during rule monitoring subtasks A7, B7 or C7, also depending on the control mode (e.g. Amalberti and Wioland, 1997).

3.5 Mathematical model of tactical ATCo

In order to establish the connections with the other ATM processes, in this subsection we describe the mathematical model of the ATCo from an input-output point of view. First we describe how initiation of cognitive activity is modelled, then the implementation of the task description and controller performance is described. The Petri Net of the tactical ATCo model is shown in Figure 2.

Initiation Three stimuli for ATCo cognitive activity are identified: ATCo's Anticipation, Automation alerts and other actions. Activity triggering situations that first have to be detected by the operator (like an aircraft severely deviating from its route) are not considered as an initiation stimulus, since general sensing is modelled as a part of the operators task and therefore the sensing activity has to be initiated first. For the occurrence of certain stimuli various other ATM modules may need to function properly, such as e.g. the ATCo HMI and surveillance for an Automation alert.

Within the Petri Net each stimulus is modelled as a place, connected with one transition that fires if initiation of the corresponding cognitive activity takes place. These transitions produce two tokens: one token returning to the stimulus place for future generation of cognitive activity and one token in a *stack* place. The *stack* places represent the situation that the respective initiated cognitive activity has to wait until the operator has completed other (more important) tasks. The places *Anticipation*, *Alert* and *Other action* represent initiation of cognitive activity by own initiative, Automation alerts and other action (e.g. a pilot request) respectively. Preconditions on occurrence of these stimuli are modelled within the respective transitions: if the preconditions are not met the transition does not fire. For example: the proper functioning of the ATCo HMI as a precondition for the occurrence of an Automation alert triggering ATCo cognitive activity is modelled as a precondition for firing of the transition connected to the *Alert* place.

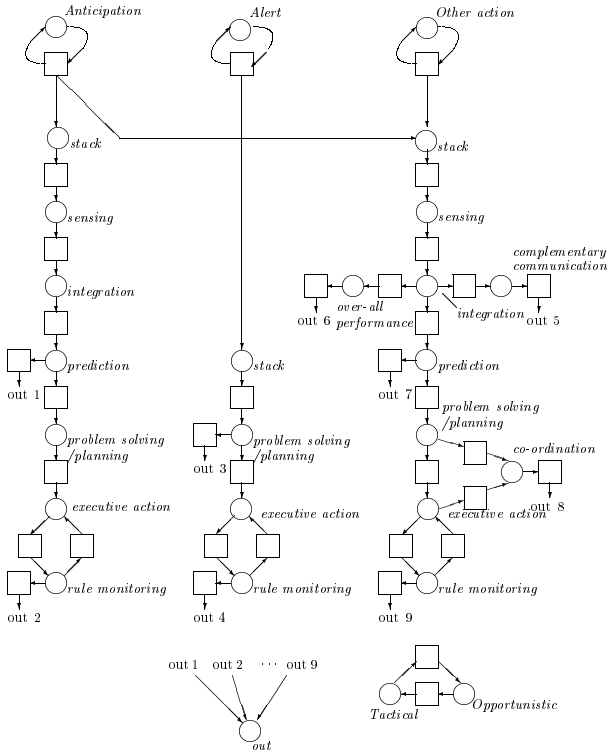

Figure 2: Petri Net of tactical ATCo model in HOMEROS.

ATCo subtasks The ATCo task has been divided into several sub-tasks which are each defined as the combination of a scenario specific purpose and a generically described cognitive activity. Three context specific purposes are modelled: ATCo to detect and correct deviations of aircraft from ATCo intentions, ATCo to react to Automation alerts (initiated by Automation tools) and ATCo to perform other control activities (initiated by own initiative or through other actions). Each subtask is represented by a place in the Petri Net, which is named after the cognitive activity it represents. The tokens then model cognitive activity on the subtask that corresponds to the place that they reside in. Some cognitive activities may be performed for several purposes, leading to several places with the same name. Below we describe the places with respect to the cognitive activities that they represent.

The places named *sensing* represent the situation that the ATCo is gathering information to improve his picture of the traffic situation. The places named *integration* represent the

situation that the ATCo incorporates the newly obtained information into this mental picture. The place named *communication* represents the situation that the ATCo makes his knowledge of the situation available to the pilots. The place *named over-all performance* describes the evaluation of sector performance as a whole. In the *prediction* place, the ATCo extrapolates his picture of the traffic to the future, while in the *problem solving/planning* place he synthesizes solutions to possible (future) problems. In the *executive action* place the operator gives clearances to aircraft, followed by the *monitoring* places where it is verified whether the aircraft adhere to these clearances. In the *out* place the tokens are collected after performance.

Whenever one subtask is logically performed after performing another (e.g. *prediction* is performed after *integration*) and they have the same scenario specific purpose a transition is drawn between those two subtasks.

Subtask scheduling We next incorporate the scheduling rules. Scheduling depends on the relative priority of a subtask and the possible concurrent performance of two subtasks. The relative priority is modelled as a colour type that is associated with the tokens that represent cognitive activity on subtasks. This colour type is a number 1,2,…. where low numbers correspond to high priority. The priority colours are up-dated whenever a new token is initiated and when a token is collected in the *out* place, according to a suitable set of assumptions. According to the scheduling rules either the token that has priority 1 is performed exclusively or the tokens with priority 1 and 2 are performed concurrently, with the token with priority 2 being the secondary task.

We assume that for each subtask the time needed to complete it has a certain probability density, given the current control mode of the ATCo and possible concurrent performance of another subtask. In the Petri Net, the duration of performing a subtask is modelled as a delay in the firing of the transition that has the subtask as input place. Transitions with a token in the input place that does not have priority 1 or 2 have 'infinite' delays. Transitions with a token in the input place that has priority 1 has a delay corresponding to the normal duration of the subtask, given the control mode. Delays of transitions with a token in the input place that has priority 2 either have an 'infinite' delay or a delay that may be longer than when the corresponding subtask is performed exclusively. This depends on the extent to which the subtasks with priority 1 and 2 may be performed concurrently. Hence in the Petri Net, each transition has a delay that is a function of the priority of the token in the input place, the current control mode and the place that the token with priority 1 resides in.

The ATCo's executive actions (i.e. the clearances given) are also modelled as a colourtype associated with the tokens in the subtasks. This colourtype is a set of paired numbers describing the type of clearance given and the aircraft that the clearance is given to. The decision to give no clearance at all is also modelled as an executive action and has a separate colour value. In the present model it is assumed that the type of clearance given is determined during the *executive action* subtask only and that it depends on the control mode only. So the firing of the transitions after the *executive action* places also affects the Petri nets of other ATM modules: completion of executive action means that a decision to give a clearance to an aircraft has been carried out and therefore the firing of these transitions describe the ATCo control actions.

ATCo control modes In the model, ATCo performance depends on the control mode, scheduling rules and results in a clearance. In the DCPN model of the ATCo two control modes are identified, which are each represented by a place in the Petri Net: the place named *Tactical* models the situation that the controller has a relatively high degree of control and the place named *Opportunistic* models a relatively low degree of control. The control mode may influence ATCo performance in all aspects. The switching between control modes is modelled by transitions between the *Tactical* and *Opportunistic* places. The resulting subnet contains one token, the place of which defines the current degree of control. The firing of the transitions between the control modes depends on the number of tokens in the *stack* places (indicating the subjectively available time) and the number of times that *monitoring* was followed by another *executive action* during the last few minutes (indicating the outcome of previous actions measured as the number of clearances that the controller considers to be insufficiently effective). Details for this type of modelling appeared to be available through human factors ATC expert knowledge.

4. REDUCTION OF THE ATCO MODEL

In this section we explain how the ATCo model that was developed in Section 3 is reduced by applying appropriate model aggregations. The motivation for this reduction is that the complexity of the original model results from a detailed modelling which is judged to be unnecessary for the application at hand. This makes the resulting reduced model interesting in its own right.

First we explain how the subtasks are clustered into a new set of subtasks, and how scheduling simplifies accordingly. Second, the Petri Net for the ATCo reduced model is given. Third, within an en-route context we compare the relevant model characteristics to verify that the model based on the reduced task description is indeed an appropriate approximation.

4.1 Aggregation of subtasks

In the previous section, Table 3 and Table 4 show that due to the possibility of concurrent performance of subtasks the number of required assumptions concerning concurrent subtask performance equals ½n(n-1) and the number of required assumptions concerning pre-emption equals n(n-1), with n the number of identified subtasks. For the present 19 subtasks, this means a total of 342 rules concerning task scheduling in the model. This large number of rules may severely complicate the stochastic analysis which is required for risk evaluation. Therefore, it is desirable to reduce the complexity of the model without compromising conservativeness or psychological validity.

This reduction of the full model is achieved by decreasing the level of detail at which the air traffic control task is described

and the way performance of these tasks is scheduled according to single-task performance.

The approach taken is to group the 19 subtasks into a smaller number of clusters of subtasks. The following clusters are identified:

Cluster	Initial subtasks
$Monitoring_A$	A1-A3
$Communication_A$	A5-A7
$Communication_B$	B5-B7
Complementary $Communication_C$	C4
$Communication_C$	C6
$Co\text{-}ordination_C$	C8
$Miscellaneous_C$	C1-C3, C5, C7, C9, C10

Table 5: Clustering of the subtasks.

Next, we need to identify how task scheduling at the level of clusters of subtasks takes place. First, concurrent performance of clusters of subtasks is investigated using Table 3. This is done conservatively using the principle that if one combination of the clustered subtasks cannot be performed concurrently, then the whole clusters of subtasks cannot be performed concurrently. Application of this principle yields the following table:

	Mon_A	Com_A	Com_B	CpC_C	Com_C	$Coor_C$	$Misc_C$
Mon_A	-	n	n	n	n	n	y
Com_A	n	-	n	n	n	n	n
Com_B	n	n	-	n	n	n	n
CpC_C	n	n	n	-	n	n	n
Com_C	n	n	n	n	-	n	n
$Coor_C$	n	n	n	n	n	-	n
$Misc_C$	y	n	n	n	n	n	-

Table 6: Concurrent performance of clusters of subtasks, this table is derived from Table 3 and Table 5.

In a similar fashion, we identify a new table (Table 7) for the pre-emption between clusters of subtasks. The following rule is applied: if any subtask in some cluster A pre-empts all subtasks in some other cluster B, then cluster A pre-empts cluster B. Otherwise, cluster A does not pre-empt cluster B.

	Mon_A	Com_A	Com_B	CpC_C	Com_C	$Coor_C$	$Misc_C$
Mon_A	-	n	n	n	n	n	n
Com_A	Com_A	-	n	Com_A	Com_A	Com_A	Com_A
Com_B	Com_B	Com_B	-	Com_B	Com_B	Com_B	Com_B
CpC_C	CpC_C	n	n	-	n	n	CpC_C
Com_C	Com_C	n	n	Com_C	-	Com_C	Com_C
$Coor_C$	$Coor_C$	n	n	$Coor_C$	n	-	$Coor_C$
$Misc_C$	n	n	n	n	n	n	-

Table 7: Pre-emption between clusters of subtasks, this table is derived from Table 4 and Table 5.

Table 7 implies that the cluster $Miscell_C$ does not pre-empt any other cluster. Moreover, $Miscell_C$ is pre-empted by all other clusters, except $Monitoring_A$. Furthermore, it follows from Table 3 that $Monitoring_A$ and $Miscell_C$ can be performed concurrently. From this we conclude that performance of the subtasks in the cluster $Miscell_C$ does not conflict with other subtasks at cluster level. Since the cluster $Miscell_C$ itself does not contain subtasks which are directly relevant for safe separation, we can therefore discard this cluster in the model without compromising conservativeness. Therefore, we do not take into account this cluster in the sequel.

Now inspect Table 6 and Table 7 again. Perhaps surprisingly, we see that concurrent performance of the remaining clusters of subtasks is not possible. Moreover, the remaining pre-emption rules boil down to a fixed priority list where $Monitoring_A$ has lowest and $Communication_B$ has highest priority. Apparently, similar principles underly Table 3 and Table 4, although the construction of these tables was done before and independently from the present analysis.

We conclude that at the level of clustered tasks, the complexity of the scheduling principle is reduced significantly, without compromising conservativeness. In summary, the main model simplifications are

- Reduction from 19 subtasks to 6 clusters of subtasks.
- Concurrent task performance is simplified into single task performance.
- Pre-emption rules for each combination of subtasks are simplified into a fixed priority list.

4.2 Reduced ATCo model

On the basis of the aggregation a reduced model of the ATCo can now be developed. Six main ATCo cognitive tasks are identified, which describe the operator performance at a cognitive level. For each task, we assumed a relative priority ranking, an average duration and the percentage of his time that the operator would spend on the task if uninterrupted.

Task	Prio	Description
$Monitoring_A$	6	Visual anticipation and detection of deviations from the ATCo intention
$Communication_A$	2	Communicate clearance with an aircraft that deviated severely visually from ATCo intention
$Communication_B$	1	Communicate clearance with aircraft for which an Automation alert was issued
$Complementary\ communication_C$	5	General complementary communication with pilots
$Communication_C$	3	General communication of executive action (i.e. clearances)
$Co\text{-}oordination_C$	4	General coordination with planner controller, controllers of other sectors.

Table 8: Six main cognitive tasks.

The ATCo performs these tasks one at a time, according to the given priorities. Task scheduling is kept straightforward: high priority tasks are performed first, possibly pre-empting a low priority task. Two important aspects of performance are incorporated as well: the influence of the control mode and the possibility of erroneous clearances. The Petri Net describing the discrete modes for the ATCo model is given in Figure 3

Two control modes are considered: *Tactical* and *Opportunistic*, which reflect the degree of control. In the *Tactical* mode, the ATCo takes his time and makes little errors. In the *Opportunistic* mode, the general tasks (marked subscript C) are performed faster, but the chances on errors are also larger. The switching between the control modes depends on the subjectively available time (measured as the

number of tasks waiting to be performed) and the outcome of previous actions (measured as the number of corrective actions, i.e. $Communication_A$ and $Communication_B$, taken by the ATCo during the last two minutes). If the subjectively available time is short or if the outcome of previous actions is bad then the ATCo switches to *Opportunistic* control mode.

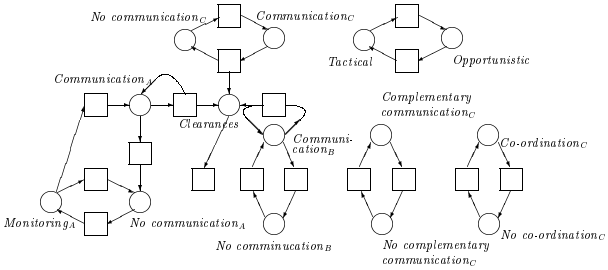

Figure 3: Petri Net of reduced ATCo model in HOMEROS.

ATCo erroneous clearances are taken into account as follows: the ATCo may give a different clearance than he intended to (e.g. switching heading and speed), or he may give the clearance to a different aircraft than he intended to (call-signs mixed up). These errors are incorporated as random variations in the ATCo actions. The error types are represented as a colour value of the tokens in the place *Clearances*.

The switching between modes is affected by several other modules, such as Aircraft evolution, Surveillance, ATC system, R/T local, R/T global, Performance of pilot. Surveillance output (i.e. the estimated aircraft state is input for the visual detection of severe deviations by the ATCo. The ATC system must be *Working* for the ATCo to be able to do his job. The R/T modules and Pilot module together form the Decision Making loop or DM-loop. If all modules in the DM-loop are Working, *Relaxed*, *Delaying* or *Busy* for a given aircraft, then the ATCo is able to give a clearance to that aircraft.

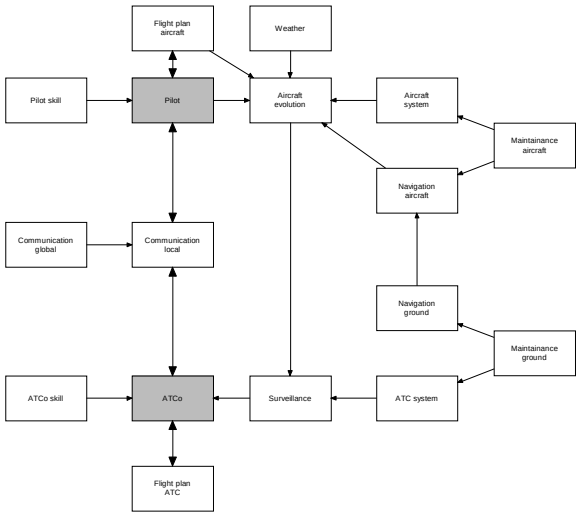

Figure 4: Functional representation of conventional ATC.

4.3 Comparison against statistical data

Next we evaluated for the ATCo routine monitoring concept the period to detect severe deviations such that a comparison with available statistical data is possible (George et al., 1973).

A full and reduced ATCo performance model was developed on the basis of the cognitive principles identified in Section 2 and integrated with appropriate Petri Net models for the other relevant components in conventional ATC (see Figure 4).

Comparison, in Figure 5, with the model based results shows that the detection time results of both the original and the reduced ATCo model agree quite well with the measured data. It should be noticed that in (George et al., 1973) only very few detection times beyond 150 s were measured. This is most probably due to the limited number of measurements made in combination with the low probability of such long detection times. Although they have low probability, the longer detection times add significantly to the risk, and Figure 5 shows that model based results do extend to these low probability values. We may conclude that both the full and the reduced model curves agree quite well with the statistical data. This clearly contributes to gaining confidence in the model-based approach taken.

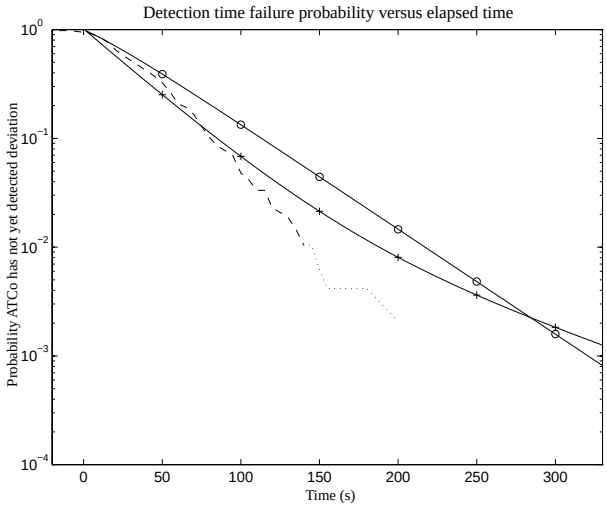

Figure 5: ATCo detection time of severe deviations of the full model (line marked '+'), of the reduced model (line marked 'o') and of statistical data, (George et al., 1973), (dashed/dotted line, the dotted part representing data based on less than 5 measurements).

Discussion of model reduction results

In this section we have shown how to derive a reduced model of the ATCO performance from a more detailed ATCo model which was developed in Section 3. This reduction is based on using a less detailed decomposition of the air traffic control task and simplifying concurrent task performance into single task performance (i.e. one task at a time). From Figure 5 it appears that this reduced model yields slightly more conservative ATCo detection time results. Therefore we conclude that for the particular application considered here, incorporation of concurrent task processing into the ATCo performance model is not necessary for avoiding overly conservative risk estimates. Obviously, incorporation of concurrent processing into human performance models may be essential for other applications such as detailed workload assessment.

5. EXAMPLE APPLICATION

In this section we show TOPAZ based assessment results for accident risk and ATCo actions, for an hypothetical ATM scenario that consists of two en-route traffic streams of RNP1 equipped traffic, flying in opposite direction, all at one single flight level.

5.1 Hypothetical ATC example

The rather hypothetical example has been developed by Eurocontrol with the aim to learn understanding how ATC influences accident risk, and how far the nominal separation S between opposite RNP1 traffic streams can safely be reduced. The specific details of this scenario are:

- Straight route, with two traffic lanes
- ATCo expects all aircraft to stay on these lanes
- Parameter S denotes distance between the two lanes (see Figure 6)
- Opposite traffic flows along each lane
- Aircraft fly at one flight level only
- Traffic flow per lane is 3.6 aircraft/hour
- All aircraft nominally perform RNP1
- None of the aircraft are TCAS equipped
- Target level of safety is 5.10^{-9} accidents/flight hour.
- 15 aircraft per sector/ATCo
- There are no military aircraft

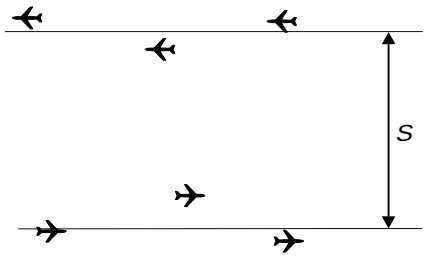

Figure 6: Opposite direction traffic in a dual lane structure.

This exemplar scenario is considered for the following three ATM concepts:

- Procedural separation only. In this case there is no ATC surveillance system. This is the type of situation encountered with traffic over the North Atlantic.
- STCA-only based ATC. In this case there is a radar based surveillance and R/T communication, but it is assumed that ATC is doing nothing with this information unless its STCA system issues an alert; thus assuming no monitoring by ATCo.
- Routine monitoring based ATC. The same as in B, but now without STCA system. Thus aircraft deviations are only identified through routine monitoring.

5.2 Accident risk

For each of the three ATM concepts the TOPAZ methodology and tool set have been used to assess accident risk for the above scenario, as a function of the spacing parameter S. The accident risk result is presented as the graph marked 'ATCo routine monitoring' in Figure 7. In Figure 7, there also is a horizontal line that represents the target level of safety (TLS) which has been specified at $5 \cdot 10^{-9}$ expected accidents per flight hour in (ICAO, 1998).

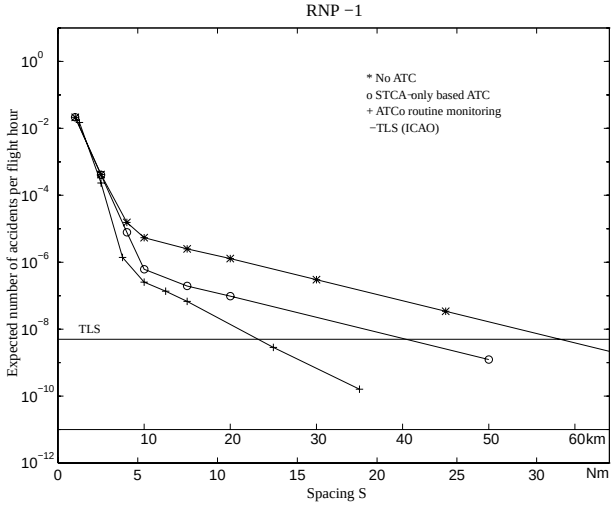

Figure 7. Accident risk versus route spacing, the graphs marked 'No ATC' and 'STCA-only based ATC' have been taken from (Everdij et al., 1997). The graph 'ATCo routine monitoring' is from (Daams et al., 1999b). The TLS value used is defined in (ICAO, 1998).

Qualitative uncertainty analysis

Absolute usage of the risk curves without taking into consideration existing bias and/or uncertainty can inspire undue conclusions. Due to a model based quantitative risk assessment approach, it is possible to bring the model assumptions made to the foreground and subsequently perform an uncertainty analysis of the model versus reality.

With the TOPAZ methodology, the starting point for such uncertainty analysis consists of the following:

- description of nominal operation and procedures,
- list of hazards identified for the operation considered,
- list of assumptions made when building the Petri net,
- Petri net specification (local and interactions),
- list of parameters and values used during the numerical evaluation, and their sources.

For the routine monitoring concept there are more than 200 hazards (about 50% is human related), about 25 model assumptions, and about 100 model parameters (about 20% for the reduced ATCo model).

The qualitative uncertainty analysis that can be performed works as follows. First, for each hazard it is specified how it is incorporated in the Petri Net or not (due to a model assumption listed). The result is that for each parameter and for each assumption the related hazards are identified. The subsequent steps are:

- Per assumption, perform a qualitative assessment of its uncertainty impact on the risk,
- Per parameter value, perform a qualitative assessment of the uncertainty in relation to the applicable hazards,
- Per parameter value assess the impact of this uncertainty impact on the risk.

At this moment, this qualitative uncertainty analysis has not yet been applied to the TOPAZ evaluated en-route examples. However, it has successfully been applied in a Wake vortex

risk assessment study (see Kos et al., 2000). On the basis of this experience we expect that the main contribution to uncertainty will come from unmodelled hazards (either due to model assumptions or due to missing hazards), rather then from parameter value uncertainty. For the curves in Figure 7 this means that for the time being they should be interpreted in a relative way only.

Analysis of risk curves

Inspection of Figure 7 yields that the TLS is reached for a route spacing of about 24 km (13 Nm), which is a significant improvement of the values of the No ATC curve (TLS reached at about 58 km (~32 Nm)) and the STCA-only based ATC curve (TLS reached at about 40 km (~22 Nm)). Obviously, for busy fixed route situations over the continent, procedureal separation is not very helpful. STCA-only based ATC neither helps a lot. The improvement provided by the routine monitoring shows that it is much more effective in safely managing deviations from centerline than reacting to STCA alerts only. Apparently, STCA really is a safety net only.

We also observe that the risk reduction provided by monitoring based ATC increases as route spacing increases. This is in contrast to the STCA-only based control strategy, where the ATCo prevents a fixed ratio of the deviating aircraft that reach the other route from collision. The reason for this increasing risk reduction is that the number of severe deviations that are detected before the aircraft reaches the other route increases faster with route spacing than the decrease in the number of deviating aircraft that reach the other route. Hence, the slope of the risk figure depends on the slope of the ATCo detection time instead of the slope of the non-nominal lateral deviation probability density function. Consequently, accident risk may be further reduced by changing the ATM design and in particular the role of the controller such that the ATCo detection time is improved.

Safety criticality analysis

Further evaluation showed that safety criticality lies with the *Sharp turn* type of deviations. This is caused by the fact that during the *Sharp turns* the aircraft deviates from the route much faster than in the case of a general *Non-nominal* deviation. We evaluated for S = 24 km (13 Nm) the risk involved with the *Sharp turns* to be a factor 15 higher than for the *Non-nominal* deviations.

In the present model, the *Sharp turns* are caused by erroneous ATCo clearances and aircraft flightplan errors, whereas the *Non-nominal* deviations are caused by degraded navigation systems, degraded aircraft systems etc. Hence from the safety criticality result we conclude that the most risky situations originate in the human factor rather than in degraded performance of technical systems.

5.3 ATCo effort and effect

The ATC effort is related to the number of ATC actions normalised by the theoretical minimum of ATC actions required for averting all accidents (i.e. one action per accident that would occur if there were no ATC). This is approximately equal to the number of ATC actions required to avert one accident (as almost all potential accidents should be averted).

$$\text{Effort: } \rho_a = \frac{\text{ATC actions}}{\text{Accidents, without ATC}} \left(\approx \frac{\text{ATC actions}}{\text{Averted accidents}}\right)$$

Next we express the ATC effect as the factor of accident risk reduction achieved by ATC:

$$\text{Effect: } \rho_b = \frac{\text{Accidents, without ATC}}{\text{Accidents, with ATC}}$$

Graphs for the metrics ρ_a and ρ_b for STCA-only based ATC and routine monitoring are given in Figure 8.

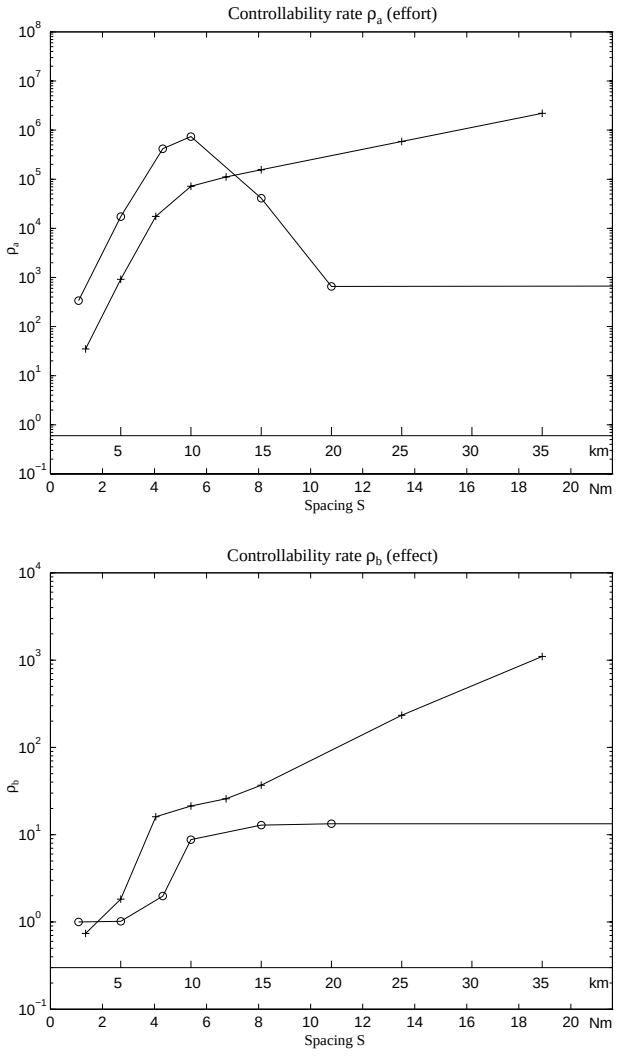

Figure 8. Effort ρ_a and effect ρ_b for routine monitoring (+ graph) and STCA-only based ATC (o graph).

From the ρ_b curves in Figure 8 we conclude that the monitoring strategy yields a better risk reduction for *all* spacings. Secondly, inspection of the ρ_a graphs yields that for small spacings monitoring requires even less effort than STCA-only. For larger spacings monitoring requires more effort than STCA-only. We conclude that for small spacings, monitoring is to be preferred (more effect, less effort), while for larger spacings the situation is less clear (more effect, but also more effort).

A remark should be made concerning spacings below 2 Nm (where the ρ_b has negative values for monitoring). Notice that

these very low spacings are not realistic for the monitoring concept, since for these spacings aircraft may collide while remaining within the safe boundary around the lanes (whence the ATCo does not take action to prevent these collisions). We therefore disregarded these very small spacings.

6. CONCLUDING REMARKS

This paper applied state of the art psychology in human cognition/performance modelling for application to accident risk modelling. This led to the development of mathematical human cognition/performance models for a tactical ATCo in a conventional en-route ATC situation. This model is shown to be of great use in the evaluation of accident risks for an ATM scenario with the tactical ATCo performing routine monitoring to detect and correct for severely deviating aircraft.

6.1 Validation issues

In this work we took a model-based approach towards the assessment of concepts such as accident risk and controllability in ATM situations. This makes the approach a formal one: for the model, accident risk and ATCo effort and effect indicators are unambiguously defined. If numerical evaluations of the model are carried out in a verifiably correct way, then the validity of the results depends on the verifiability of the model only.

The main problem thus is how to verify that the model 'matches' reality sufficiently well, with respect to the intended use of the model. It should be stressed that an absolute 'match' is not feasible, however this is also not neccesary. Instead a case that the model is sufficiently realistic for its purposes should be built, by testing both the assumptions made during model development and relevant characteristics of the eventual model. The confidence in the model should then be based on the quality of the arguments for its validity (i.e. the 'test results'). This model validation approach is currently under development. On the basis of the human cognition modelling and the controllability results in the present report, we recognise a contribution to this approach which consists of comparing relevant model characteristics with human-in-the-loop measurements in the case of human controllability evaluation. Such comparison should always be treated with care as the results may be sensitive to the context.

For the present model, three tests of its validity have been carried out: in (Daams et al., 1998b) the human performance modelling approaches that underly the ATCo model used have been shown to be sufficiently powerful to explain ATCo related hazards in en-route ATM. Secondly, in Figure 5 ATCo detection time which is a relevant model characteristic for accident risk was compared against controller-in-the-loop based data from the literature. Thirdly, during the whole development of the ATCo model a Human Factors specialist has been actively involved and the results have been reviewed by an operational expert. Obviously, further confidence building can and should be done, e.g. on the basis of detailed reviews with a number of experts and comparison of a range of model characteristics with additional empirical data.

6.2 Human cognition modelling

When designing advanced ATM, it is important to understand the safety issues already at a conceptual level. Because of the extreme low probability of accidents in existing ATM practice, statistical data from practical situations is limited and analysing accident reports alone is not sufficient to understand safety at the level of the interactions between the various ATM components. For advanced ATM designs, data concerning unsafe events may even be lacking at all. Therefore, some kind of modelling approach is required to optimize for capacity and separation criteria without compromising safety.

Since in about eighty percent of the reported accidents humans were part of the cause, it is imperative to properly incorporate the human factor into the models used for risk assessment. In this report, we therefore investigated three complementary psychological models, and we combined them into a single mathematical model of a tactical ATCo in a conventional en-route context.

Because monitoring activity is typically performed as an integrated part of the tactical ATCo job, it is necessary to also take into account other ATCo activities that may interfere with monitoring. This was accomplished through our contextual model of ATCo performance that takes into account the interfering tasks at a cognitive level, thus minimizing the level of modelling detail required to take into account the interfering tasks. We also showed that this advanced ATCo performance model can be included in an accident risk model for the conventional en-route ATC situation considered, and that the time needed for the ATCo to detect a severe deviation as predicted by the model agrees rather well with statistical data. We also demonstrated that we could use the model to evaluate accident risk for the ATM scenario, and that the results provide valuable insight and feedback to ATM designers.

We conclude that the use of advanced psychological models in accident risk modelling is feasible, thus extending the applicability of the accident risk modelling approach to situations where isolated models of individual human actions do not suffice.

7. REFERENCES

AGARD, A designer's guide to human performance modelling, AGARD Advisory report 356, December 1998.

Amalberti, R, Wioland, L, Human error in aviation, In: Aviation safety, pp. 91-108, H. Soekkha (Ed.), 1997.

Ammerman, H.L., Fairhurst, W.S., Hostler, C.M., Jones, G.W., FAA air traffic control concepts volume VI: ARTCC/HOST en route controllers. DOT/FAA/AP/87-01. Washington DC: FAA, 1987.

Bainbridge, L., The change of concepts needed to account for human behaviour in complex dynamic tasks, Proc. 1993 Int. Conf. on Systems, Man and Cybernetics, vol 1, pp 126-131.

Bakker, G.J., Blom, H.A.P., Air Traffic Collision risk modelling, Proc. 32nd IEEE Conf. on Decision and Control, pp. 1464-1469, 1993.

Biemans, M.C.M., Daams, J., HOMEROS: Human Operator Modelling to Evaluate Reliability, Organisation and Safety, RHEA WP6 subtask report, NLR, June 1997.

Blom, H.A.P., Bakker, G.J., Blanker, P.G.J., Daams, J., Everdij, M.H.C., Klompstra, M.B., Accident risk assessment for advanced

ATM, Proc. 2nd USA/Europe ATM R&D Seminar, Orlando, FAA/Eurocontrol, 1998.

Blom, H.A.P., Everdij, M.H.C., Daams, J., Modern Safety Cases for a new operation in air traffic, ARIBA consolidation report Part II, NLR, 1999.

Buck, S., Biemans, M.C.M., Hilburn, B.G., van Woerkom, P.Th.L.M., Synthesis of functions, NLR Report TR 97054 L, 1996.

Cohen, S., Hockaday, S. (Eds.), A concept paper for separation safety modelling, FAA/Eurocontrol, May 1998.

Corker, K.M., Pisanich, G., Bunzo, M., Human Factors in advanced ATM system simulation studies, Proc. 1st USA/Europe ATM R&D Seminar, 1997.

Cox, M. Task analysis of selected operating positions within UK air traffic control, Royal Air Force Institute of Aviation Medicine, Report No. 749, Farnborough, 1994.

Daams, J., Bakker, G.J., Blom, H.A.P., Safety evaluation of an initial free flight scenario with TOPAZ (Traffic Organization and Perturbation AnalyZer), NLR Report TR-98098, 1998a.

Daams, J., Nijhuis, H.B., Blom, H.A.P., Accident risk assessment with a human cognition model using TOPAZ (Traffic Organization and Perturbation AnalyZer), NLR Report, 1998b.

Daams, J., Bakker, G.J., Blom, H.A.P., Safety evaluation of encounters between free flight equipped aircraft in a dual route structure, NLR Report TR-99577, 1999a.

Daams, J., Nijhuis, H.B., Blom, H.A.P., Human operators controllability of ATM safety, Report ARIBA WP4, NLR, 1999b, http://www.nlr.nl/public/hosted-sites/ariba/

EATCHIP, Model for Task and Job Descriptions of Air Traffic Controllers, Report HUM.ET.ST01.1000-REP-01, Eurocontrol, Brussels, 1996.

Embrey, D., Hyphreys, P., Rosa, E., Kirwan, B., Rea, K., SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgement, NUREG/CR-3518 US Nuclear Regulatory Comm., 1984.

Endsley, M.R., Rodgers, M.D., Situation awareness information requirements for en route air traffic control, Texas Univ., Lubeck, Dept. of Industrial Eng., 1994.

Everdij, M.H.C., Bakker, G.J., Blom, H.A.P., Application of collision risk tree analysis to DCIA/CRDA with support of TOPAZ, NLR report CR 96784, 1996.

Everdij, M.H.C., Blom, H.A.P., Klompstra, M.B. Dynamically Coloured Petri Nets for Air Traffic Management safety purposes, Proc. 8th IFAC Symp. on Transportation Systems, Chania, Crete, 1997a, pp. 184-189.

Everdij, M.H.C., Bakker, G.J., Blom, H.A.P., Blanker, P.J.G., Demonstration report in preparation to Designing EATMS inherently safe, Report TOSCA II WPR/4/01 Part I, NLR, 1997b.

George, P.H., Johnson, A.E., Hopkin, V.D., Radar monitoring of parallel tracks, automatic warning to controllers of track deviations in a parallel track system, EEC Report No 67, Bretigny, 1973.

Gibson, J.J., The ecological approach to visual perception, Lawrence Erlbaum Ass., 1986.

Haraldsdottir et al., Air Traffic Management Concept Baseline Definition, NEXTOR Report RR-97-3, Boeing, 1997.

Hollnagel, E., Human Reliability analysis, context and control, Academic press, London, 1993.

Hollnagel, E., Cognitive Reliability and Error Analysis Method CREAM, Elsevier Science Ltd, 1998.

ICAO, Annex 11 - Air Traffic Services, 12th edition, incorporating amendments 1-38, Green pages, attachment B, paragraph 3.2.1., July 1998.

Isaac, A., Ruitenberg, B., Air traffic control: human performance factors, Ashgate, 1999.

Jackson, A. The role of the controller in future ATC systems with enhanced information processing capabilities, EEC report No 224, 1989.

Kilner, A., Hook, M., Duck, R., Workload Assessment, Report TOSCA II WPR/8/01 Part I, NATS, 1997.

Kirwan, B., A guide to practical human reliability assessment, Taylor and Francis, 1994.

Kos, J., H.A.P. Blom, L.J.P. Speijker, M.B. Klompstra, and G.J. Bakker. Probabilistic wake vortex induced accident risk assessment, Proc. 3rd USA/Europe ATM R&D Seminar, 2000.

Norman, D., Bobrow, D., On data-limited and resource limited processing, J. of Cognitive Psychology, 7, pp 44-60, 1975.

Odoni, A.R. et al., Existing and required modelling capabilities for evaluating ATM systems and concepts, Report MIT, 1997.

Rasmussen, J., Skills, rules and knowledge: Signals, signs and symbols, and other distinction in human performance models, IEEE Tr. on System, Man and Cybernetics, Vol. 13, pp 257-266, 1983.

Reason, J., Human Error, Cambridge University Press, 1990.

Swain, A.D., Guttman, H.E., A handbook of human reliability analysis with emphasis on nuclear power plant applications, USNRC-Nureg/CR-1278, 1983.

Wickens, C.R., Attention and skilled performance. In: Human Skills (Chapter 4), Holding, D. (Ed.) New York, Wiley, 1989.

Wickens, C.R., Engineering, Psychology and Human Performance, Columbus: Merrill, 1992.

Wickens, C.D., Mavor, A.S., Parasuraman, R., McGee, J.P. (Eds.), The future of Air Traffic Control, Human Operators and Automation, National Academy Press, Washington DC, 1998.

Wiliams, J.C., A data-based method for assessing and reducing human error to improve operational performance, Proc. IEEE 4th Conf. on human factors and power plants, 1988, pp. 436-450.

ACRONYMS

ATC	Air Traffic Control
ATCo	Air Traffic Controller
ATM	Air Traffic Management
CAA	Civil Aviation Authority
DCPN	Dynamically Coloured Petri Net
DM	Decision Making
FMS	Flight Management System
FPM	Flight Path Monitoring
HEART	Human Error Assessment and Reduction Technique
HEI	Human Error Identification
HEP	Human Error Probability
HMI	Human Machine Interface
HOMEROS	Human Operator Models to Evaluate Reliability, Organisation and Safety
PRA	Probabilistic Risk Assessment
PSF	Performance Shaping Factors
R/T	Radio/Telephony
RNP	Required Navigation Performance
SLIM	Success Likelihood Index Methodology
STCA	Short Term Conflict Alert
TCAS	Traffic Collision Avoidance System
THERP	Technique for Human Error Rate Prediction
TLS	Target Level of Safety
TOPAZ	Traffic Organization and Perturbation AnalyZer