Safety Risk Assessment by Monte Carlo Simulation of Complex Safety Critical Operations

Henk A.P. Blom, Sybert H. Stroeve and Hans H. de Jong
blom@nlr.nl; stroeve@nlr.nl; hdejong@nlr.nl
National Aerospace Laboratory NLR
Amsterdam, The Netherlands

Abstract

This paper gives an overview of performing safety risk assessment of a safety critical operation with support of Monte Carlo (MC) simulation. The approach is outlined for an air traffic example involving aircraft departing from a runway, which is occasionally crossed by taxiing aircraft. At the airport considered, a Runway Incursion Alert System (RIAS) is installed to warn the air traffic controller in case of impending runway incursions. The paper explains the key issues to be mastered in performing a MC simulation supported safety risk assessment of this kind of operation. To begin with, one has to develop an appropriate simulation model, and a sound way to speed up the MC simulation based on this model. Complementary, one has to validate the simulation model versus the real operation, and the simulation supported approach has to be embedded within the safety risk assessment of the total operation. For this application example MC simulation results are given and the way of feedback to the design of the operation is outlined.

1 Introduction

Among the class of complex and safety critical industries, air traffic is an interesting example that poses exceptional challenges to advanced design. By its very nature, each aircraft has its own crew, and each crew is communicating with several human operators in different air traffic management (ATM) and airline operational control (AOC) centres on the ground in order to timely receive instructions critical to a safe flight. In addition, from an organisational perspective, air traffic involves interactions between many stake holders: pilots, air traffic controllers, airline operation centres, airport authorities, government regulators and the public travelling. Figure 1 highlights this characteristic feature of interplay between distributed decision making and safety criticality both for air traffic and for other complex or safety-critical industries, such as finance and nuclear and chemical plants. Among the safety critical industries, air traffic stands out regarding the many distributed levels of interactions in control and decision making.

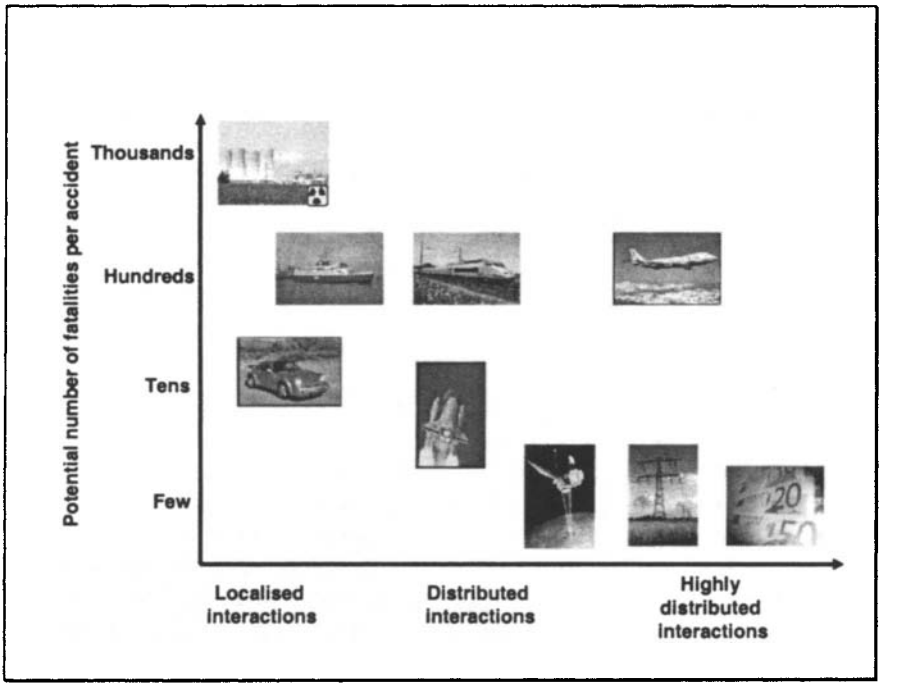

Figure 1: Air traffic compared with other complex and/or safety-critical industries in terms of potential number of fatalities per accident and the level of distributed interactions

The implication is that safety of air traffic is the result of interactions between multiple human operators, procedures (including spacing and separation criteria), and technical systems (hardware and software) all of which are highly distributed. Since safety depends crucially on the interactions between the various elements of a system, providing safety is more than making sure that each of these elements function properly. It is imperative to understand the safety impact of these interactions, particularly in relation to non-nominal situations.

Traditional ATM design approaches tend to be bottom-up, that is starting from developing concept elements aimed at increasing capacity, and next to extend the design with safety features. The advantage of the traditional approach is that advanced design developments can be organised around the clusters of individual elements, i.e., the communication cluster, the navigation cluster, the surveillance cluster, the automation tools cluster, the controllers/pilots and their human machine interfaces (HMIs), the advanced procedures, etcetera. The disadvantage of this traditional approach is that it fails to fully address the impact of interactions between controllers, pilots and ATM systems on safety.

A goal oriented approach would be to design ATM such that safety has been built in at the capacity-level required. From this perspective, safety assessment forms a primary source of feedback (Figure 2) in the development of advanced ATM

designs. An early guidance of ATM design development on safety grounds can potentially avoid a costly redevelopment program, or an implementation program that turns out to be ineffective. Although understanding this idea is principally not very difficult, it can be brought into practice only when an ATM safety assessment approach is available that provides appropriate feedback to the ATM designers from an early stage of the concept development (Figure 2). This feedback should provide information on which safety-capacity issues are the main contributor to unsafety.

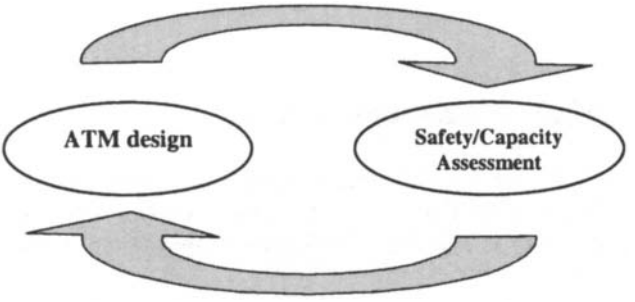

Figure 2: Safety feedback based ATM design.

For oceanic air traffic, the civil aviation community has developed a mathematical model to estimate mid-air collision risk levels as a function of spacing values in route structures (ICAO, 1988). This model is known as the Reich collision risk model; it assumes that the physical shape of each aircraft is a box, having a fixed orientation, and the collision risk between two aircraft is approximated by an expression that has proven to be of practical use in designing route structures (Hsu, 1981). Apart from the approximation, the most severe shortcoming is that the Reich model does not adequately cover situations where interaction between pilots and controllers play a crucial role, e.g. when controllers monitor the air traffic through surveillance systems and provide tactical instructions to the aircraft crews. In order to improve this situation, NLR has developed a safety risk assessment methodology which provides safety risk feedback to advanced air traffic operation design. The resulting safety risk assessment methodology has been named TOPAZ, which stands for Traffic Organization and Perturbation AnalyZer (Blom, 2001a). Within TOPAZ, Petri net modelling and Monte Carlo simulation has proven to deserve a key role in modelling and assessment of advanced air traffic operations on safety risk (Bakker and Blom, 1993; Blom et al., 2001b, 2003a,b,c; Everdij&Blom, 2002, 2003, 2005; Stroeve et al., 2003; Blom&Stroeve, 2004). In this respect it is relevant to notice that the use of Petri nets has been shown to work well in modelling safety critical operations in nuclear and chemical industries (e.g. Labeau et al., 2000).

The aim of this paper is to explain how the TOPAZ methodology effectively uses Monte Carlo simulation in safety risk assessment of an advanced air traffic operation. Emphasis is on how Monte Carlo simulation of safety risk works and how this is embedded within a complete safety risk assessment cycle.

This paper is organized as follows. First, section 2 provides an overview of the steps of the TOPAZ safety risk assessment cycle and for which step Monte Carlo simulation is of direct use. Next, section 3 provides an overview of how to develop a Monte Carlo simulation model of a given operation. In order to keep the explanation concrete, a particular example is introduced first. Subsequently section 4 provides an overview of key issues that have to be taken into account when using a Monte Carlo simulation supported safety risk assessment. Section 5 presents Monte Carlo simulation results for the particular example identified in section 3. Finally, conclusions are drawn in section 6.

2 Safety Risk Assessment Steps

An overview of the steps in a TOPAZ safety risk assessment cycle is given in Figure 3. Although the cycle itself is very much in line with the established risk assessment steps (e.g. Kumamoto and Henley, 1996), some of these steps differ significantly.

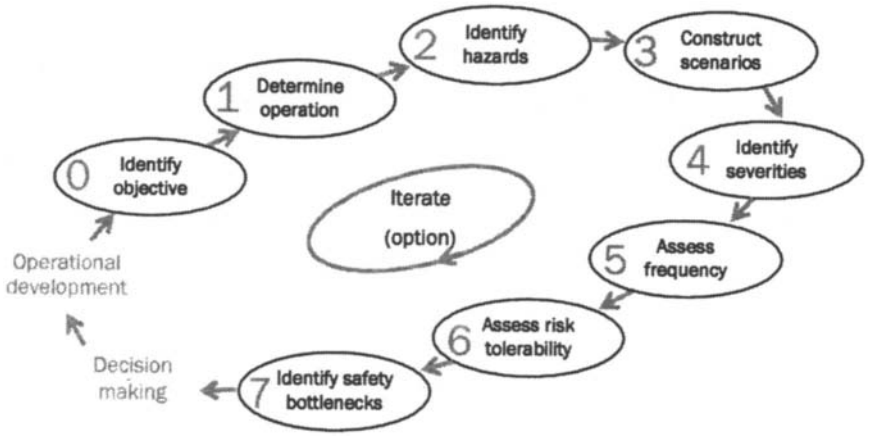

Figure 3: Steps in TOPAZ safety risk assessment cycle.

In step 0, the objective of the assessment is determined, as well as the safety context, the scope and the level of detail of the assessment. The actual safety assessment starts by determining the operation that is assessed (step 1). Next, hazards associated with the operation are identified (step 2), and aggregated into safety relevant scenarios (step 3). Using severity and frequency assessments (steps 4 and 5), the safety risk associated with each safety relevant scenario is classified (step 6). For each safety relevant scenario with a (possibly) unacceptable safety risk, the main sources contributing to unsafety (safety bottlenecks) are identified (step 7), which help operational concept developers to learn for which safety issues they should

develop improvements in the ATM design. If the ATM design is changed, a new safety risk assessment cycle of the operation should be performed in order to investigate how much the risk posed by previous safety issues has been decreased, but also to assess any new safety issues that may have been introduced by the enhancements themselves.
The following subsections present the risk assessment steps of a TOPAZ cycle in more detail. Then it also becomes clear that Monte Carlo simulation plays a key role in step 5: assess frequency.

Step 0: Identify objective
Before starting the actual safety assessment, the objective and scope of the assessment are set. This should be done in close co-operation with the decision makers and designers of the advanced operation. Also, the safety context must be made clear, such that the assessment is performed in line with the appropriate safety regulatory framework.
An important issue for setting the safety context is the choice of safety criteria with respect to which the assessment is performed. Depending of the application, such criteria are defined for particular *flight condition* categories (e.g. flight phases or sub-phases) and for particular *severity* categories (e.g. accident, serious incident). Typically, within the chosen context, these criteria define which *flight condition/severity* categories have to be evaluated and which frequency level forms the Target Level of Safety (TLS) threshold per *flight condition/severity* category.

Step 1: Determine operation
Step 1 serves for the safety assessors to obtain a complete and concise overview of the operation, and to freeze this description during each safety assessment cycle. Main input to step 1 is a description of the operational concept from the designers, while its output is a sufficiently complete, structured, consistent and concise description of the operation considered. The operation should be described in generic terms, the description should provide any particular operational assumption to be used in the safety assessment, and the description has to be verified by the operational concept designers. Typically during this step, holes and inconsistencies in the concept as developed are also identified and immediately fed back to the design team

Step 2: Identify hazards
The term hazard is used in the wide sense; i.e. an event or situation with possibly negative effects on safety. Such a non-nominal event or situation may evolve into danger, or may hamper the resolution of the danger, possibly in combination with other hazards or under certain conditions. The goal of step 2 is to identify as many and diverse hazards as possible. Hazard identification brainstorming sessions are used as primary means to identify (novel) hazards.
In system engineering, the functional approach to hazard identification is well-known. In this approach it is attempted to determine all possible failure conditions and their effects, for each function that plays a role in the operation, including the human operator tasks. Unfortunately, the approach cannot identify all hazards related

to an operation that involves human operators. An important reason for this is that the performance of air traffic controllers and pilots depend on their (subjective) situational awareness. From a human cognition perspective a particular act by an air traffic controller or pilot can be logical, while from a function allocation perspective the particular act may be incorrect. Such incidents are often called "errors of commission" (Sträter et al., 2004). An example of error of commision in the crossing operation is that because of the complicated taxiway structure, the pilot thinks to be taxiing far from the runway, while in reality, he starts crossing the runway without noticing any of the runway signs.

Another well-known technique of hazard identification is the HAZOP (HAZard and OPerability) method. With this method, hazards are identified and analyzed using sessions with operational experts. At the same time, the experts come up with potential solutions and measures to cope with the identified hazards (Kletz, 1999). The advantage of HAZOP with respect to the functional approach is that also non-functional hazards are identified during the brainstorm with operational experts. However, in applying HAZOP, one needs to take care that hazard analysis and solution activities do not disturb the hazard identification process, which could leave certain hazards unidentified or inappropriately "solved". Leaving such latent hazards in a design typically is known to be very costly in safety critical operation.

Based on the experience gained in using the hazard identification part of HAZOP in a large number of safety analyses and on scientific studies of brainstorming, NLR has developed a method of hazard identification for air traffic operations – by means of pure brainstorming sessions (De Jong, 2004). In such a session no analysis is done and solutions are explicitly not considered. An important complementary source is formed by hazards identified in previous studies on related operations. For this purpose, hazards identified in earlier studies are collected in a TOPAZ database.

Step 3: Construct scenarios

When the list of hazards is as complete as reasonably practicable, it is processed to deal with duplicate, overlapping, similar and ambiguously described hazards. First, per flight condition selected in Step 0, the relevant scenarios which may result from the hazards are to be identified using a full list of potentially relevant scenarios, such as for instance 'conflict between two aircraft merging onto one route' or 'aircraft encounters wake vortex of parallel departure'. Per *flight condition*, each potentially relevant scenario is subsequently used as crystallisation point upon which all applicable hazards and their combined effects are fitted. If hazards are not appropriately addressed by the crystals developed so far, then additional crystallisation points are defined. The output of such crystallisation process is a bundle of event/condition sequences and effects per crystallisation point, and these are referred to as a safety relevant scenario. This way of constructing scenarios aims to bring into account all relevant ways in which a hazard can play a role in each *flight condition/severity* category.

In order to cope with the complexity of the various possible causes and results, clusters of similarly crystallised hazards are identified. A cluster of hazards could for instance be the set of 'events causing a missed approach to deviate from the normal path'. An example is given in Figure 4. It should also be noticed that the same cluster

of hazards may play a role in one or more safety relevant scenerios.
Each of the identified hazards can be of the following types:

- a root hazard (cluster), which may cause a safety relevant scenario; or
- a resolution hazard (cluster), which may complicate the resolution of a safety relevant scenario.

For an appropriate safety risk assessment, all combinations of root and resolution hazards have to be evaluated in the next steps.

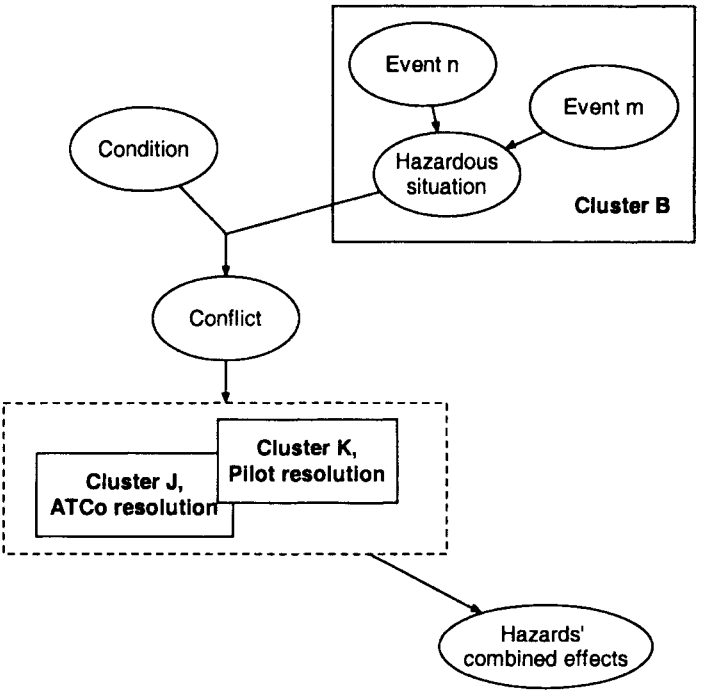

Figure 4: Example of a safety relevant scenario diagram.

Step 4: Identify severities

For each of the safety relevant scenarios identified in step 3, it is determined which of the *severity* categories selected in step 0 are applicable to its possible effects. Safety experts should assess which of the severities are applicable for each safety relevant scenario, by consultation of and review by operational experts. For each safety relevant scenario the effects and their severities depend on many factors, such as the conditions under which the scenario starts and evolves, the geometry of the scenario, and the possibilities of (timely) resolution of the conflict. Therefore, a range of *severities* may apply to a safety relevant scenario. If necessary, the structuring of the events in the safety relevant scenarios of step 3 are updated such that each applicable *severity* category is linked to the occurrence of specific event sequences.

Step 5: Assess frequency

Next, for each possible severity outcome of each safety relevant scenario, the occurrence frequency is evaluated by making use of an appropriate tree per safety relevant scenario. The probability of the top event in the tree is expressed as a sum of a product of probabilities of applicable conditional events. For assessing the factors in these trees, primary sources of data are operational experts and databases. Examples of databases are aviation safety databases, local controller reporting system(s), et cetera. For appropriate use of such data dedicated operational expertise is taken into account. Hence, important input for the frequency assessments is always formed by interviews with operational experts (including experienced pilots and controllers) who are as much as is possible familiar with the operation under consideration. Qualitative expressions are to be translated in quantitative numbers when the selected safety criteria of step 0 also are expressed in numbers. Complicating factors in assessing at once the frequency of a conflict ending in a given severity can be that there is often little or no experience with the new operation, and that the situation may involve several variables. This holds especially for the more severe outcomes of a safety relevant scenario, since these situations occur rarely, and consequently little information is at hands about the behaviour of air traffic controllers and pilots in such situations. For these difficult safety relevant scenarios it is logical to make use of Monte Carlo simulation of safety risk. This approach has three clear advantages: 1) the quality of the risk estimate improves; 2) it is possible to estimate a 95% confidence interval; and 3) once a MC simulation tool for a particular application has been developed it can be re-used for assessing safety risk for similar applications. The next sections explain for an example safety risk assessment by MC simulation.

Step 6: Assess risk tolerability

The aim of this step is to assess the tolerability of the risk for each of the *flight condition/severity* categories selected in step 0. First the total risk per *flight condition/severity* category is determined by summing over the assessed risk contributions per safety relevant scenario for that *flight condition/severity* category. This summation takes into account both the expected value and the 95% confidence interval of the risk summation. Next for each *flight condition/severity* category the total risk expected value and the 95% confidence interval are compared against the TLS selected in step 0.

Step 7: Identify safety bottlenecks

From the risk tolerability assessment, it follows which safety relevant scenario(s) contribute(s) most to the expected value and the 95% confidence interval of the risks that has been qualified as being not below the TLS. For each safety relevant scenario the hazards or conditions that contribute most to the high risk level or confidence interval are identified and localised during step 7. These are referred to as the safety bottlenecks. If desired, this may also be done for assessed risk levels that are just below the TLS. The identification and localisation of safety bottlenecks is important as it gives operational concept designers directions for searching potential risk

mitigating measures of the operation, and it gives the safety assessment experts the hazards and conditions for which the reduction of uncertainty has priority.

3 Monte Carlo Simulation Model

3.1 Active Runway Crossing Example

The Monte Carlo simulation-based risk assessment approach will be illustrated for an active runway crossing operation. This example accounts for a number of interacting human agents (pilots and controllers). The runway configuration of the active runway crossing operation considered is shown in Figure 5. The configuration takes into account one runway, named runway A, with holdings for using the runway from two sides (A1 and A2) and with crossings (C1, C2, D1 and D2) and exits (E1, E2, E3 and E4). The crossings enable traffic between the aprons and a second runway, named runway B. Each crossing has remotely controlled stopbars on both sides of the runway. Also the holdings have remotely controlled stopbars and each exit has a fixed stopbar.

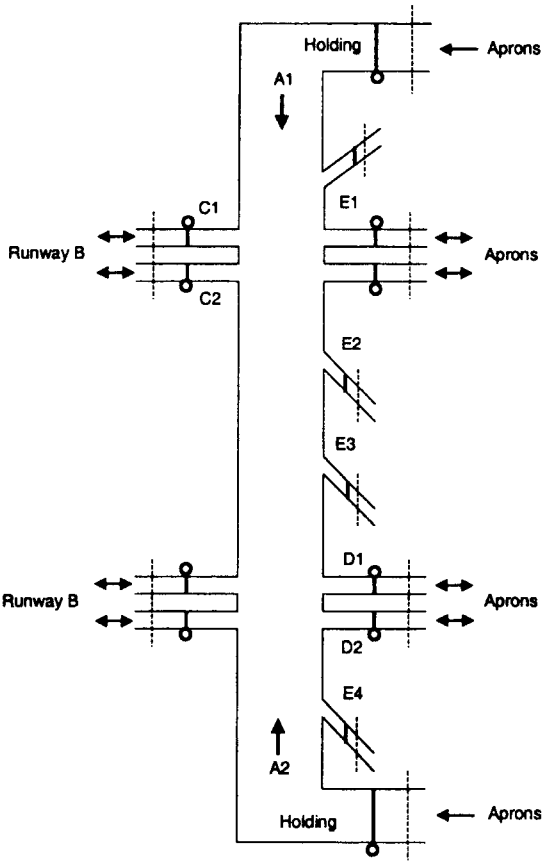

Figure 5: Runway configuration of active runway crossing procedure.

The involved human operators include the start-up controller, the ground controller, the runway A controller, the runway B controller, the departure controller, and the pilots flying (PF's) and pilots not flying (PNF's) of aircraft taking off and aircraft crossing. Communication between controllers and aircraft crews is via standard VHF R/T (Very High Frequency Receiver/Transmitter). Monitoring by the controllers can be by direct visual observation under sufficiently good visibility conditions; it is supported by ground radar surveillance. The runway A controller is supported by a runway incursion alert system and a stopbar violation alert system. The runway A controller manages the remotely controlled stopbars and the runway lighting. Monitoring by the aircraft crews is by visual observation, supported by the VHF R/T party-line effect.
In the runway crossing operation considered, the control over the crossing aircraft is transferred from the ground controller or the runway B controller (depending on the direction of the runway crossing operation) to the runway A controller. If the runway A controller is aware that the runway is not used for a take-off, the crew of an aircraft intending to cross is cleared to do so and subsequently the appropriate remotely controlled stopbar is switched off. The PNF of the crossing aircraft acknowledges the clearance and the PF subsequently initiates the runway crossing. When the crossing aircraft has vacated the runway, then the PNF reports this to the runway A Controller. Finally, the control over the aircraft is transferred from the runway A controller to either the runway B controller or the ground controller.

3.2 Safety Relevant Scenarios

Prior to the development of a quantitative accident risk model for the active runway crossing operation considered, all risk assessment steps had been performed using an expert-based approach. In this study the following safety relevant scenarios were found:

- Scenario I: Aircraft erroneously in take-off and crossing aircraft on runway;
- Scenario II: Aircraft erroneously crossing and other aircraft in take-off;
- Scenario III: Aircraft taking off and runway unexpectedly occupied;
- Scenario IV: Aircraft crossing and runway unexpectedly occupied by aircraft;
- Scenario V: Aircraft crossing and vehicle on runway;
- Scenario VI: Collision between aircraft sliding off runway and aircraft near crossing;
- Scenario VII: Aircraft taking off and vehicle crossing;
- Scenario VIII: Jet-blast from one aircraft to another; and
- Scenario IX: Conflict between aircraft overrunning/climbing out low and aircraft using a nearby taxiway.

From this expert-based study it followed that of all identified safety relevant scenarios, for scenarios I, II and III it was difficult to assess the risk sufficiently accurate using an expert based approach. For these three scenarios it is therefore useful to assess the risk through performing Monte Carlo simulations.
In this paper, we focus on the details of a Monte Carlo simulation accident risk

model for scenario II. In this scenario there is one aircraft that takes off and has been allowed to do so and there is one aircraft that crosses the runway while it should not. Taxiing along a straight line over one of the standard runway crossings (i.e., via C1, C2, D1 or D2 in Figure 5) is considered.

3.3 Multi-Agent Situation Awareness in the Simulation Model

The safe organisation of co-operation between pilots and controllers in air traffic depends to a large extent on the "picture" or situation awareness (SA) maintained by each of the pilots and controllers. When a difference, even a small one, sneaks into the individual pictures and remains unrecognised, this may create unnoticed miscommunication and a subsequent propagation and increase in differences between the individual pictures. Eventually the situation may spiral out of control, with potentially catastrophic results. Hence any mismatch between individual pictures forms a serious hazardous condition in maintaining a safe organisation. Many hazards identified for the runway crossing operation were of this type.

Endsley (1995) has defined human SA as the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future. Stroeve et al. (2003) and Blom and Stroeve (2004) have captured these perception, comprehension and projection notions of SA mathematically in terms of three components: State SA, Mode SA and Intent SA. They also extended this single (human) agent SA concept to a multi-agent SA concept for operations involving multiple humans and systems, inclusive the basic updating mechanisms of such multi-agent SA.

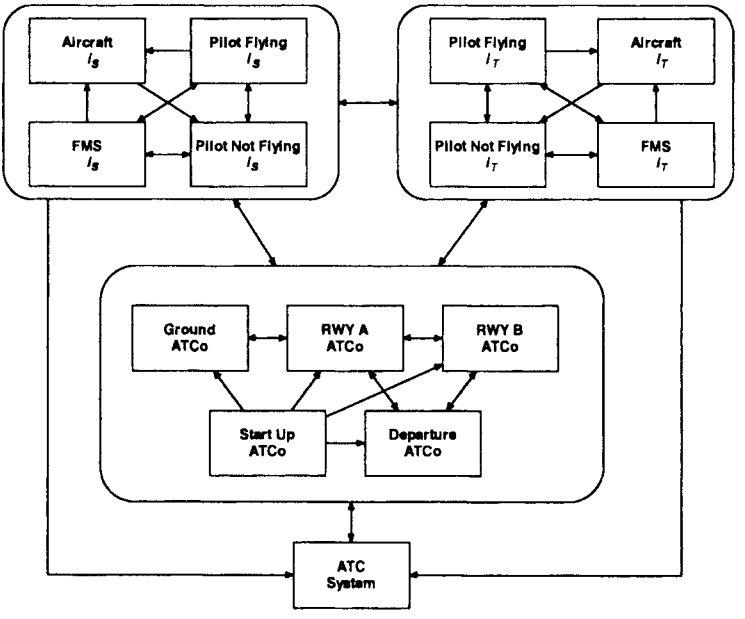

Figure 6: Relations between agents identified for the active runway crossing operation.

As depicted in Figure 6, for the active runway crossing operation we identified a need to model 10 agent types (7 humans and 3 systems) and their interactions:

- Pilots flying;
- Pilots not flying;
- (Each) aircraft;
- Aircraft's flight management systems (FMS);
- Runway A controller;
- Runway B controller;
- Ground controller;
- Departure controller;
- Start-up controller;
- ATC system, which we broadly define to include airport manoeuvre control systems, air traffic communication and surveillance systems, airport configuration and environmental conditions.

3.4 Dynamic Stochastic Modelling

The Monte Carlo simulations are based on dynamic stochastic models of all relevant agents. These simulation models are mathematically specified using the Dynamically Coloured Petri Net (DCPN) formalism (Everdij and Blom, 2003, 2005). A high-level overview of the agents modelled is provided next.

Taking-off Aircraft

The model of the taking-off aircraft represents the ground run, airborne transition and airborne climb-out phases and includes the possibility of a rejected take-off. The taking-off aircraft initiates its take-off from a position near the runway threshold and may have a small initial velocity. The aircraft may have diminished acceleration or deceleration power. Two types of aircraft are included in the model: medium-weight aircraft and heavy-weight aircraft.

Taxiing Aircraft

The model of the taxiing aircraft represents aircraft movements (hold, acceleration, constant speed, deceleration) during taxiing. The taxiing aircraft enters the taxiway leading to a runway crossing at a position close to the remotely controlled stopbar, with a normal taxiing speed or initiates taxiing from stance. The entrance time of the crossing aircraft is uniformly distributed around the take-off start time. The taxiing aircraft may have diminished deceleration power. Two types of aircraft are included in the model: medium-weight aircraft and heavy-weight aircraft.

Pilot Flying of Taking-off Aircraft

Initially, the pilot flying (PF) of a taking-off aircraft has the SA that taking-off is allowed and initiates a take-off. During the take-off the PF monitors the traffic situation on the runway visually and via the VHF communication channel. The PF starts a collision avoiding braking action if a crossing aircraft is observed within a critical distance from the runway centre-line or in reaction to a call of the controller, and if it is decided that braking will stop the aircraft in front of the crossing aircraft.

Further details of taking-off aircraft PF model are given by (Stroeve et al., 2003).

Pilot Flying of Taxiing Aircraft
Initially, the PF expects that the next airport way-point is either a regular taxiway or a runway crossing. In the former case the PF proceeds taxiing and in the latter case the PF may have the SA that crossing is allowed. The characteristics of the visual monitoring process of the PF depend on the intent SA. In case of awareness of a conflict, either due to own visual observation or due to a controller call, the PF stops the aircraft, unless it is already within a critical distance from the runway centre-line. Further details of taxiing aircraft PF model are given by (Stroeve et al., 2003).

Runway Controller
The runway A controller visually monitors the traffic and has support from a stopbar violation alert and a runway incursion alert. If the controller is aware that a taxiing aircraft has passed the stopbar, a hold clearance is given to both taxiing and taking off aircraft. Further details of the runway controller model are given by (Stroeve et al., 2003).

Radar Surveillance System
The model of the radar surveillance system represents position and velocity estimates for both aircraft. There is a probability that radar surveillance is not available, resulting in track loss. Radar surveillance data is used as basis for ATC stopbar violation alerting and ATC runway incursion alerting.

ATC Alerts
Two types of ATC alerts are included in the model: a stopbar violation alert and a runway incursion alert. A stopbar violation alert is presented to the controller if surveillance data indicates that an aircraft has passed an active stopbar. There is a probability that the stopbar violation alert system does not function, implying that there will be no alert. A runway incursion alert is presented to the controller if radar surveillance data indicates that the taxiing aircraft is within a critical distance of the runway centre-line and the taking-off aircraft has exceeded a velocity threshold in front of the runway crossing. There is a probability that the runway incursion alert system does not function, implying that there will be no alert.

VHF Communication Systems
The model for the VHF communication system between the runway controller and the aircraft crews accounts for the communication system of the aircraft, the communication system of the controller, the tower communication system, the frequency selection of aircraft communication system and the VHF communication medium. The nominal status of these communication systems accounts for direct non-delaying communication. The model accounts for a probability of delay in or failure of the communication systems.

4 Use of Simulation Model in Risk Assessment

Once the simulation model has been specified, there are several important aspects that have to be taken into account during the preparation, execution and interpretation of the Monte Carlo simulations. This section explains these aspects.

4.1 Does the Simulation Model Cover the Identified Hazards?

During step 2 of the safety assessment cycle, a lengthy list of hazards, including non-nominal situations, has been identified. These hazards contribute individually and possibly in combination with other hazards to the safety risk of the operation considered. Hence it is quite important to verify prior to performing the simulations that the hazards identified in step 2 of the assessment cycle are covered by the model. The verification process consists of specifying per hazard how it is captured by the simulation model. A special class of hazards is formed by the situation awareness related hazards. Table 1 shows three of such situation awareness related hazards and includes a short explanation how these hazards are covered by the simulation model.

Table 1: Examples of situation awareness related hazards and their simulation model.

Pilots become confused about their location at the airport because of complexity of the airport layout.	State SA of the PF of a taxiing aircraft is that its aircraft is at a location that differs from the actual location.
Crew of taxiing aircraft is lost and therefore not aware of starting to cross a runway.	Intent SA of PF is and stays taxiing while PF starts crossing the runway.
RIAS is switched off by maintenance and controllers are not informed.	RIAS working or not is not connected to Mode SA of controllers.

Inevitably this verification of each hazard against the model will lead to the identification of hazards that are not (yet) covered by the simulation model. For non-covered hazards the simulation model developers should consider to further extend the simulation model prior to performing Monte Carlo simulations.

4.2 Parametrisation of the Simulation Model

During the mathematical specification of the simulation model there is no need to bother about the correct parameter values to be used during the Monte Carlo simulation. Of course, this is addressed prior to running the simulations. In principle there are three kinds of sources for parameter values. The ideal source would consist of sufficient statistical data that has been gathered under the various contextual conditions for which the risk assessment has to be performed. In practice such ideal sources almost never exist. Instead one typically has to work with limited statistical data that has been gathered under different conditions. Fortunately there often are two complementary sources: domain expertise and scientific expertise (on safety and human factors). In the context of Monte Carlo

simulation this means one fuses statistical and expertise sources into a probability density function for the possible values of each parameter. Typically the mean or mode of such a density function is then used as the best estimate of the parameter value to be used when running the Monte Carlo simulation.

4.3 Speeding up Monte Carlo Simulations

Air traffic is a very safe means of transport. Consequently, the risk of collision between two aircraft is extremely low. The assessment of such low collision risk values through straightforward Monte Carlo simulation would need extremely lengthy computer simulation periods. In order to reduce this to practicable periods, five to six orders of magnitude in speeding up the Monte Carlo simulation are needed. The basis for realizing such speed-up factors in Monte Carlo simulation consists of decomposing accident risk simulations in a sequence of conditional Monte Carlo simulations, and then to combine the results of these conditional simulations into the assessed collision risk value. For the evaluation of logical systems good decomposition methods can often be obtained by Fault and Event Tree Analysis. Because air traffic operations involve all kinds of dependent, dynamic and concurrent feedback loops, these logic-based risk decomposition methods cannot be applied without adopting severe approximations, typically by assuming that events/entities happen/act independent of each other.
The stochastic analysis framework, that has shown its value in financial mathematics (e.g. Glasserman, 2004), is exploited by the TOPAZ methodology to develop Monte Carlo simulation models and appropriate speed-up factors by risk decomposition. The power of these stochastic analysis tools lies in their capability to model and analyse in a proper way the arbitrary stochastic event sequences (including dependent events) and the conditional probabilities of such event sequences in stochastic dynamic processes (Blom et al., 2003c; Krystul&Blom, 2004). By using these tools from stochastic analysis, a Monte Carlo simulation based risk assessment can mathematically be decomposed into a well-defined sequence of conditional Monte Carlo simulations together with a subsequent composition of the total risk out of these conditional simulation results. The latter composition typically consists of a tree with conditional probabilities to be assessed at the leaves, and nodes which either add or multiply the probabilities coming from the subbranches of that node. Within TOPAZ such a tree is referred to as a collision risk tree (Blom et al., 2001, 2003).

For the active runway crossing example, the particular conditions taken into account for this risk decomposition are:

- The type of each aircraft (either a medium-weight or a heavy-weight);
- The intent SA of the PF of a crossing aircraft concerning the next way-point (*Taxiway/Crossing*) and concerning allowance of runway crossing (*Allowed/Not Allowed*);
- The alert systems (functioning well or not);
- The remotely controlled stopbar (functioning well or not); and
- The communication systems (functioning well or not).

Based on the simulation model and the accident risk decomposition, Monte Carlo simulation software is developed to evaluate the event probabilities and the conditional collision risks, and to compose this with the help of the collision risk tree into the collision risk value assessed for the simulation model.

4.4 Validation of the assessed risk level

For operations as complex as the active runway example considered, a simulation model will always differ from reality. Hence, validation of the MC simulation results does not mean that one should try to show that the model is perfect. Rather one should identify the differences between the simulation model and reality, and subsequently analyse what the effects of these differences are in terms of *bias* and *uncertainty* at the assessed risk level of the model. If the bias and uncertainty fall within acceptable bounds, then the assesed risk levels are valid for the specified application. Otherwise one should improve the MC simulation model on those aspects causing the largest *bias* and *uncertainty* influence on the assessed risk level.
Five types of differences between simulation model and the real operation can be distinguished (Everdij and Blom, 2002):

- Numerical approximations;
- Parameter values;
- Assumptions on the model structure;
- Non-covered hazards;
- Differences between the real operational concept and the operational concept modelled.

Thinking in terms of these differences makes it possible to consider the validation problem as a problem of making the differences specific, assessing each difference and its effect on the collision risk, and subsequently decide if this is accurate enough (valid) or not (invalid) for the purpose aimed at. The effects of differences on the collision risk can mathematically be expressed in terms of bias and uncertainty that has to be taken into account when using the simulation model assessed risk value for decisions about reality:

- *Bias*. The accident risk as defined by the simulation model is systematically higher or lower than it is for the real operation.
- *Uncertainty*. In addition to a systematic bias, the differences between simulation model and reality may induce uncertainty in the difference between the safety risk of the real operation and the safety risk resulting from the simulation model.

With this, the validation of a simulation based accident risk assessment has largely become a bias and uncertainty assessment process. Within TOPAZ, a bias and uncertainty assessment method has been developed which consists of the following steps:

- Identify all differences between the simulation model and reality;
- Assess how large these differences are, or how often they happen;
- Assess the sensitivity (or elasticity) of the risk outcome of the simulation model to changes in parameter values;

- Assess the effect of each difference on the risk outcome, using model sensitivity knowledge and complementary statistical and/or expert knowledge;
- Combine the joint effects of all differences in bias and uncertainty factors, and compensate the risk value of the model with these bias and uncertainty factors.

The result is an expected value of risk for the real operation, including a 95% confidence interval of other possible risk values. If the bias or the 95% confidence interval of the combined effects, or the bias and uncertainty of individual differences is too large, then these differences have to be taken into account in the decision making process regarding the acceptability and/or further design of the operation considered.

5 Monte Carlo Simulation Results

This section presents collision risk results obtained by Monte Carlo simulation with a computer implementation of the mathematical model of the active runway example of section 3. In order to relate these results to an actual operation, a bias and uncertainty assessment remains to be performed; however, this falls outside the scope of this paper.

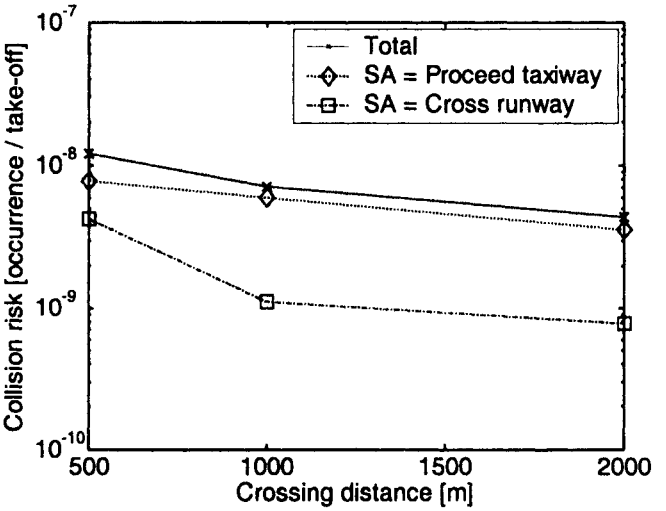

Figure 7: Contributions to the total collision risk by the simulation model for the cases that the SA of the PF of the taxiing aircraft is to proceed on a taxiway, or to cross the runway.

5.1 Assessed risk levels

Figure 7 shows the accident risk as function of the position of the runway crossing with respect to the runway threshold. The probability of a collision decreases for positions of the crossing distances further from the threshold. Figure 7 also shows the decomposition of the total risk for the cases that the pilot flying of the taxiing

aircraft either thinks to be proceeding on a normal taxiway (without being aware to be heading to a runway crossing) or where the pilot intends to cross the runway (without being aware that crossing is currently not allowed). The largest contribution to the risk is from the situation that the pilot thinks to be proceeding on a normal taxiway. The relative size of this contribution depends on the crossing distance and varies from 64% for crossing at 500 m to about 83% for crossing at 1000 or 2000 m.

A more complete overview of the contributions to the collision risk is provided by a projected version of the collision risk tree in Figure 8. It shows the contributions of events related to the situation awareness of the pilot of the taxiing aircraft (*Cross runway/Proceed runway*) and the functioning of ATC alert and communication systems (*Up/Down*). The collision risk results in the leaves of the tree are the product of the probability of the event combination indicated and the Monte Carlo simulation based collision risk given the event combination. The results in Figure 8 show that the risk is dominated by situations with a pilot flying of a taxiing aircraft having an erroneous situation awareness and the ATC alert and communication systems working nominally.

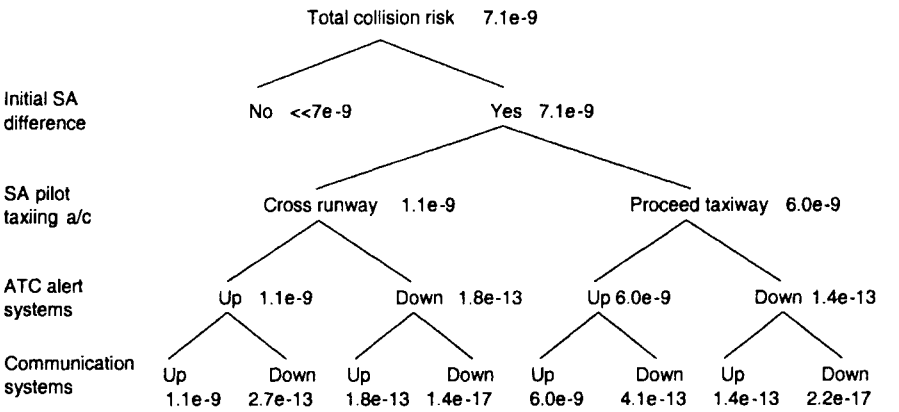

Figure 8: Projected version of the collision risk tree for the active runway crossing example, showing the contributions to the collision risk for various combinations of events related to pilot situation awareness and functioning of ATC alert and communication systems. The values are for a crossing distance of 1000 m.

5.2 Who contributes to safety risk reduction?

Based on results of the accident risk model, it is possible to attain insight in the accident risk reducing performance of involved human operators and technical systems. Table 2 shows conditional collision risks for the situation that an aircraft taxies towards a runway crossing at a distance of 1000 m from the runway threshold while the pilot has the situation awareness to taxi on a normal taxiway. The conditional collision risks in Table 2 refer to cases where the model either does ('yes') or does not ('no') involve the indicated human operators actively monitoring for traffic conflicts. A risk reduction percentage is determined by comparing the

conditional collision risk with the situation in which none of the human operators is actively monitoring. In this case, a collision is only avoided by the lucky circumstances that the taxiing aircraft just passes in front of or behind the taking-off aircraft (case 0 in Table 2). From the results in Table 2 a number of model-based insights into the operation can be attained:

- It follows from case 1 that 99.8% of the accidents can be prevented by the combined effort of all human operators and alert systems.
- It follows from a comparison of cases 1 and 5 that in the normal situation that all human operators are actively monitoring, ATC alert systems (runway incursion or stopbar violation) have a modest effect on the achieved risk.
- It follows from a comparison of cases 1 and 4, and cases 5 and 8, that the risk reduction that can be achieved by the tower controller in addition to the risk reduction of both pilots is very small.
- It follows from comparison of cases 1 and 3, and cases 5 and 7 that the pilot of the taxiing aircraft has the largest capability to prevent a collision in this context. Thus, resolution of the conflict is most likely to be by the human operator whose wrong situation awareness initiated the conflict.

Table 2: Risk reduction achieved in the simulation model by various combinations of involved human operators when the PF of a taxiing aircraft intends to proceed on a normal taxiway under good visibility (crossing is at 1000 m from runway threshold.)

0	no	no	no	$8.9\ 10^{-2}$	-
1	yes	yes	yes	$1.7\ 10^{-4}$	99.8%
2	yes	no	yes	$4.0\ 10^{-4}$	99.6%
3	no	yes	yes	$9.4\ 10^{-3}$	89.4%
4	yes	yes	no	$2.3\ 10^{-4}$	99.7%
5	yes	yes	yes	$2.2\ 10^{-4}$	99.8%
6	yes	no	yes	$1.7\ 10^{-3}$	98.1%
7	no	yes	yes	$1.1\ 10^{-2}$	87.9%
8	yes	yes	no	$2.3\ 10^{-4}$	99.7%

5.3 Comparison against expert based results

In the earlier conducted expert based safety risk assessment of the active runway crossing operation, it was concluded that both the pilots and the runway controller make large contributions to the prevention of a collision in the scenario aircraft erroneously crossing and other aircraft in take-off. In hindsight, it can be concluded that in the expert based safety risk assessment, the total effect of the pilots and the runway controller in preventing a collision turns out to be overestimated under good visibility condition. It is the simulation based approach that makes clear that although the runway controller identifies a good share of the conflicts, its

contribution to timely conflict resolution is relatively small. One significant part of the instruction issued by the runway controller appears to concern conflicts that are already solved by the pilots. And another significant part of the instructions issued by the runway controller appear to arrive too late for the pilots to successfully avoid a collision. Because of this, the effective contribution by the runway controller towards reducing collision risk is relatively small.

6 Concluding remarks

This paper has given an overview of performing safety risk assessment and providing feedback to the design of advanced air traffic operations with support of Monte Carlo simulation. The motivation for developing such a Monte Carlo simulation approach towards safety risk assessment was the identified need for modelling stochastic dynamic events and interactions between multiple agents (humans and systems) in advanced air traffic operations. The distributed and dynamical interactions pose even greater challenges than those seen in, for instance, nuclear and chemical industries (e.g. Labeau et al., 2000). The paper has explained the key issues to be mastered in performing a Monte Carlo simulation supported safety risk assessment of air traffic operations, and how this fits within a full safety risk assessment cycle. The steps to be followed in developing an appropriate Monte Carlo simulation model has been outlined, including a short overview of multi-agent situation awareness modelling, which plays a key role in the safe organization of cooperation between many pilots and controllers in air traffic. The paper also has explained the need for using stochastic analysis tools in order to develop the necessary speed-up of the Monte Carlo simulations, and has shown a feasible way to validate the simulation model versus the real operation. This assessment approach has been applied to an air traffic example involving aircraft departing from a runway that is occasionally crossed by taxiing aircraft. The results obtained demonstrate the feasibility and value of performing Monte Carlo simulation in accident risk assessment for safety relevant scenarios that are difficult to assess expert based, because of many interacting agents.

References

G.J. Bakker and H.A.P. Blom, (1993). 'Air Traffic Collision Risk Modeling, Proc. 32nd IEEE Conf. on Decision and Control, pp. 1464-1469

H.A.P. Blom G.J. Bakker, P.J.G. Blanker, J. Daams, M.H.C. Everdij and M.B. Klompstra (2001a). Accident risk assessment for advanced air traffic management. In: Donohue GL and Zellweger AG (eds.), Air Transport Systems Engineering, AIAA, pp. 463-480.

H.A.P. Blom, J. Daams and H.B. Nijhuis (2001b), Human cognition modelling in air traffic management safety assessment, Eds: G.L. Donohue and A.G. Zellweger, Air transport systems engineering, AIAA, pp. 481-511.

H.A.P. Blom, S.H. Stroeve, M.H.C. Everdij and M.N.J. van der Park (2003a), Human cognition performance model to evaluate safe spacing in air traffic, Human Factors and Aerospace Safety, Vol. 3, pp. 59-82

H.A.P. Blom, M.B. Klompstra and G.J. Bakker (2003b), Accident risk assessment of simultaneous converging instrument approaches, Air Traffic Control Quarterly, Vol. 11, pp. 123-155.

H.A.P. Blom, G.J. Bakker, M.H.C. Everdij and M.N.J. van der Park (2003c), *Collision risk modelling of air traffic*, Proc. European Control Conf. 2003 (ECC03), Cambridge, UK.

H.A.P. Blom and S.H. Stroeve (2004). Multi-agent situation awareness error evolution in air traffic. Proc. 7th Conference on Probabilistic Safety Assessment & Management, Berlin, Germany

H.H. De Jong (2004). Guidelines for the identification of hazards; How to make unimaginable hazards imaginable? National Aerospace Laboratory NLR, Contract report for EUROCONTROL, NLR-CR-2004-094

M.R. Endsley (1995). Towards a theory of situation awareness in dynamic systems. Human Factors, 37(1): 32-64

M.H.C. Everdij and H.A.P. Blom (2002), Bias and uncertainty in accident risk assessment. National Aerospace Laboratory NLR, NLR-TR-2002-137.

M.H.C. Everdij and H.A.P. Blom (2003). Petri-nets and hybrid-state Markov processes in a power-hierarchy of dependability models. In: Engel, Gueguen, Zaytoon (eds.), Analysis and design of hybrid systems, Elsevier, pp. 313-318

M.H.C. Everdij and H.A.P. Blom (2005), Piecewise deterministic Markov processes represented by dynamically coloured Petri nets, Stochastics Vol. 77, pp.1-29

P. Glasserman (2004), Monte Carlo methods in financial engineering, Springer.

D.A. Hsu (1981). 'The evaluation of aircraft collision probabilities at intersecting air routes', J. of Navigation, Vol.34, pp.78-102

ICAO (1988). Review of the General Concept of Separation Panel, 6th meeting, Doc 9536, Volume 1, ICAO, Montreal.

T. Kletz (1999), Hazop and Hazan; identifying and assessing process industry hazards, The Institution of Chemical Engineers, 4th ed.

J. Krystul and H.A.P. Blom (2004). Monte Carlo simulations of rare events in hybrid systems. Hybridge report D8.3, http://hosted.nlr.nl/public/hosted-sites/hybridge/

H. Kumamoto and E.J. Henley (1996), Probabilistic Risk Assessment and management for engineers and scientists, IEEE Press.

P.E. Labeau, C. Smidts and S. Swaminathan (2000), Dynamic reliability: towards an in-tegrated platform for probabilistic risk assessment, J. Reliability Engineering and System Safety, Vol. 68, pp. 219-254

O. Sträter, V. Dang, B. Kaufer and A. Daniels (2004). On the way to assess errors of commission. Reliability Engineering and System Safety 83:129-138

S.H. Stroeve, H.A.P. Blom and M.N.J. Van der Park (2003). Multi-agent situation awareness error evolution in accident risk modelling. 5th USA/Europe ATM R&D Seminar, Budapest