NLR Air Transport Safety Institute
Research & Consultancy

Executive summary

CONTRASTING SAFETY ASSESSMENTS OF A RUNWAY INCURSION SCENARIO: EVENT SEQUENCE ANALYSIS VERSUS MULTI-AGENT DYNAMIC RISK MODELLING

Problem area
In the safety literature it has been argued, that in a complex socio-technical system safety cannot be well analysed by event sequence based approaches, but requires to capture the complex interactions and performance variability of the socio-technical system. In order to evaluate the quantitative and practical consequences of these arguments, this study compares two approaches to assess accident risk of an example safety critical sociotechnical system.

Description of work
This study contrasts an event sequence based assessment with a multi-agent dynamic risk model (MA-DRM) based assessment, both of which are performed for a particular runway incursion scenario. The event sequence analysis uses the well-known event tree modelling formalism and the MA-DRM based approach combines agent based modelling, hybrid Petri nets and rare event Monte Carlo simulation. The comparison addresses qualitative and quantitative differences in the methods, attained risk levels, and in the prime factors influencing the safety of the operation.

Results and conclusions
The assessments show considerable differences in the accident risk implications of the performance of human operators and technical systems in the runway incursion scenario. In contrast with the event sequence based results, the MA-DRM based results show that the accident risk is not manifest from the performance of and relations between individual human operators and technical systems. Instead, the safety risk emerges from the totality of the performance and interactions in the agent based model of the safety critical operation considered, which coincides very well with the argumentation in the safety literature.

Applicability
Safety assessment of air traffic operations.

Report no.
NLR-TP-2013-284

Author(s)
S.H. Stroeve
H.A.P. Blom
G.J. Bakker

Report classification
UNCLASSIFIED

Date
July 2013

Knowledge area(s)
Vliegveiligheid (safety & security)

Descriptor(s)
Dynamic risk model
Event tree
Runway incursion
Accident risk
Petri net

This report is based on an article published in Reliability Engineering and System Safety, Volume 109, pages 133-149, January 2013.

NLR Air Transport Safety Institute
Research & Consultancy

NLR Air Transport Safety Institute

Anthony Fokkerweg 2, 1059 CM Amsterdam,
P.O. Box 90502, 1006 BM Amsterdam, The Netherlands
Telephone +31 88 511 35 00, Fax +31 88 511 32 10, Web site: http://www.nlr-atsi.nl

NLR Air Transport Safety Institute
Research & Consultancy

NLR-TP-2013-284

CONTRASTING SAFETY ASSESSMENTS OF A RUNWAY INCURSION SCENARIO: EVENT SEQUENCE ANALYSIS VERSUS MULTI-AGENT DYNAMIC RISK MODELLING

S.H. Stroeve
H.A.P. Blom
G.J. Bakker

This report is based on an article published in Reliability Engineering and System Safety, Volume 109, pages 133-149, January 2013.

Customer	National Aerospace Laboratory NLR
Owner	National Aerospace Laboratory NLR
Division	Air Transport
Distribution	Unlimited
Classification of title	Unclassified
	July 2013

Approved by:

Author	Reviewer	Managing department
S.H. Stroeve	External anonymous revierwers	A.D.J. Rutten
Date: 8-7-2013	Date: 8-7-2013	Date: 8-7-2013

SUMMARY

In the safety literature it has been argued, that in a complex socio-technical system safety cannot be well analysed by event sequence based approaches, but requires to capture the complex interactions and performance variability of the socio-technical system. In order to evaluate the quantitative and practical consequences of these arguments, this study compares two approaches to assess accident risk of an example safety critical sociotechnical system. It contrasts an event sequence based assessment with a multi-agent dynamic risk model (MA-DRM) based assessment, both of which are performed for a particular runway incursion scenario. The event sequence analysis uses the well-known event tree modelling formalism and the MA-DRM based approach combines agent based modelling, hybrid Petri nets and rare event Monte Carlo simulation. The comparison addresses qualitative and quantitative differences in the methods, attained risk levels, and in the prime factors influencing the safety of the operation. The assessments show considerable differences in the accident risk implications of the performance of human operators and technical systems in the runway incursion scenario. In contrast with the event sequence based results, the MA-DRM based results show that the accident risk is not manifest from the performance of and relations between individual human operators and technical systems. Instead, the safety risk emerges from the totality of the performance and interactions in the agent based model of the safety critical operation considered, which coincides very well with the argumentation in the safety literature.

Contents

1	Introduction	1
2	Active runway crossing operation	4
3	Event sequence based safety study of the runway incursion scenario	6
3.1	Event tree	6
3.2	Quantification of the event tree	8
3.3	Accident risk reduction contributions of entities in the event tree	10
4	Multi-agent dynamic risk model based approach	12
4.1	Agent based modelling and simulation	12
4.2	Stochastic hybrid automata	13
4.3	Petri net based specification of a GSHP	14
4.4	Rare event Monte Carlo simulation of SDCPN	15
4.5	Evaluating differences between the SDCPN based model and the real operation	16
5	MA-DRM based safety study of the runway incursion scenario	18
5.1	SDCPN model of the runway incursion scenario	18
5.1.1	Taking-off aircraft (AC-TO)	19
5.1.2	Taxiing aircraft (AC-TX)	20
5.1.3	ATC system	20
5.1.4	Pilot flying of taxiing aircraft (PF-TX)	20
5.1.5	Pilot flying of the taking-off aircraft (PF-TO)	21
5.1.6	Runway controller (ATCo-R)	22
5.2	Accident risk assessment results for the runway incursion scenario	22
5.3	Analysis of agent based events in the runway incursion scenario	23
5.3.1	ATC alerts	26
5.3.2	Runway controller (ATCo-R)	26
5.3.3	Pilot flying of taking-off aircraft (PF-TO)	26
5.3.4	Pilot flying of taxiing aircraft (PF-TX)	27
5.3.5	Taking-off aircraft (AC-TO)	27
5.3.6	Taxiing aircraft (AC-TX)	28

5.4	Risk effects due to placing agents out of the loop or monitoring role	30
6	COMPARISON OF THE SAFETY ASSESSMENT STUDIES	35
6.1	Comparison of the architecture of the models	36
6.1.1	Model complexity	36
6.1.2	Dynamics	36
6.1.3	Performance variability	36
6.1.4	Interactions / concurrency	37
6.1.5	Emergent behaviour	38
6.2	Comparison of the use of the models	38
6.2.1	Expertise and techniques needed	38
6.2.2	Variety of contextual conditions	39
6.2.3	Transparency	40
6.2.4	Uncertainty	41
6.3	Comparison of the risk results of the models	42
6.3.1	Differences in findings	42
6.3.2	Comparison against real data	43
6.3.3	Feedback to design	43
7	DISCUSSION	44
8	REFERENCES	47

This page is intentionally left blank.

1 INTRODUCTION

The man-made disasters theory of Turner [1] gives early descriptions of how the objective of safely operating technological systems could be subverted by normal organizational processes due to unintended and complex interactions between contributory preconditions. Also Perrow [2] describes accidents as the consequence of complex interactions and tight couplings in sociotechnical systems in his normal accident theory. Building forward on the notion of normal accidents, Hollnagel [3] argues that performance in sociotechnical systems is necessarily variable due to the performance variability of its entities and the complexity of their interactions. Dekker [4] uses complex systems theory to qualitatively discuss safety in complex organizations and accidents as emergent properties.

In order to cope quantitatively with the challenge of safety risk assessment for a complex sociotechnical system, Zio [5] has performed a systematic analysis of the various issues that have to be addressed. Based on this analysis, Zio [5] identifies a need for a methodology that integrates dynamic and stochastic behaviour and that automatically generates various scenarios through dynamic simulation. Such methodologies use models of controlled process dynamics and human operator behaviour during safety-relevant scenarios, and simulation techniques such as Monte Carlo simulation to evaluate the models. Advantages of a dynamic simulation approach indicated by Zio are (1) the identification of a broad range of accident scenarios, (2) the exclusion of oversimplifying assumptions about process evolution, since processes are simulated directly by dynamic models, and (3) the retrieval of additional information on time-dependent probability density functions of process states from the analysis.

In spite of these valuable views on safety of complex sociotechnical systems, it still is common practice to adopt classical event sequence based approaches for safety assessment. Such classical approaches use sequential cause-effect propagation of technical failures, human errors, contextual conditions and conflict resolving actions to model the development of accidents. Well-known techniques are fault trees and event trees, which represent relations between event occurrences and use event probabilities to achieve quantification of risk levels [6]. An advantage of these techniques is that the resulting tree structure is easy to understand by a large audience, but recognized limitations include the difficulty to represent the dynamics and interdependencies between entities of

safety relevant scenarios, and their limited account of human performance [3, 7, 8]. If the argumentations by [1-5] are correct, then one should expect that the adoption of an event sequence based approach for a safety risk assessment of a complex sociotechnical system would have significant impact on its outcomes. The aim of this paper is to evaluate this expectation by conducting a systematic comparison between an event sequence based approach and an advanced dynamic simulation based approach, both applied to the same safety critical air traffic control (ATC) example scenario.

The safety critical ATC example is a runway incursion scenario in the context of an active runway crossing operation. A runway incursion is defined by the International Civil Aviation Organization (ICAO) as “Any occurrence at an aerodrome involving the incorrect presence of an aircraft, vehicle or person on the protected area of a surface designated for the landing and take-off of aircraft” [9]. Within air traffic, runway incursion is recognised as an important safety issue [9, 10].

The development of the event sequence and MA-DRM based safety assessment approaches for the runway incursion scenario stemmed from a need to conduct a safety risk assessment of an active runway crossing operation at Amsterdam airport; these developments have been reported in preceding studies [11-13]. The design of this active runway crossing operation included an ATC alert system, which was aimed at minimizing the risk of runway incursions. During the development of the infrastructure and the operational concept for the active runway crossing operation, a series of risk assessment studies has been conducted [14]. Initial safety studies included event sequence based safety risk assessment of various safety relevant scenarios of the active runway crossing operation [11]. Having recognized the difficulty in capturing the complexity of possible runway incursion within the active runway crossing example, [12, 13] developed a MA-DRM approach for this very same scenario. Because both approaches have been developed for the same runway incursion scenario, and under the very same set of identified hazards, the approaches and results of these two studies provide a suitable basis for the comparison of event sequence and MA-DRM based risk assessment approaches for a complex safety critical socio-technical system example.

The comparison in this paper focuses on the risk of an accident in the runway incursion scenario as assessed by both approaches. It addresses differences in the methods, differences in the risk results attained and differences in the understanding of the factors influencing the safety of the operation. The

comparison is backed-up by dedicated simulations of the model of [13], which provide additional insights in the relation between the accident risk and events in the scenario, and the effect of the roles of agents in the scenario on the accident risk.

This paper is organized as follows. Section 2 describes the active runway crossing considered. Section 3 describes the methods and results of the event sequence based safety assessment of the runway incursion scenario. Section 4 introduces the MA-DRM based safety assessment approach. Section 5 describes the application of the MA-DRM based safety assessment of the runway incursion scenario. Section 6 compares the methods and results of both approaches. Section 7 provides a discussion of the implications of the differences identified between the two approaches.

Parts of the results in this paper have been presented in conference papers [15-17].

2 ACTIVE RUNWAY CROSSING OPERATION

In the runway incursion scenario considered, an aircraft is taking off and a taxiing aircraft is crossing the runway while it should not; thus the runway incursion is due to the taxiing aircraft. It may occur in the context of the active runway crossing operation depicted in Figure 1. The runway is used for departures and has a taxiway that crosses the runway at a distance of 1000 m from the runway threshold. The runway crossing has stopbars that are remotely controlled by the runway controller. The scenario considered is under good visibility conditions.

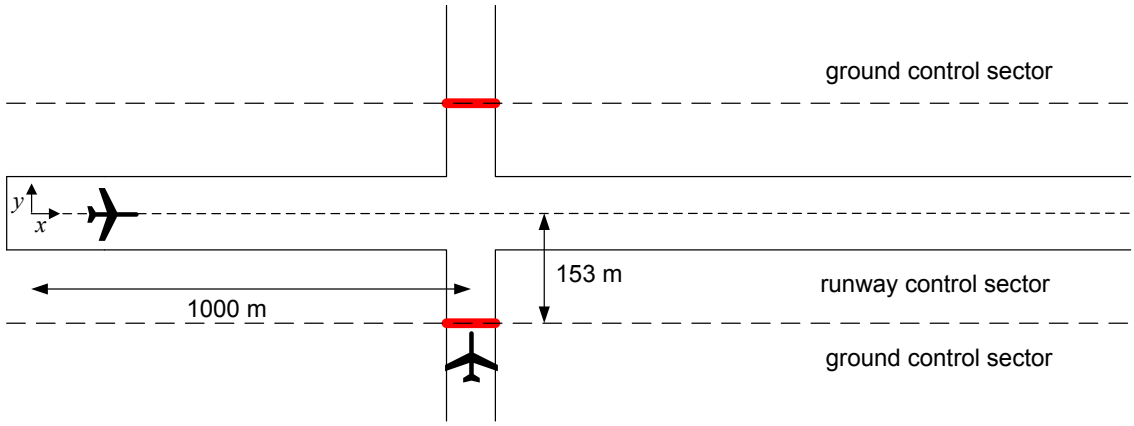

Figure 1: Schematic overview of the traffic situation considered. The taking-off aircraft accelerates along the runway while the crew of the taxiing aircraft intends to proceed along the taxiway towards the active runway.

The main human operators involved in the runway crossing operation are the pilots of the taking-off aircraft, the pilots of the taxiing aircraft, the runway controller and the ground controllers responsible for traffic on nearby taxiways. The pilots are responsible for safe conduct of the flight operations and should actively monitor for potential conflicting traffic situations. The runway controller is responsible for safe and efficient traffic handling on the runway and the runway crossings; the ground controllers are responsible for the traffic on the taxiways in the surroundings of the runway.

Standard communication, navigation and surveillance systems are used: communication between controllers and crews is by radio/telecommunication (R/T) systems, the pilots use their knowledge on the aerodrome layout and/or

their maps for taxiing, and ground radar tracking data of all aircraft and sufficiently large vehicles on the airport surface is shown on displays of the runway and ground controllers. The ATC system may generate two types of alerts to warn the runway controller: (1) a stopbar violation alert for the situation that an aircraft crosses an active stopbar in the direction of the runway, and (2) a runway incursion alert for the situation that an aircraft is on the runway in front of an aircraft that has initiated to take off.

3 EVENT SEQUENCE BASED SAFETY STUDY OF THE RUNWAY INCURSION SCENARIO

This section describes the event tree that was developed for the runway incursion scenario and its risk results. Section 3.1 presents the structure of the developed event tree. Section 3.2 presents the quantification of the event tree. Section 3.3 presents an analysis of the accident risk reduction achieved by pilots, controller and ATC alert system.

3.1 EVENT TREE

The developed event tree is shown in Figure 2. The starting event Q_0 in this tree is the situation that the taxiing aircraft starts crossing while it should not. The crossing is initiated by the pilots without contacting the runway controller, e.g. due to a misunderstanding of the ground controller. Subsequent events in this tree capture possible contributions to resolution of the runway incursion conflict by the pilots of both aircraft directly or following a call by the runway controller, who may have recognized the conflict directly or via an alert. The branching points in the event tree differentiate between early, medium and late recognition of the conflict by the pilots and the runway controller; early, medium and late communication between the controller and the pilots; and early and medium-timed alerts warning the controller (events Q_1 to Q_{12}). This approach was chosen as a systematic means to get hold on the variety in the timing of conflict detection and resolution events by the human operators in combination with the timing of the alerts and the remaining braking distance. Figure 2 shows that there are 27 event sequences S1 to S27. The outcomes of these event sequences are classified in the categories “No conflict”, “Early resolution”, “Medium resolution”, “Late resolution” and “Accident”. Figure 2 shows that there are six event sequences that lead to an accident between the two aircraft.

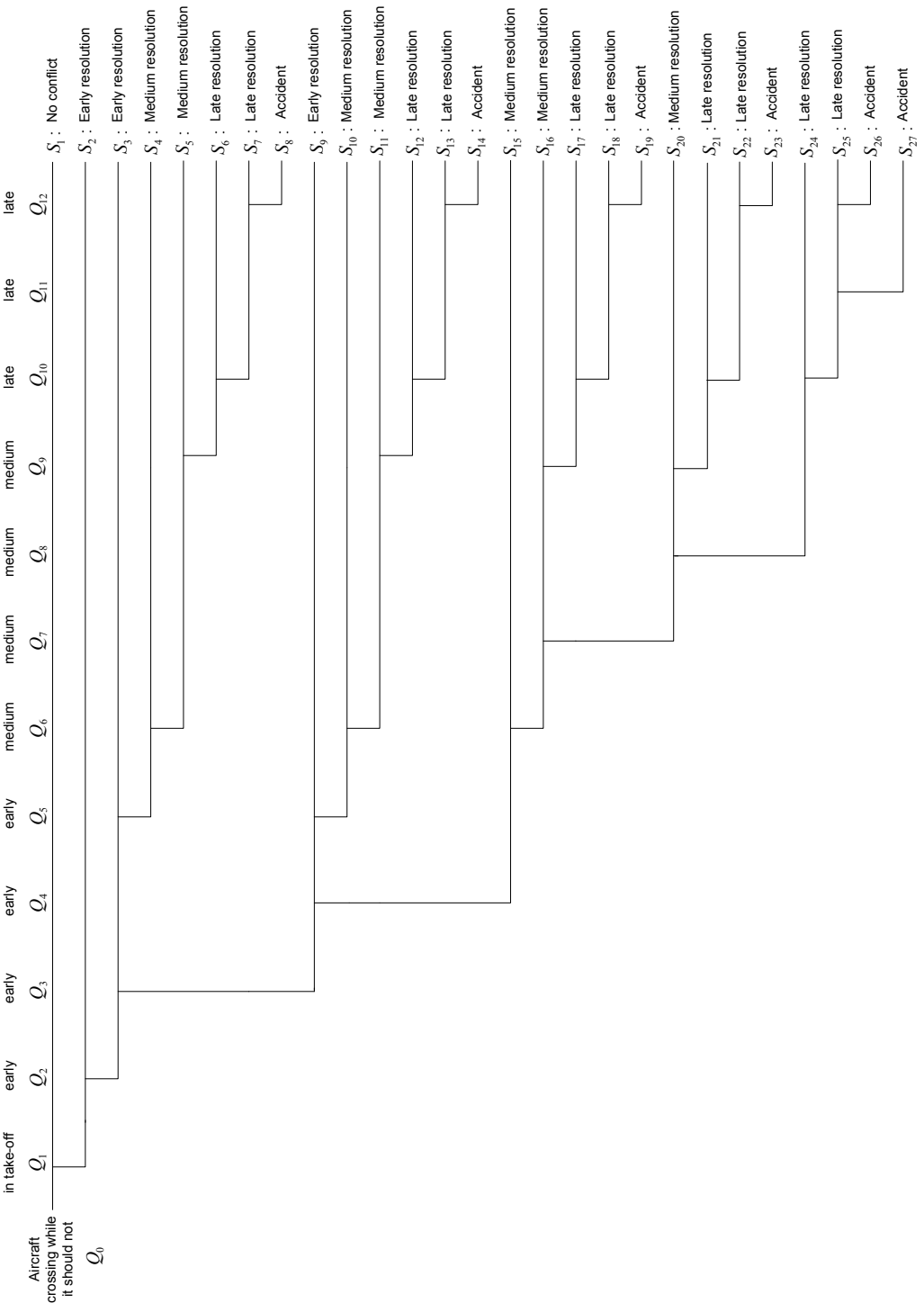

Figure 2: Event tree for resolution and recognition of the runway incursion scenario

3.2 QUANTIFICATION OF THE EVENT TREE

Quantification of an event tree such as in Figure 2 implies evaluation of the probabilities of the event sequences. For instance, the probability of event sequence S_3 (early resolution of the conflict following early recognition and communication by the controller) is evaluated as

$$P(S_3) = P(Q_0) \cdot P(\bar{Q}_1 \mid Q_0) \cdot P(\bar{Q}_2 \mid Q_0, \bar{Q}_1) \cdot P(Q_3 \mid Q_0, \bar{Q}_1, \bar{Q}_2) \cdot P(Q_5 \mid Q_0, \bar{Q}_1, \bar{Q}_2, Q_3) \quad (1)$$

Note that the event probabilities are conditional upon prior events in the tree. For instance, the probability of event Q_3 is conditional on the occurrence of Q_0 and the non-occurrence of Q_1 and Q_2: $P(Q_3 \mid Q_0, \bar{Q}_1, \bar{Q}_2)$, meaning the probability that the controller recognizes the conflict at an early stage, given there is an aircraft crossing while it should not, there is an aircraft in take-off and the conflict has not been recognized and resolved at an early stage by the pilots.

Table 1: Estimated lower and upper bounds of event probabilities, which are obtained by expert elicitation and are used in the quantification of the event tree of Figure 2.

Event		Event probability	
		Lower bound	Upper bound
Q_1	No aircraft in take-off	0.75	0.75
Q_2	Pilots recognize conflict at early stage	0.5	0.7
Q_3	Controller recognizes conflict at early stage	0.1	0.2
Q_4	Alert system warns controller at early stage	0.95	0.99
Q_5	Communication leads to resolution at early stage	0.8	0.9
Q_6	Pilots recognize and resolve conflict at medium stage	0.9	0.99
Q_7	Controller recognizes conflict at medium stage	0.2	0.4
Q_8	Alert system warns controller at medium stage	0.9	0.99
Q_9	Communication leads to resolution at medium stage	0.6	0.8
Q_{10}	Pilots recognize and resolve conflict at late stage	0.9	0.99
Q_{11}	Controller recognizes conflict at late stage	0.5	0.75
Q_{12}	Communication leads to resolution at late stage	0.4	0.6

In the event sequence based assessment, lower and upper bounds of the event probabilities were estimated by expert (controller and pilot) elicitation [11]. In particular, interviews were conducted in which the runway incursion scenarios were discussed with the experts and their opinions were asked about the

frequency of events in these scenarios. The thus found ranges of the event probabilities are shown in Table 1. Depending on the agent and the early/medium/late stage, the probabilities of the events (leading to resolution of the conflict) are in the range of 0.1 to 0.99. It can be noticed that the communication-related events Q_5, Q_9 and Q_{12}, and the pilot recognition-related events Q_6 and Q_{10} are used at multiple places in the event tree and that the probabilities of these events have been assumed to be the same, irrespective of the preceding chain of events.

Using these event probabilities in the event tree structure of Figure 2, the probabilities of the event sequences in the scenario are computed. For each of the outcome categories, the following conditional probabilities given the runway incursion are shown in Table 2: (1) a lower bound , which is based on the upper bounds of the event probabilities of Table 1, (2) an upper bound, which is based on the lower bounds of the event probabilities of Table 1, and (3) a geometric mean of these bounds. The geometric mean is calculated to support the comparison with the risk point estimate obtained by the MA-DRM and it assumes a multiplicative uncertainty range (i.e. a same factor above and below the geometric mean).

Table 2: Conditional probabilities given the runway incursion scenario per event tree outcome category, which are the sums of the probabilities of the related event sequences. For each category, the lower and upper bounds and their geometric mean are shown.

Event tree outcome category	Related event sequences	Probability		
		Lower bound	Geometric mean	Upper bound
No conflict	S_1	7.5 E-1	7.5 E-1	7.5 E-1
Early resolution	S_2, S_3, S_9	2.2 E-1	2.3 E-1	2.4 E-1
Medium resolution	S_4, S_5, S_{10}, S_{11}, S_{15}, S_{16}, S_{20}	8.0 E-3	1.5 E-2	2.8 E-2
Late resolution	S_6, S_7, S_{12}, S_{13}, S_{17}, S_{18}, S_{21}, S_{22}, S_{24}, S_{25}	1.6 E-5	1.3 E-4	1.1 E-3
Accident	S_8, S_{14},S_{19}, S_{22}, S_{26}, S_{27}	6.5 E-8	2.2 E-6	7.3 E-5

3.3 ACCIDENT RISK REDUCTION CONTRIBUTIONS OF ENTITIES IN THE EVENT TREE

To obtain a better insight in the accident risk reduction contributions of the alert system, the controller and the pilots, the accident risks of the event tree (Figure 2) are calculated for cases where one or several of these entities do not timely detect the conflict. These conditions are achieved as follows:

- The alert system does not timely warn the controller, is achieved by setting the probability of events Q_4 and Q_8 to zero.
- The controller does not recognize the conflict him/herself, is achieved by setting the probability of events Q_3 , Q_7 and Q_{11} to zero.
- The pilots do not recognize the conflict themselves, is achieved by setting the probability of events Q_2 , Q_6 and Q_{10} to zero.

The geometric mean of the conditional probability of an accident given the runway incursion scenario is shown in Figure 3 for each of the eight possible combinations of these conditions. Figure 3 also presents for each combination of conditions, the risk increase factor with respect to the case evaluated in Section 3.2 (referred to as case B1), where all entities contribute to detection and resolution of the conflict.

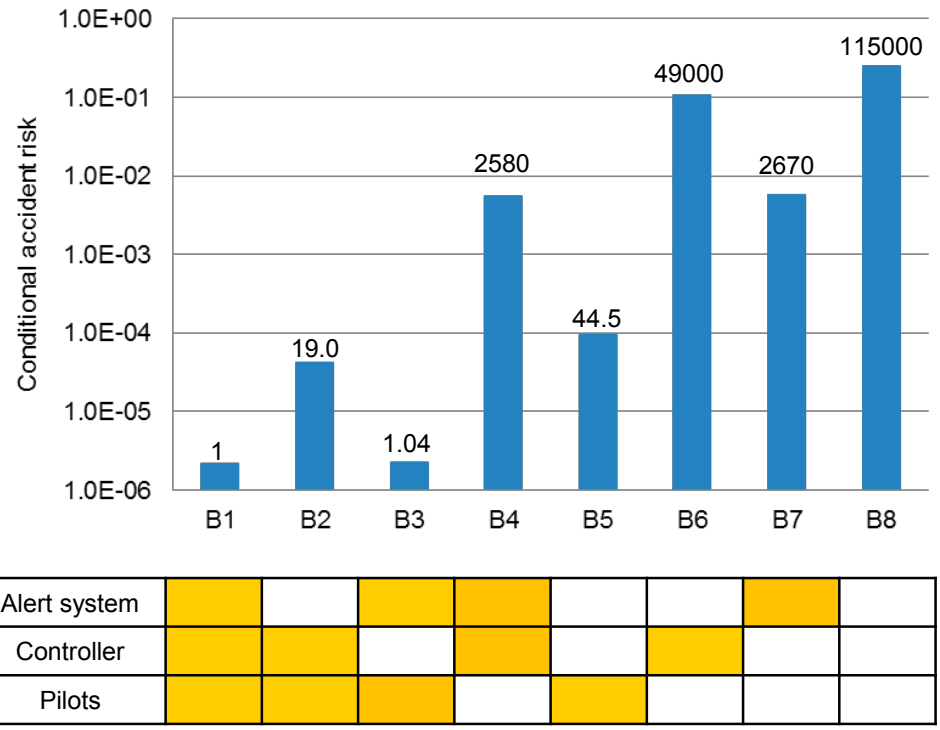

Figure 3: Conditional accident risk results of the event tree in Figure 2 for various cases where the alert system, the controller and the pilots can (filled box) or cannot (open box) independently detect a conflict. The value on top of each bar is the risk increase factor with respect to case B1.

The results given in Figure 3 show that according to the event sequence based safety assessment, the accident risk is reduced very strongly (by a factor 115,000, see case B8) by the combined contributions of the pilots, controller and alert system. The results given in Figure 3 also show that the pilots have by far the largest contribution in reduction of the accident risk (by a factor 2580, case B4). The alert system leads to a considerable risk reduction by a factor 19.0 (case B2), or even by a factor 44.5 in the case that the controller would not detect conflicts him/herself (case B5). The controller makes relatively small contributions to reduction of the accident risk: by a factor 1.04 if the controller is supported by the alert system (case B3) and by a factor 2.35 if the controller is not supported by the alert system (as follows from a comparison of the accident risk of cases B5 and B2). These small risk reductions by the controller are in line with the small probability values of events Q_3 , Q_7 ,and Q_{11} (see Table 1), which reflect the assumption made in the event sequence based study, that as the pilot of the taxiing aircraft starts crossing without contacting the runway controller, the runway controller is not very likely to timely observe the conflict by own visual monitoring.

4 MULTI-AGENT DYNAMIC RISK MODEL BASED APPROACH

The MA-DRM based assessment approach has been developed as the quantitative part of the Traffic Organization and Perturbation AnalyZer (TOPAZ) methodology for the analysis of accident risk in ATM [13, 18, 19]. The MA-DRM based assessment approach makes use of an agent based model (ABM) that is embedded in a stochastic analysis framework. The ABM approach is described in Section 2.1. The stochastic analysis framework is described in Section 2.2. Next Section 2.3 describes a Petri net approach toward specifying an ABM in the stochastic analysis framework. Section 2.4 explains how the stochastic analysis framework is exploited to conduct rare event Monte Carlo simulations with the Petri net based model. Finally Section 2.5 describes how differences between the Petri net based model and reality are taken into account in the MA-DRM based risk assessment.

4.1 AGENT BASED MODELLING AND SIMULATION

The sociotechnical system views of [1-4] match very well with the complexity science view that in systems of multiple agents, behaviour may emerge from the interactions between these agents. Hence, ABM approaches have been used for the analysis of a variety of sociotechnical systems [20], including evacuation induced traffic flows, stock markets, organizational design, and diffusion and adoption of innovation. The benefits of ABM over other modelling techniques are captured by Bonabeau [20] in three statements: (i) ABM captures emergent phenomena; (ii) ABM provides a natural description of a system; and (iii) ABM is flexible. In [21] it is further argued that an ABM approach is well suited for applications involving interactions between functionally or geographically distributed autonomous subsystems. This makes ABM simulation a logical choice for the evaluation of future advanced ATC designs. For example, Shah et al. [22] showed that ABM simulation offers the capability to integrate cognitive and technological models that interact in an ATC environment. Simulation of such interacting models can predict the results of transformations in procedures and technology and such emergent behaviour typically cannot be found by examining the behaviour of the individual agents alone.

In the development of an MA-DRM approach within the TOPAZ methodology, an ABM has explicitly been embraced in [12] in order to extend the human directed

situation awareness (SA) model of Endsley [23] to a multi-agent SA propagation model, which covers both human and technical agents. The motivation for developing this extension was twofold: 1) Endsley [24] showed that more than 60% of the causal factors underlying aircraft accidents involving major air carriers in USA involved problems with proper SA; and 2) TOPAZ experience showed that many hazards in multi-agent ATM operations stem from SA inconsistencies between agents. The multi-agent SA model of [12] makes explicit that in a multi-agent system, SA propagates from one agent to another agent, during which errors may sneak in the SA's of the agents without being noticed by any of the agents.

4.2 STOCHASTIC HYBRID AUTOMATA

In an ABM simulation, a collision between a pair of aircraft occurs when the joint state of the simulated aircraft hits a critical subset of their joint state space. In systems theory, the estimation of the probability of reaching a given subset of the state space within a given time period is known as a problem of probabilistic reachability analysis, e.g. [25]. Because of the huge dimensionality of a multi-agent model of a complex sociotechnical system, existing probabilistic reachability approaches, e.g. [26], fall short for determining the accident risk.

In safety-critical industries (e.g., nuclear, chemical), reachability analysis is addressed by methods that are known as dynamical approaches towards probabilistic risk analysis (PRA), e.g. [27]. These dynamical PRA methods represent the dynamic evolution between discrete events by ordinary differential equations. In stochastic control theory these are known as piecewise deterministic Markov process [28, 29]. For safety modelling of air traffic operations, it may as well be needed to incorporate Brownian motion in the piecewise deterministic Markov process model, e.g. to represent the effect of random wind disturbances on aircraft trajectories [30].

The class of systems which incorporates Brownian motion within piecewise deterministic Markov processes, has been defined as a stochastic hybrid automaton [31]. Such automaton has a hybrid state consisting of two components: a continuous valued state component and a discrete valued state component. The continuous state evolves according to a stochastic differential equation (SDE), where the vector field and drift factor depend on both hybrid state components. Switching from one discrete state to another discrete state is governed by a probability law or occurs when the continuous state hits a pre-specified boundary. Whenever a switching occurs, the hybrid state is reset instantly to a new state according to a probability measure which depends itself

on the past hybrid state. Complementary dynamic and stochastic effects are induced by the interaction between the hybrid state components. A key quality of this type of stochastic hybrid automaton is that it generates a process, which is named generalised stochastic hybrid process (GSHP), and for which it has been proven that it satisfies the strong Markov property [32, 33].

4.3 PETRI NET BASED SPECIFICATION OF A GSHP

For the modelling of accident risk of safety-critical operations in nuclear and chemical industries, the most advanced approaches use Petri nets as model specification formalism, and stochastic analysis and Monte Carlo simulation to evaluate the specified model [27]. Since their introduction as a systematic way to specify large discrete event systems, Petri nets have shown their usefulness for many practical applications in different industries, e.g. [34]. Various types of Petri net modelling have also found their way into reliability and safety applications, e.g. [35-38].

Although Petri nets have much in common with automata, there also are significant differences. Cassandras and Lafortune [39] explain that both Petri nets and automata have their specific advantages. Petri net is more powerful in the development of a model of a complex system, whereas automata are more powerful in supporting analysis. In order to combine the advantages offered by both approaches, there is need for a systematic way of transforming a Petri net model into an automata model. Such a transformation would allow using Petri nets for the specification and automata for the analysis. For a timed or stochastic Petri net with a bounded number of tokens and deterministic or Poisson process firing, such a transformation exists [39].

A Petri net consists of places (drawn as circles) and transitions (drawn as squares), which are connected by arcs (drawn as directed arrows). The places represent discrete states; a token (drawn as a black dot) in a place represents that discrete state to be currently active. The transitions can remove tokens from places and produce tokens for places in the direction of the arcs, representing jumps between discrete states. In order to make this basic Petri net formalism useful for modelling of air traffic operations, we need various extensions, including a one-to-one transformation to the stochastic hybrid automaton setting of GSHP. Jensen [40] introduced the extension of attaching a colour to each token in a Petri net, where the colour assumes values from a finite set. Tokens and the attached colours determine which transitions are enabled. Upon firing by a transition, new tokens and attached colours are produced as a function of the removed tokens and colours. Haas [41] extended this colour idea to

(stochastically) timed Petri nets where the time period between enabling and firing depends of the input tokens and their attached colours. Both in [40, 41], a colour does not change as long as the token to which it is attached remains at its place. Everdij and Blom [42, 43] defined a dynamically coloured Petri net (DCPN) by incorporating two additional extensions: (1) a colour assumes values from a Euclidean state space, its value evolves as solution of a differential equation and influences the time period between enabling and firing; (2) the new tokens and attached colours are produced as random functions of the removed tokens and colours. Subsequently, the DCPN has been further extended to a stochastically and dynamically coloured Petri net (SDCPN) by allowing a colour to evolve as a solution of a stochastic differential equation [44]. Also, it has been proven that an SDCPN-generated process (e.g. through Monte Carlo simulation) is mathematically equivalent to a GSHP [44]. Therefore SDCPN generated processes inherit the stochastic analysis power of GSHP as well as of stochastic hybrid automata [45]. This inheritance distinguishes SDCPN from various other hybrid Petri net modelling extensions, e.g. [34]. Finally, complementary SDCPN features have been developed [46] that allow a hierarchical and compositional approach in specifying a multi-agent model as an SDCPN.

4.4 Rare event Monte Carlo simulation of SDCPN

Based on the SDCPN specification of the ABM, Monte Carlo simulation software is developed. Air traffic is a very safe means of transport and the probability of a 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. Therefore, a speed-up method for Monte Carlo simulation of the SDCPN is required. For collision risk assessment in ATM, such speed-up has been achieved by risk decomposition and by an interacting particle system (IPS) approach, both of which are concisely explained next.

Risk decomposition consists of decomposing accident risk simulations in a sequence of conditional Monte Carlo simulations and combining the results of these conditional simulations into the assessed collision risk value. The strong Markov property of an SDCPN generated process allows to properly estimate the conditional risk given a specific event sequence, (including dependent events) and the conditional probabilities of such event sequences [47].

The IPS approach supports probability estimation of collision probability in ATM scenarios by introducing a sequence of intermediate aircraft encounter conditions that are always preceding a collision. The collision probability is

determined as the product of conditional probabilities of reaching these intermediate encounter conditions. The conditional probabilities are estimated by simulating in parallel several copies of the process, i.e. each copy is considered as a particle following the trajectory generated by the process dynamics [48]. Cerou et al. [49] have proven that under certain conditions this IPS approach yields unbiased risk probabilities, which distinguishes IPS from the popular Restart method [50]. The main condition that is required to ensure unbiased estimation, is that the simulated process must have the strong Markov property, which property holds true for the SDCPN generated GSHP [32, 33].

4.5 Evaluating differences between the SDCPN based model and the real operation

By the very nature of any model, there are differences between a real operation and a model of the operation. This means that the effects of these differences remain to be taken into account in the risk assessment. In the MA-DRM based risk assessment this is pursued by a systematic assessment of the bias and uncertainty in the risk that is expected to be inferred by potential differences between SDCPN based model and reality (Figure 4).

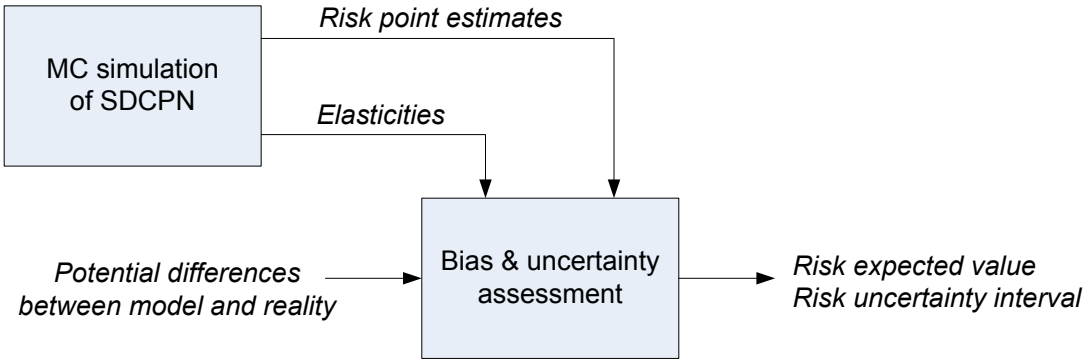

Figure 4: *In the MA-DRM based risk assessment, MC simulation of the SDCPN plays a dual role by providing risk point estimates and elasticities of the risk with respect to model parameter values.*

Figure 4 shows that rare event Monte Carlo simulation of the SDCPN is used for two purposes: 1) to assess the model based point estimate of the collision risk; and 2) to assess the model elasticities (log-sensitivities) from input to output. The assessed elasticity values are used to evaluate the impact on the assessed risk level of the differences between model and reality. The specific steps in the bias and uncertainty assessment are [51, 52]:

1. *Identify potential differences between model and reality.* This concerns differences between: i) the values assumed in the SDCPN simulation and the

real parameter values; ii) differences between SDCPN structure assumed and structure in reality; iii) differences due to hazards that are not modelled by the SDCPN; and iv) differences between the operational concept assumed in the SDCPN and the real operational concept.

2. *Assess the size/probability of the differences.* For each parameter value a bias factor and a corresponding uncertainty interval and a bias factor are assessed. For other types of differences a value is assessed for the probability that the difference applies in the case considered.
3. *Assess the elasticity (log-sensitivity) of assessed risk level for changes in parameter values*. Additional Monte Carlo simulations are conducted with the SDCPN in order to assess the elasticities (log-sensitivities) of the SDCPN assessed accident risk to changes in its parameter values.
4. *Assess the effect of each potential parameter value difference on the risk outcome.* The bias and uncertainty interval of each parameter value are combined with the risk elasticities. In order to find the bias and uncertainty interval in the risk for the parameter value considered.
5. *Assess effect of the non-parameter differences.* For the non-parameter types of differences, a conditional risk bias given the difference exists is assessed and this is combined with the probability that the difference exists.
6. *Determine the joint effect of all differences*. The joint effect of all differences on the bias and uncertainty interval of the risk is determined [51, 52].

5 MA-DRM BASED SAFETY STUDY OF THE RUNWAY INCURSION SCENARIO

This section describes the results obtained by the MA-DRM based safety assessment of the runway incursion scenario. Section 5.1 describes the SDCPN of the runway incursion scenario. Section 5.2 describes the accident risk results achieved using the MA-DRM based approach. Section 5.3 provides a further analysis of agent based events in Monte Carlo simulations of the SDCPN. Section 5.4 describes an analysis of changes in risk results when one or several agents are placed out of the loop or monitoring roles.

5.1 SDCPN MODEL OF THE RUNWAY INCURSION SCENARIO

The main agents in the MA-DRM of the runway incursion scenario are the aircraft taking-off and taxiing, the pilots flying of the aircraft, the runway controller and the ATC system. These agents and the interactions between them are shown in Figure 5. A summary of the agents and illustrative examples of the associated SDCPN models are provided next.

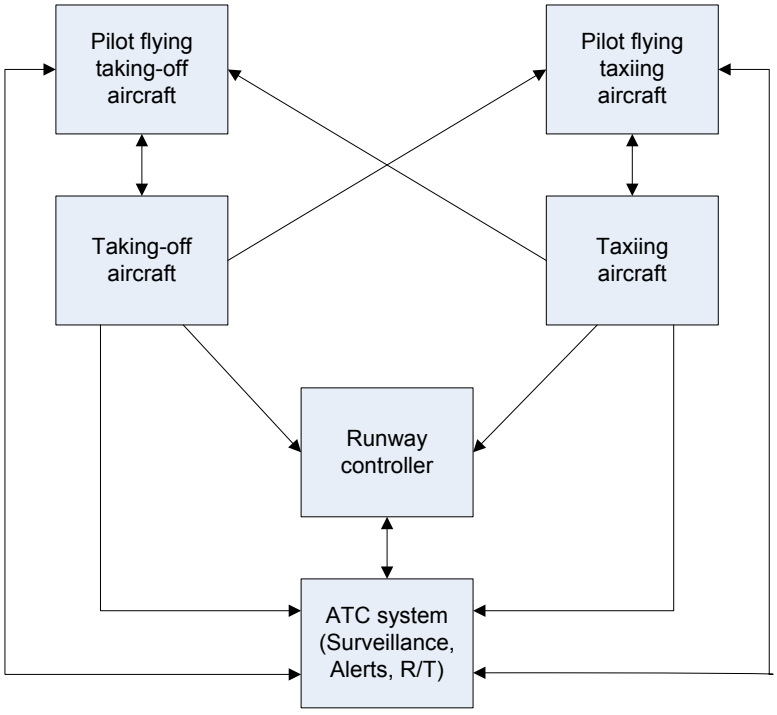

Figure 5: Interactions between the agents of the MA-DRM of the runway incursion: aircraft, pilots flying, runway controller and ATC system.

5.1.1 TAKING-OFF AIRCRAFT (AC-TO)

The model of the taking-off aircraft represents the ground run, airborne transition and airborne climb-out phases during takeoff and includes the possibility of a rejected takeoff. The aircraft initiates takeoff from a position near the runway threshold and it may be medium-weight or heavy-weight. Figure 6 shows a part of the Petri Net for the evolution of the taking-off aircraft, which illustrates that the modes *Ground Run* may be followed by the modes *Rejected Takeoff* and *Hold*, or by the modes *Airborne Transition* and *Airborne Climb-out*, dependent on actions of the agent Pilot Flying Taking-off Aircraft. Examples of selected differential equations associated with these modes are:

$$\begin{aligned}
\dot{x}_t^{\text{AC-TO}} &= v_t^{\text{AC-TO}} \cos\psi_t^{\text{AC-TO}} \cos\gamma_t^{\text{AC-TO}} \\
\dot{z}_t^{\text{AC-TO}} &= v_t^{\text{AC-TO}} \sin\gamma_t^{\text{AC-TO}} \\
\dot{v}_t^{\text{AC-TO}} &= a_t^{\text{AC-TO}} \\
\dot{\gamma}_t^{\text{AC-TO}} &= \begin{cases} 0 & \text{if } \textit{Ground Run} \\ v_t^{\text{AC-TO}} / R_t^{\text{AC-TO}} & \text{if } \textit{Airborne Transition} \end{cases}
\end{aligned} \qquad (2)$$

where $x_t^{\text{AC-TO}}$ is the position along the runway of the aircraft, $z_t^{\text{AC-TO}}$ is the vertical position of the aircraft, $v_t^{\text{AC-TO}}$ is the aircraft speed, $\gamma_t^{\text{AC-TO}}$ is the flight-path angle, $\psi_t^{\text{AC-TO}}$ is the heading, $a_t^{\text{AC-TO}}$ is the acceleration and $R_t^{\text{AC-TO}}$ is the flight-path radius during airborne transition. In the model of the taking-off aircraft it has thus been assumed that the aircraft accelerates along the runway and moves along a circular flight-path during the airborne transition phase.

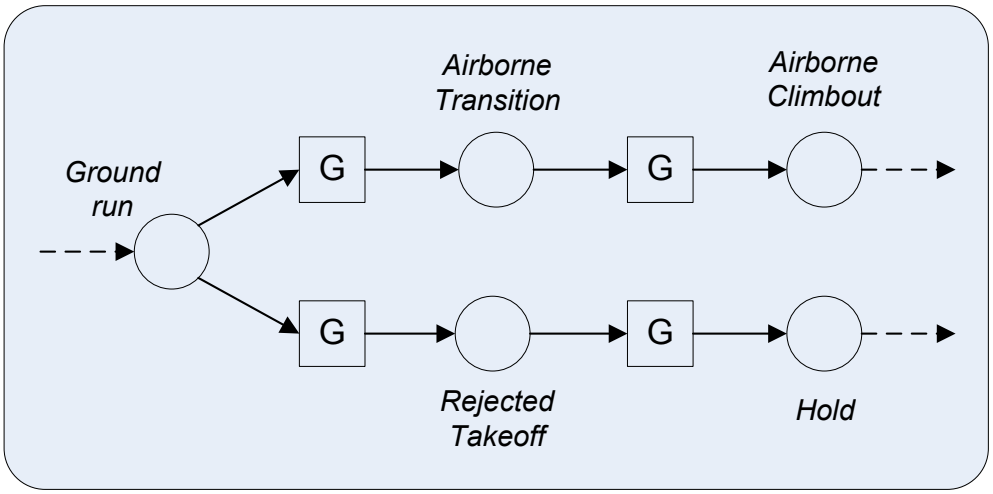

Figure 6: Part of the Petri Net for the model of the trajectory of the taking-off aircraft.

5.1.2 TAXIING AIRCRAFT (AC-TX)

The model of the taxiing aircraft represents aircraft movements during taxiing, including braking as a means to avoid a collision. The aircraft enters the taxiway leading to the runway crossing at a position close to the remotely controlled stopbar (see Figure 1), where its entrance time is uniformly distributed around the take-off time of AC-TO. The aircraft may be medium-weight or heavy-weight.

5.1.3 ATC SYSTEM

The model of the ATC system includes components for the surveillance system, the alert system and the R/T system.

- The model of the surveillance system provides position and velocity estimates for both aircraft. There is a chance that the surveillance system is not available, resulting in track loss. Surveillance data is used by the ATC alert system.
- A stopbar violation alert (SVA) becomes active if the surveillance data indicate that AC-TX has passed an active stopbar. A runway incursion alert (RIA) becomes active if the surveillance data indicate that AC-TX is within a critical distance of the runway centre-line and AC-TO has exceeded a velocity threshold in front of the runway crossing. There is a chance that the alerts are not well functioning.
- The model for the R/T 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 and the frequency selection of the aircraft communication system. The nominal status of these communication systems accounts for direct non-delaying communication. The model accounts for the chance of delay or failure of the communication systems.

5.1.4 PILOT FLYING OF TAXIING AIRCRAFT (PF-TX)

The model for the performance of PF-TX accounts for performance of tasks such as auditory monitoring, visual monitoring, crew coordination, aircraft control, and conflict detection and reaction. The model includes dynamic representations of situation awareness about AC-TO, AC-TX and controller calls, a cognitive control mode of the pilot and task scheduling by the pilot. In the conflict scenario considered, PF-TX intends to continue taxiing on a regular taxiway (whereas actually the aircraft is on the runway crossing). During taxiing PF-TX visually monitors the traffic situation. In particular, at stochastically distributed times $t_{k,\text{PF-TX}}^{\text{visual}}$:

$$t^{\text{visual}}_{k,\text{PF-TX}} = t^{\text{visual}}_{k-1,\text{PF-TX}} + \tau^{\text{interval-vis}}_{k,\text{PF-TX}} + \tau^{\text{duration-vis}}_{k,\text{PF-TX}} \qquad (3)$$

where $\tau^{\text{interval-vis}}_{k,\text{PF-TX}}$ is the interval to the previous visual monitoring action, which is chosen from an exponential probability distribution, and $\tau^{\text{duration-vis}}_{k,\text{PF-TX}}$ is the duration of the visual monitoring action, which is chosen from a uniform probability distribution, the agent PF-TX updates situation awareness components:

$$\begin{aligned}
\hat{x}^{\text{AC-TO}}_{t,\text{PF-TX}} &= x^{\text{AC-TO}}_{t} + \varepsilon^{\text{AC-TO},\hat{x}}_{t,\text{PF-TX}} \\
\hat{v}^{\text{AC-TO}}_{t,\text{PF-TX}} &= v^{\text{AC-TO}}_{t} + \varepsilon^{\text{AC-TO},\hat{v}}_{t,\text{PF-TX}} \\
\hat{y}^{\text{AC-TX}}_{t,\text{PF-TX}} &= y^{\text{AC-TX}}_{t}
\end{aligned} \qquad (4)$$

where $\hat{x}^{\text{AC-TO}}_{t,\text{PF-TX}}$ is the situation awareness about the position of AC-TO, $\hat{v}^{\text{AC-TO}}_{t,\text{PF-TX}}$ is the situation awareness about the speed of AC-TO, $\hat{y}^{\text{AC-TX}}_{t,\text{PF-TX}}$ is the situation awareness about the position of the own aircraft and $\varepsilon^{\text{AC-TO},\hat{x}}_{t,\text{PF-TX}}$ and $\varepsilon^{\text{AC-TO},\hat{v}}_{t,\text{PF-TX}}$ are noise contributions in the position and speed estimation processes, respectively. Based upon the situation awareness, PF-TX detects a conflict if AC-TX is within a minimum distance of the runway, AC-TO approaches towards AC-TX and the speed of AC-TO exceeds a threshold value, or due to an R/T call of ATCo-R:

$$\hat{\theta}^{\text{conflict}}_{t,\text{PF-TX}} = \begin{cases} Conflict & \text{if } (\hat{\theta}^{\rho,\text{ATCo-R}}_{t,\text{PF-TX}} = Hold) \vee \{(\left|\hat{y}^{\text{AC-TX}}_{t,\text{PF-TX}}\right| < d^{\text{confl}}_{\text{PF-TX}}) \\ & \quad \wedge (\hat{v}^{\text{AC-TO}}_{t,\text{PF-TX}} > v^{\text{det-TO}}_{\text{PF-TX}}) \wedge (\hat{x}^{\text{AC-TO}}_{t,\text{PF-TX}} < \hat{x}^{\text{AC-TX}}_{t,\text{PF-TX}})\} \\ No\,Conflict & \text{else} \end{cases} \qquad (5)$$

where $\hat{\theta}^{\rho,\text{ATCo-R}}_{t,\text{PF-TX}}$ is the situation awareness of a controller call, $d^{\text{confl}}_{\text{PF-TX}}$ is a minimum distance to the runway and $v^{\text{det-TO}}_{\text{PF-TX}}$ is a minimum speed of AC-TO above which it is recognized as taking-off. Following conflict detection, PF-TX starts a full braking action unless AC-TX already is within a critical distance of the runway centre-line; otherwise it continues and may pass the runway in front of AC-TO.

5.1.5 Pilot flying of the taking-off aircraft (PF-TO)

The model structure of PF-TO is similar to that of PF-TX. Initially, PF-TO is aware that take-off is allowed and initiates a take-off. During the take-off, PF-TO visually monitors the traffic situation on the runway at stochastically distributed times. PF-TO may detect a conflict if AC-TX is observed to be within a critical distance of the runway or due to an R/T call by the runway controller (ATCo-R). Following conflict detection, PF-TO starts a collision avoiding braking action if it is expected that braking will stop AC-TO in front of AC-TX; otherwise it continues and may fly over AC-TX.

5.1.6 RUNWAY CONTROLLER (ATCO-R)

The model for the performance of ATCo-R accounts for the performance of tasks such as visual monitoring, communication with aircraft crews, ATC coordination, and conflict detection and reaction. The model includes dynamic representations of the situation awareness about the aircraft and the alerts, a cognitive control mode and task scheduling. ATCo-R visually monitors the traffic situation on the runway and is supported by ATC alerts. ATCo-R may detect a safety-critical situation if AC-TX is observed to have passed the stopbar, or due to a stopbar violation alert, or due to a runway incursion alert. Following detection of the safety-critical situation, ATCo-R instructs both AC-TX and AC-TO to hold.

5.2 ACCIDENT RISK ASSESSMENT RESULTS FOR THE RUNWAY INCURSION SCENARIO

The Monte Carlo simulation software that was developed for the SDCPN model described above uses risk decomposition as MC simulation speed-up method. The particular conditions taken into account for this risk decomposition are [13]:

- Type of each aircraft (medium-weight or heavy-weight).
- Remotely controlled stopbar (functioning or not).
- Communication systems (functioning or not).
- ATC alert system (functioning or not).
- Situation awareness of the PF-TX concerning allowance of runway crossing (allowed/not allowed).
- Situation awareness of PF-TX concerning the next waypoint (taxiway/crossing).

The risk assessment takes into account the risk contributions of combinations of these conditions and it includes an evaluation of conditional accident probabilities given each condition.

For the comparison with the risk results of the event tree approach, we focus on the condition that the pilot flying of the taxiing aircraft intends to proceed on a normal taxiway (i.e. without being aware to be heading to the runway crossing). In this situation the pilot of the taxiing aircraft starts to cross the runway without contacting the runway controller, which is the condition considered in the event sequence based risk assessment. For this condition a total of $3.9 \cdot 10^{7}$ Monte Carlo simulations runs were done and a total of 7000 collisions were observed. The point estimate of the accident probability given the condition considered is $1.8 \cdot 10^{-4}$. Given the large number of collisions observed in the Monte Carlo simulations, the statistical error in this point estimate is negligible.

In the bias and uncertainty assessment a total of 306 potential differences between model and reality were assessed, consisting of 175 parameter values and 131 other types. In support of this assessment, interviews with pilots and controllers were conducted about operational aspects, such as the way and timing of recognition of conflicts and subsequent actions [13]. The point estimate and 95% uncertainty interval of the conditional accident probability given the runway incursion scenario are shown in Figure 7. It also shows the conditional accident probability results achieved by the event tree based study. Full comparison of both approaches is done in Section 6.

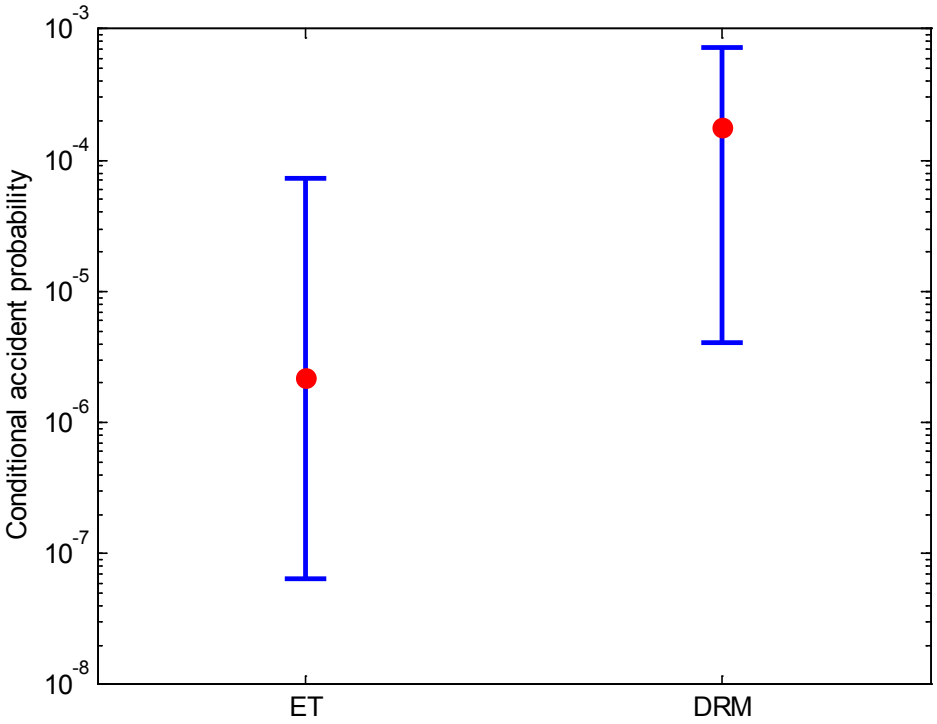

Figure 7: Conditional accident probability results of the event tree based study (lower/upper bound and geometric mean) and of the MA-DRM based study (95% uncertainty interval and point estimate).

5.3 Analysis of agent based events in the runway incursion scenario

In the previous section we showed results for the accident risk and related sensitivity and uncertainty results. To improve the insight in the interactions between the agents in the MA-DRM, the relation of this performance with the accident risk and to support the comparison with the event sequence based analysis, we analyse event occurrences in the Monte Carlo simulations of the MA-DRM. Descriptions of the defined events E_q are provided in Table 3. These

events consider conflict detection by the ATC alert system; conflict detection by the runway controller, either by own observation or via an alert of the ATC alert system; conflict detection by the pilots flying of each aircraft, either by own observation or via a call by the runway controller; braking actions by the pilots flying of each aircraft, and starts and ends of aircraft movements.

Table 3: Description of events tracked in the Monte Carlo simulations of the MA-DRM.

Event	
ID	**Description**
E_1	PF-TO detects conflict
E_1'	PF-TO detects conflict by own observation
E_2	PF-TO initiates rejected take-off (RTO)
E_3	PF-TX detects conflict
E_3'	PF-TX detects conflict by own observation
E_4	PF-TX Initiates braking
E_5	ATCo-R detects conflict
E_5'	ATCo-R detects conflict by own observation
E_6	ATCo-R warns PF-TO
E_7	ATCo-R warns PF-TX
E_8	Stopbar violation alert (SVA) is active
E_9	Runway incursion alert (RIA) is active
E_{10}	AC-TO starts takeoff run
E_{11}	AC-TO comes to stance
E_{12}	AC-TX starts taxiing
E_{13}	AC-TX comes to stance
E_{coll}	Collision

In a Monte Carlo simulation run, the time τ_q of the first occurrence of event E_q, the time τ_{coll} of a collision event E_{coll} and the positions of both aircraft at the times τ_q and τ_{coll} were recorded (when the events occurred). This provides information on the first times when the agents could become aware of the conflict and the resolution actions they could then implement. A total of 10 million Monte Carlo simulation runs were performed for the condition that the pilot flying of the taxiing aircraft has the intent to proceed on a normal taxiway.

In these runs a total of 1809 collisions were counted, which is consistent with the risk point estimate of $1.8 \cdot 10^{-4}$ for this condition as found earlier in Section 5.2.

Figure 8 shows the events, the relations between the events and the event probabilities resulting from the Monte Carlo simulations of the MA-DRM. For each event, two probability results are shown: the unconditional probability of the event and the conditional probability given the occurrence of a collision. Almost all event probabilities shown in Figure 8 result from the interactions in the MA-DRM and could not have been predicted a priori. Only the results for events E_{10} and E_{12} were known before the MC simulations, as the modelled scenario considers the conflict between an aircraft taking-off with an aircraft taxiing, and the conditional probability of E_{coll} given a collision equals one by definition. Key observations and explanations of the results given in Figure 8 are provided next for each of the agents.

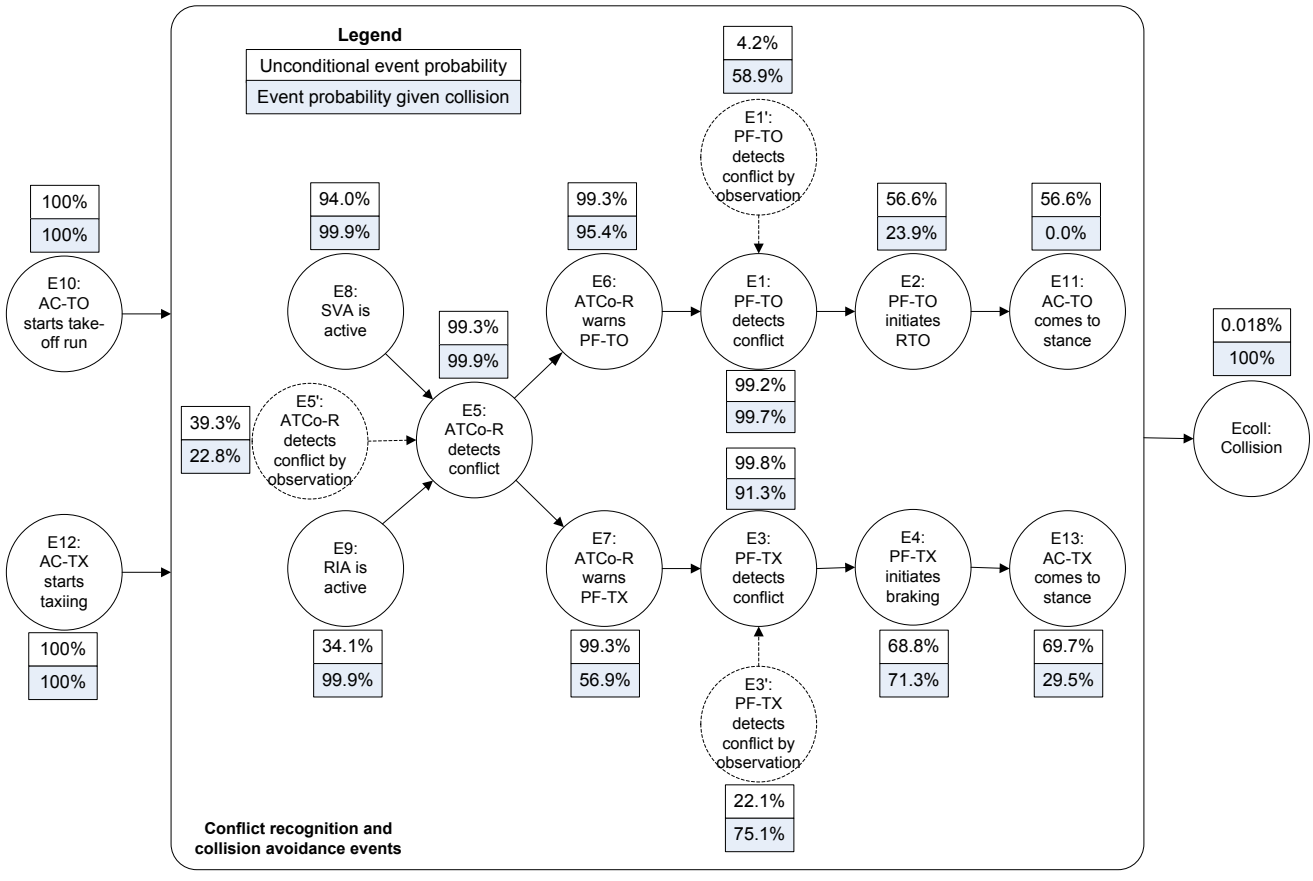

Figure 8: Relations between and probability results of events in the MC simulations of the MA-DRM. For each event, the unconditional event probability and the conditional event probability given the occurrence of a collision are shown.

5.3.1 ATC ALERTS

The stopbar violation alert is active (event E_8) in 94.0% of all scenarios and in 99.9% of the cases ending in a collision. Mostly, it is not activated in situations that AC-TX stops close after the stopbar, such that the alert threshold has not yet been passed.

The runway incursion alert is active (event E_9) in 34.1% of all scenarios and in 99.9% of the cases ending in a collision. It is not activated in situations where AC-TX taxies in front of AC-TO while it has not initiated take-off, or when AC-TX taxies after AC-TO has passed the crossing position.

5.3.2 RUNWAY CONTROLLER (ATCO-R)

ATCo-R detects the conflict (event E_5) in 99.3% of all simulated conflict scenarios. Here, ATCo-R detects the conflict by own observation (event E_5') in 39.3% of all cases, whereas in the remaining 60.0% ATCo-R detects the conflict via the ATC alerting systems.

In the simulation runs ending in a collision, ATCo-R detects the conflict (event E_5) in 99.9% of these cases. Here, ATCo-R detects the conflict by own observation (event E_5') in 22.8% of these cases and via the ATC alert system in 77.1% of these cases. Thus for the conditional case given a collision it is found in hindsight that the probability of conflict detection by ATCo-R is considerably larger than in the unconditional case and the contribution of the ATC alert system to detection of the conflict by ATCo-R is somewhat higher than in the unconditional case.

The controller warns the pilots of the aircraft (events E_6, E_7) in 99.3% of all simulated conflict scenarios, which is equal to the detection rate by the controller (event E_5). In the runs ending in a collision, the probability of a warning is decreased to 95.4% for PF-TO and to 56.9% for PF-TX. A factor contributing to the larger decrease for PF-TX is that in this conflict scenario PF-TX is not on the R/T frequency of ATCo-R and their communication is thus delayed.

5.3.3 PILOT FLYING OF TAKING-OFF AIRCRAFT (PF-TO)

PF-TO detects the conflict (event E_1) in 99.2% of all simulated conflict scenarios. Here, PF-TO detects the conflict by own observation (event E_1') in only 4.2% of all cases, whereas in the remaining 95.0% of all cases PF-TO detects the conflict via ATCo-R. Although PF-TO is very frequently monitoring the traffic situation and

ATCo-R needs time to recognize the conflict and to warn PF-TO, the PF recognizes AC-TX as conflicting only if it is within a critical distance of 90 m to the runway centreline and ATCo-R can recognize AC-TX as conflicting as soon as it has passed the stopbar

Of the simulation runs ending in a collision, in hindsight we observe that PF-TO detects the conflict (event E_1) in 99.7% of these cases, PF-TO detects the conflict by own observation (event E_1') in 58.9% and via the controller in 40.8%. Thus for the conditional case given a collision it is found in hindsight that the probability of conflict detection by PF-TO is higher than in the unconditional case and the contribution of ATCo-R to detection of the conflict by PF-TO is significantly lower than in the unconditional case.

5.3.4 PILOT FLYING OF TAXIING AIRCRAFT (PF-TX)

PF-TX detects the conflict (event E_3) in 99.8% of all simulated conflict scenarios. Here, PF-TX detects the conflict by own observation (event E_3') in 22.1% of the cases, whereas in the remaining 77.7% of all cases PF-TX detects the conflict via ATCo-R. Although ATCo-R needs time to recognize the conflict and to warn PF-TX, the PF detects the conflict situation if it is recognized that AC-TO is taking off, whereas ATCo-R can already recognize the conflict as soon as the taxiing aircraft has passed the stopbar.

Of the simulation runs ending in a collision, in hindsight we can see that PF-TX detects the conflict (event E_3) in 91.3% of these cases, PF-TX detects the conflict self (event E_3') in 75.1% and via the controller in 16.2%. Thus for the conditional case given a collision it is found in hindsight that the probability of conflict detection by PF-TX is considerably lower than in the unconditional case and the contribution of ATCo-R to detection of the conflict by PF-TX is also significantly lower than in the unconditional case.

5.3.5 TAKING-OFF AIRCRAFT (AC-TO)

PF-TO initiates a rejected take-off (RTO) (event E_2) in 56.6% of all cases and also in 56.6% of all cases AC-TO comes to stance (event E_{11}). For the cases ending in a collision, an RTO was initiated in 23.9% of the cases and the aircraft came to stance in 0.0% of the cases.

5.3.6 TAXIING AIRCRAFT (AC-TX)

PF-TX initiates braking (event E_4) in 68.8% of all cases and in 68.7% of all cases AC-TX comes to stance (event E_{13}). In the cases that ended in a collision, braking was initiated in 71.3% of these cases and the aircraft came to stance in 29.5% of these cases.

As indicated above, aircraft positions were recorded in the Monte Carlo simulations at the event times τ_q. Based on this recorded data, Figure 9 and Figure 10 show empirical probability density functions (PDFs) of aircraft positions conditional on the occurrence of an event, or conditional on the occurrence of an event and a collision in the simulation run. Figure 9(a,b) shows that AC-TO is predominantly within the first 500 m of the runway when the conflict is detected by PF-TO or ATCo-R. In contrast, in the cases resulting in a collision (Figure 9(c,d)) these events mostly occur when AC-TO is between 500 m and 1000 m. There is thus a considerable difference in the AC-TO position at the event occurrences for the conditional case given a collision versus the unconditional case. Figure 10(a) shows that AC-TX is often more than 100 m and almost always more than 50 m from the runway centre-line when PF-TX detects the conflict by own observation. In contrast, Figure 10(b) shows that it is quite likely that AC-TX is close or even past the runway centre-line at the time ATCo-R warns PF-TX. Figure 10(c,d) shows that in the cases which resulted in a collision, the taxiing aircraft is within a 100 m at the time of the event. There is an overlap in the cores of the PDFs for the same events in Figure 10, especially for Figure 10(b) and Figure 10(d).

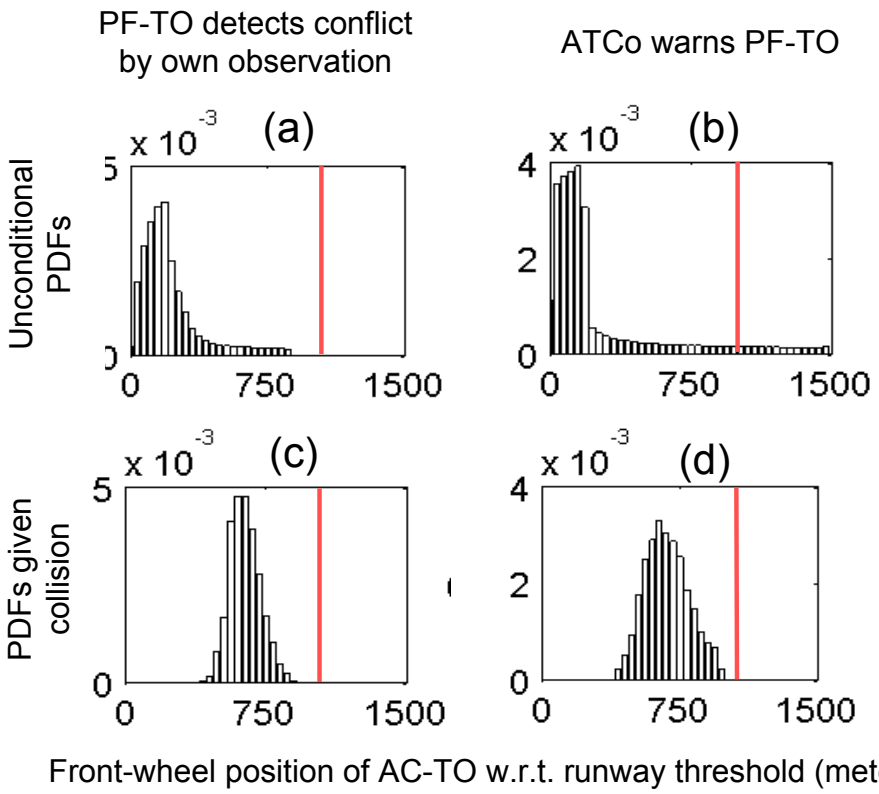

Figure 9: PDFs of the position of the front-wheel of AC-TO along the runway given PF-TO detects the conflict by own observation (event E_1', figures a and c) and given ATCo warns PF-TO(event E_6, figures b and d). The upper figures (a and b) are unconditional PDFs, the lower figures (c and d) are PDFs given a collision occurred. The horizontal axes reflect the position w.r.t. the runway threshold in metres; the red line indicates the taxiway position (at 1000 m).

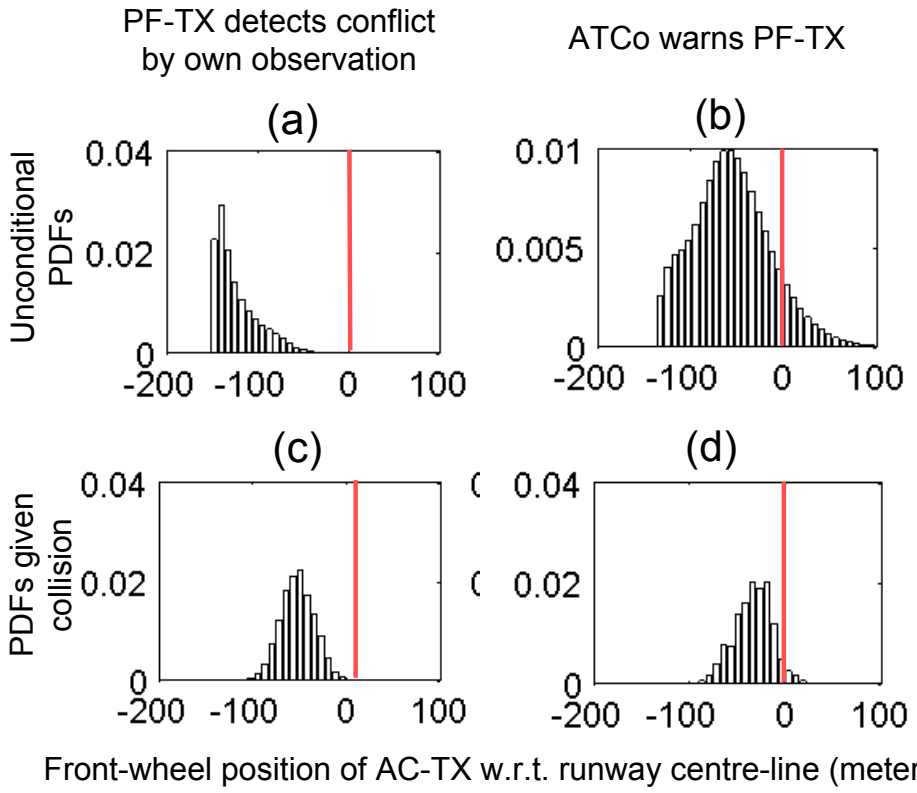

Figure 10: PDFs of the position of the front-wheel of AC-TX along the taxiway given PF-TX detects the conflict by own observation (event E_3', figures a and c) and given ATCo warns PF-TO(event E_7, figures b and d). The upper figures (a and b) are unconditional PDFs, the lower figures (c and d) are PDFs given a collision occurred. The horizontal axes reflect the position w.r.t. the runway centre-line in metres; the red line indicates the runway centre.

5.4 Risk effects due to placing agents out of the loop or monitoring role

The results of the analysis in the last section provided insight in the performance of the various agents in the runway incursion scenario and its relation with accident occurrence. These results were achieved in the setting that all agents perform in the scenario as described by the MA-DRM (Section 5.1). To better understand the potential of agents to restrict the risk increase in cases where the performance of other agents is affected, we performed additional Monte Carlo simulations in which we placed one or more agents out of the loop or monitoring role. This was done for all the agents that are capable of detecting a conflict, namely PF-TO, PF-TX, ATCo-R and ATC System. The conditions for placing these agents out of the loop or monitoring role are:

- PF-TX does not actively monitor the traffic situation visually, such that PF-TX may only detect and react to a conflict via a call of ATCo-R. PF-TX is thus placed out of the monitoring role.
- PF-TO does not actively monitor the traffic situation visually, such that PF-TO may only detect and react to a conflict via a call of ATCo-R. PF-TO is thus placed out of the monitoring role.
- ATCo-R cannot communicate with the pilots. ATCo-R is thus placed out of the loop.
- ATC Alert System does not specify alerts. The ATC Alert System is thus placed out of the loop.

These conditions refer to the situation at the start and during the runway incursion scenario and they are not assumed to hold prior to the occurrence of the runway incursion scenario. Note that for conditions where ATCo-R is out of the loop, it does not matter whether or not the ATC alerts are in the loop, as these can only be effective via ATCo-R.

For all relevant combinations of agents in or out of the loop or monitoring role, the conditional collision risk given the runway incursion scenario was determined by Monte Carlo simulation. Figure 11 shows the conditional collision risks for all twelve relevant cases C1 to C12, and it shows risk factors with respect to the lowest risk as obtained for case C1. Next we discuss the key results of Figure 11 and their relation with the results presented in Section 5.3.

Figure 11 shows that the conditional collision risk varies in the range between $1.8 \cdot 10^{-4}$ and $9.4 \cdot 10^{-2}$ depending on the agents that are in the loop or monitor role. In the extreme case that none of the agents would be actively involved in recognizing the conflict and avoiding a collision (case C12), the collision risk of the runway incursion scenario increases by a factor 522. In this case an accident is thus only prevented by chance, especially by the coincidental timing of the runway incursion with respect to the start of the take-off run.

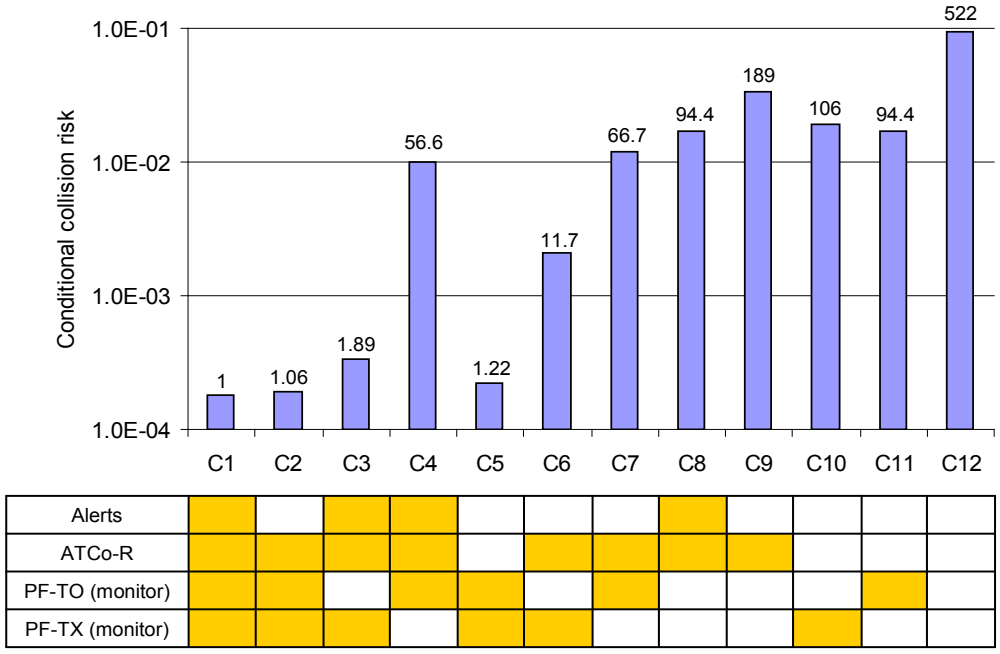

Figure 11: Conditional collision risk results for various conditions with one or more agents in (filled box) or out of (open box) the loop or monitoring role for the runway incursion scenario. Results are shown for the relevant cases, with four to zero agents in the loop / monitoring role. The value on top of each bar is the risk increase factor with respect to case C1. Since ATCo-R being out of the loop implies that alerts are also out of the loop, there are twelve combinations.

The collision risk of the runway incursion scenario increases by a factor 1.06 if the ATC alert systems are not available (case C2). Stated differently, the presence of an ATC alert system barely reduces the collision risk. This is remarkable given the results for events E_5 and E_5' in Figure 8, which show that if the ATC alert system is available, it warns ATCo-R before ATCo-R detected the conflict by own observation in 60% of the cases. Although the ATC alert system thus effectively supports ATCo-R, the results for case C2 show that the agents can well cope without the alerting system. In particular, even though the controller now regularly recognizes the conflict later, the conflict recognition time by the controller and by the pilots is only affected to a limited extent, such that the risk is increased by a factor 1.06 only.

The collision risk of the runway incursion scenario increases by a factor 1.22 if ATCo-R is out of the loop (case C5). Thus the performance of ATCo-R in the resolution of the runway scenario has a small effect only on reducing the collision risk. This result may be seen as quite surprising, given the results for events E_1, E_1', E_3 and E_3' in Figure 8, which show that the controller warns the

pilots flying of the taking-off and taxiing aircraft in 95% and 78% of all cases before they have detected the conflict themselves. Notwithstanding this good performance of the controller, if the controller is placed out of the loop in the modelled scenario, pilots can mostly detect the conflict themselves and react timely to avoid a collision, such that the risk increase is small.

The collision risk of the runway incursion scenario is increased by a factor 1.89 in the (hypothetical) case that PF-TO is not actively monitoring the traffic situation, but might still be warned by ATCo-R (case C3). If in addition to the lack of monitoring by PF-TO, ATCo-R is out of the loop (case C10), then the risk is majorly higher by a factor 56 with respect to case C3. Figure 9 shows that ATCo-R often warns PF-TO at an early stage, namely if AC-TO is well within the first 500 m of the runway. This early stage warning implies that ATCo-R can considerably restrict the risk increase of a non-monitoring PF-TO, as is manifest from the comparison of the risk factors in cases C3 and C10.

The collision risk of the runway incursion scenario is increased majorly by a factor 56.6 in the (hypothetical) case that PF-TX is not actively monitoring the traffic situation, but might still be warned by ATCo-R (case C4). If in addition to the lack of monitoring by PF-TX, ATCo-R is out of the loop (case C11), then the risk increases by a factor 1.7 with respect to case C4. Figure 10 shows that AC-TX is often close to the runway when ATCo-R warns PF-TX. This implies that warnings of ATCo-R to PF-TX are often too late to prevent AC-TX entering a collision-critical area. Therefore, ATCo-R can barely restrict the risk increase due to a non-monitoring PF-TX, as is manifest from the comparison of the risk factors in cases C4 and C11.

The collision risk of the runway incursion scenario is increased majorly by a factor 94.4 in the case that only ATCo-R would be monitoring (while supported by the ATC alert system) and the pilots of both aircraft would not be monitoring, but may be warned by ATCo-R (case C8). The attained risk level is similar to the other cases where only one human operator is actively monitoring the traffic situation (cases C10 and C11). It shows that only one human actively monitoring human cannot effectively restrict the risk increase due to the malperformance of other operators.

Cases C6, C7 and C9 represent situations where the ATC alert system is not available and also one or both of the pilots flying are not actively monitoring the traffic situation. It follows from comparison with the similar cases including the

ATC alert system (i.e. cases C3, C4 and C8, respectively) that the effect of the non-availability of the ATC alert systems varies a lot.

- In the cases without active monitoring by PF-TX (C7 versus C4) the risk increases by a factor 1.2 only, indicating that the alerts are often too late to warn the PF-TX.
- In the cases without active monitoring by PF-TO (C6 versus C3) the risk increases by a factor 6, indicating that in this context the ATC alerts often warn ATCo-R such that ATCo-R can timely warn PF-TO.
- In the cases without monitoring by both pilots (C9 versus C8) a risk increase by a factor 2 is achieved, which is intermediate between the above indicated values.

These results indicate that the potential effectiveness of the ATC alert system can be better than the factor 1.06 found in case C2 if one or both pilots underperform. In the context given it is most important for timely warning of PF-TO.

6 COMPARISON OF THE SAFETY ASSESSMENT STUDIES

In Sections 3 and 5 we presented the methods and key results of safety assessments for a particular runway incursion scenario by event tree and MA-DRM based approaches. The studies are contrasted in this section. This comparison and evaluation is structured according to the following aspects:

- Comparison of the architecture of the models:
  - Model complexity: what are the levels of complexity of the models?
  - Dynamics: how are the dynamics of sociotechnical systems represented?
  - Performance variability: how is the variability in the performance of humans and technical systems in a sociotechnical system dealt with?
  - Interactions/concurrency: how are interactions and concurrency of the performance of entities in a sociotechnical system dealt with?
  - Emergent behaviour: to what extent is emergent behaviour in the sociotechnical system considered?
- Comparison of the use of the models:
  - Expertise and techniques needed: what kinds of expertise and techniques are required?
  - Variety of contextual conditions: how are various contextual conditions dealt with?
  - Transparency: what is the level of transparency of the methods and results of the safety assessments?
  - Uncertainty: does the estimated uncertainty interval incorporate all uncertainties?
- Comparison of the risk results obtained:
  - Differences in findings: are there significant differences in the safety risk assessment results obtained?
  - Comparison against real data: to what extend is it possible to compare the results obtained against real data?
  - Feedback to design: are there significant differences in the type of feedback that can be provided to the design of the active runway crossing operation?

6.1 Comparison of the architecture of the models

6.1.1 Model complexity

The event tree is represented in a single figure (Figure 2) and can be straightforwardly understood by a large audience with only some basic background in risk models. Obviously, the model complexity of this event tree is relatively low.

The architecture of the MA-DRM addresses various modelling levels. At a high level it considers the agents and the interactions between agents (see Figure 5); this level is easily understood. At a low level, the details of the SDCPN specification are described (see Sections 4.3 and 5.1), including the modes, stochastic dynamics and interactions of the agents. This level describes the details of the complexity encountered in the runway incursion scenario. At this detailed level, the model complexity of the MA-DRM is high.

6.1.2 Dynamics

In the event tree based study, the dynamics of the runway incursion scenario are represented by 27 possible sequences of 12 events Q_1 to Q_{12} (Figure 2), where it is assumed that higher-indexed events do not occur before lower-indexed events. The event tree uses three time ranges for recognition and resolution of the conflict: early, medium and late.

In the MA-DRM based study, the dynamics of the runway incursion scenario are represented by the dynamically interacting agents. Following the SDCPN specification (Sections 4.3 and 5.1), the dynamics evolve from the differential equations of token colours and the dynamics of the token transitions between places. Examples of the dynamics are the movements of the aircraft, the updating of the situation awareness of the pilots and controller and the timing of failures of technical systems. In contrast with the event tree there are no fixed sequences for conflict detection and resolution related events. Rather these events develop in the Monte Carlo simulations of the SDCPN.

6.1.3 Performance variability

In the event tree based study, the variability in the performance of mechanisms for the recognition and resolution of the runway incursion scenario is completely defined by the structure of the event tree and the (conditional) probabilities of its events. In particular, the conditional probabilities define the level of effectiveness that a particular event can resolve the conflict or not, conditional on earlier events in the event tree.

In the MA-DRM based study, performance variability concerns the performance of individual agents, the interactions between agents and the overall performance variability of the multi-agent system. The performance of individual agents and agents' interactions are defined in the SDCPN based model, e.g. the timing of task performance by human operators, noise components in visual observation of an aircraft position by human operators, probabilistic errors in communication between pilot and controller, noise in radar surveillance systems, and variation in deceleration profiles of aircraft. The overall performance variability of the multi-agent system arises from the Monte Carlo simulations of the SDCPN.

6.1.4 INTERACTIONS / CONCURRENCY

The event tree based study combines the mechanisms to detect a conflict and avoid a collision at three stages, leading to the 27 different cases of Figure 2**.** In the quantification of the effectiveness of each of these mechanisms their interactions must be accounted for. For instance in Figure 2, the occurrence of event Q_4 is conditional on the occurrence of Q_0 and the non-occurrence of Q_1, Q_2 and Q_3, meaning that the alert system may warn the controller at an early stage, if and only if: i) there is an aircraft crossing while it should not; and ii) there is an aircraft in take-off; and iii) the conflict has not been recognized and resolved at an early stage by the pilots; and iv) the controller has not yet recognized the conflict. These types of nested relations get more complicated as one progresses along the event tree. Even with the limited number of sequences, this makes an appropriate modelling of the dependencies between the events a very difficult task.

The MA-DRM of the runway incursion scenario describes the performance of agents, including nominal and non-nominal performance modes, and the interactions between agents by the SDCPN formalism. As such it represents a broad variety of dynamic and stochastic hybrid processes in a more direct way. Examples of performance modes include particular tasks of pilots and controller or failure modes of technical systems. Examples of interactions between agents are observation of aircraft positions by pilots or communication between controller and pilots. In an SDCPN based model only a limited number of such performance modes, contextual conditions, stochastic dynamics and agents' interactions have to be defined in order to capture an in principle infinite variety of potential event sequences. Through running Monte Carlo simulations with the SDCPN model, random samples are drawn from this large variety of potential event sequences.

6.1.5 Emergent behaviour

There are various views on emergent behaviour, see [53] for a recent overview including a discussion of their applicability to air traffic. Here, we evaluate emergence according to the definitions of Bedau [54], Corning [55] and Chalmers [56]. In the event tree, the probabilities of the end events (e.g. incidents/accidents) are calculated straightforwardly and they are qualities of the same kind as the data used to obtain them: both are event probabilities. The evaluation of risk by these types of models does not require the type of simulation as described by Bedau [54] in his definition of weak emergence, nor can they be considered to be qualitative novelties and synergistic wholes composed of things of unlike kind as in the definition of emergence of Corning [55], nor can they be considered emerging truths that are unexpected given the principles governing the low-level domain as expressed in the definition of weak emergence by Chalmers [56]. In conclusion, the assessed risk level is not an emergent property in the event tree based approach.

In the MA-DRM based study, the assessed risk level emerges from the MC simulations of the MA-DRM. In line with the definition of weak emergence by Bedau [54], risk is a macrostate that can be inferred by simulation of the microdynamics of the MA-DRM. In line with the definition of Corning [55] risk is an emergent property, since it is a qualitative novelty of a completely different nature of the traffic scenario considered and it is obtained by combined effects of various elements (agents). In line with the definition of weak emergence of Chalmers [56], risk is a high-level phenomenon that arises from the low-level domain (i.e. the varying performance of the agents). In conclusion, the performance variability of the agents and interactions between the agents lead to emergent behaviour of the multi-agent system and to the assessed risk level as an emergent property in the MA-DRM based approach.

6.2 Comparison of the use of the models

6.2.1 Expertise and techniques needed

Both approaches have in common that safety analysis requires access to multi-disciplinary knowledge regarding the technical systems, human operators, environmental influences and the interactions between these entities. Also common is the expertise used at the initial phase of safety assessment of the runway incursion scenario. This consists of in-depth learning about the specifics of the ATC case considered, including identification of potential hazards. Part of this is done through collecting expert information from controllers, pilots and

the experts of the various technical systems. However, after this initial learning phase, there are significant differences between the two approaches.

In the event sequence based approach, the collected information is analysed and subsequently synthesized into a manageable number of potential event sequences. For socio-technical systems this analysis/synthesis process is an art rather than a science. This also means that the development of the event sequence tree for a socio-technical system is typically done through several iterations. Once the event tree has been frozen, the follow-up step is to collect data for the estimation of the various conditional event probabilities in the tree. For frequent events, data often is available to estimate these probabilities. For less frequent events, however, typically interviews with controllers and pilots form the main source of data collection.

The analysis and synthesis of the Petri net model and the running of rare event Monte Carlo simulations differ very much the event sequence based approach. The key difference is that there is no longer the need to think about the various combinations of dynamics and events that might happen. A Petri net based modelling approach allows to develop the model in a compositional way, agent by agent. The synthesis of the potential combinations of event sequence is simply left to the Monte Carlo simulator. Obviously, the development of an MA-DRM based safety risk assessment requires knowledge on capturing agent type specific background in a SDCPN model, and in running efficient Monte Carlo simulation, including the use of dedicated speed-up techniques to capture rare events.

6.2.2 Variety of contextual conditions

In the event tree based study, the structure of the event tree and the conditional probabilities of its events are derived for the particular contextual conditions considered, such as good visibility and the use of a runway incursion alert system. A change in such contextual conditions would imply changes in the event tree structure and parameterization. The event tree approach does not provide ways to easily reassess the structure and parameter values to account for varieties of contextual conditions.

In the MA-DRM based study, contextual conditions are often defined through model parameters for various agent entities. For instance, as shown in [13], the visibility condition for human operators is included in the SDCPN based model for the runway incursion case by a visibility distance parameter. As another example, in this study we analysed the effect of excluding a runway incursion

alert system in the operation simply by removing the possibility of these alerts to occur. The risk effects of such changes are the resultant of the agents' interactions and do not need to be pre-specified on the level of event probabilities as in an event tree based study.

6.2.3 TRANSPARENCY

Transparency is an important quality in risk assessment, and includes the following three complementary aspects:

a) Transparency of the development of the risk model architecture,
b) transparency of the quantification of the risk models,
c) transparency of the output generated by the risk models.

These three transparency aspects are compared for the event tree and MA-DRM based approaches for the example considered.

Transparency of the development of the risk model architecture

During the development of the event tree various choices were made regarding the types of events, dependencies between events and the ordering of events. Typically, these choices seem a bit arbitrary, in the sense that other choices might as well have been made.

In the MA-DRM based study, an SDCPN model has been specified, which includes explicit representation of modes, dynamics and interactions of the agents. During this development, assumptions regarding the specific modelling choices made were explicitly formulated. In contrast with the event tree development, no specification of event orderings had to be developed; the various event orderings emerged from the Monte Carlo simulations. Consequently, the SDCPN development process is more structured and transparent than it is for the event tree.

Transparency of the quantification of the risk models

In the event tree study, quantification means adopting values for the event probabilities and this was achieved by expert (controller and pilot) elicitation for the case studied. As argued above under the heading of interactions, the event probabilities are conditional upon events earlier in the event tree and appropriately accounting for these dependencies is ambiguous. Typically no clear argumentation was provided for the conditional event probabilities. Similar difficulties exist in accounting for the effect of contextual conditions on the event probabilities, such as argued above under the heading contextual

conditions. In conclusion, the quantification in the event tree approach is not transparent.

In the MA-DRM approach, quantification means adopting values for a wide range of parameters for the agent models developed. Above, under the heading of performance variability, several examples are provided of agent aspects for which parameter values have been specified. These parameters typically have a physical meaning that is dedicated for each of type of agent modelled. Part of these model parameters (e.g. aircraft performance) could be quantified quite accurately, the quantification of other model parameters (e.g. human related) involved more uncertainty. Because an SDCPN based model is nearer to capturing the physical aspects of a scenario than the event tree model, the quantification of its parameters is more objective and transparent than the quantification of the ambiguous conditional event probabilities in the event tree.

Transparency of the output generated by the risk models
The results of the event tree are the probabilities of incidents and accidents resulting from the runway incursion scenario. Given the event tree architecture and quantification, these results are calculated straightforwardly in a manner that can easily be checked by others.

The results of the MA-DRM based risk assessment are (conditional) accident risk probabilities, the risk uncertainty range, risk sensitivities of the parameters, (conditional) probabilities of agent performance-related events, (conditional) probability density functions of agent performance variables and risk probabilities of dedicated cases with agents being in or out of the loop or monitoring role. These results were achieved by Monte Carlo simulations of the MA-DRM and the associated bias & uncertainty assessment Figure 4. The results are thoroughly documented and can thereby be checked or repeated by others.

6.2.4 Uncertainty

For the event tree, the estimated uncertainty interval accounts for uncertainties in the estimated probabilities used in the numerical evaluation of the event tree. However, uncertainty regarding potentially missing safety critical event sequences and regarding lack of knowledge on dependencies between event probabilities are not taken into account.

In the MA-DRM based approach, sequences of events leading to accidents follow from the rare event Monte Carlo simulation rather than from pre-specification of limited set of sequences. Moreover, in the MA-DRM based approach potential differences between the model and the real operation are taken into account through a systematic assessment of the effects of these differences on the estimated safety risk. For the latter estimation explicit use is made of the assessed sensitivities of the risk to changes in the physical parameters of the model.

6.3 Comparison of the risk results of the models

6.3.1 Differences in findings

Figure 7 shows that the accident risk was assessed to be considerably lower by the event tree based assessment in comparison with the MA-DRM based assessment. In particular the mean risk assessed by the event tree approach is a factor 82 below the risk point estimate of the MA-DRM approach. Moreover, comparison of results of Figure 3 and Figure 11 shows that the differences in risk reduction factors by the agents are even larger. The combined action of pilots, controller and ATC alert system was assessed to reduce the accident risk by a factor 115,000 in the event tree based approach (case B8), whereas the risk reduction factor following the MA-DRM based approach is limited to 522 (case C12). The risk reduction by the ATC alert system was assessed to be a factor 19 by the event tree based approach (case B2) versus only a factor 1.06 by the MA-DRM based approach (case C2). This small factor is striking, since the MA-DRM based results in Figure 8 show that in about 94% of the runway incursion scenarios at least one of the alert types is active and in 60% of the scenarios the alert system warns the controller before (s)he has detected the conflict independently. In spite of this alert activity, the pilots are mostly able to timely detect the conflict themselves and to avoid a collision independent of the ATC alert system. This relationship has not been captured in the event tree based assessment. A similar situation exists for the risk reduction by the alert-supported controller, which was assessed to be a factor 44.5 by the event tree based approach (case B5 in Figure 3), but only a factor 1.22 by the MA-DRM based approach (case C5 in Figure 11). The additional results in Figure 8 show that the controller detects the conflict and warns the pilots in 99.3% of the cases and that in 95% and 78% of the cases the controller is able to warn the pilots flying of the taking-off or taxiing aircraft, respectively, before the pilots have detected the conflict independently. In spite of this laudable performance of the controller in the model, the accident risk would only increase by a factor 1.22 if

the controller would not play a role at all in the resolution of the runway incursion scenario. This limited capability of the controller to avoid the occurrence of an accident is illustrated by the decrease in the probability that the controller is able to warn the pilots prior to detection of the conflict by the pilots themselves from 95% to 41% (taking-off aircraft) and from 78% to 16.2% (taxiing aircraft) for the simulation runs that ended in an accident (Figure 8). These types of results follow from the MA-DRM based approach by considering the totality of the performance and interactions of all agents in huge numbers of simulations. Such relationships have not been captured in the event tree based study.

6.3.2 Comparison against real data

Although runway incursion data are gathered and analysed as part of safety management in ATC [57], the number of resulting collisions is (fortunately) far too low for statistically meaningful comparison at the level of conditional collision risk results for a runway incursion scenario. This applies to both safety assessment approaches. However, for various submodels of the total model, comparison against real data often is feasible. The latter is easier for the MA-DRM approach because this simply asks for conducting Monte Carlo simulations with a part of the complete SDCPN, and subsequently to compare the results obtained with real data obtained for the subsystem considered. Such a localized comparison is not feasible for an event sequence based approach.

6.3.3 Feedback to design

The results of the event tree study are the risk levels and the main events contributing to these risk levels. As has been explained in [14], the assessed risk levels were possibly unacceptable in the event tree based study and hazards contributing to these risk levels include pilots' selection of an incorrect R/T frequency and unavailability of the R/T frequency.

The feedback to the design provided by the MA-DRM based safety assessment is more diverse than that of the event tree analysis. As the SDCPN based model represents the dynamic performance of interacting agents rather than merely the occurrence of events, the feedback to the design reflects this rich variety of performance aspects. For instance, the feedback includes the unexpected ineffectiveness of the ATC alert system, the sensitivity of the risk for system settings (e.g. alert threshold and system availability settings), and the contributions of human operators to reducing the accident risk in the runway incursion scenario.

7 Discussion

In this paper we compared risk assessment studies of a particular runway incursion scenario by an event tree based approach versus a MA-DRM based approach. This comparison was done at a qualitative level as well as for the particular quantitative differences attained. Already at a qualitative level it has been shown that for the considered runway incursion scenario the event tree based risk model has clear limitations with regard to the representation of the dynamics in the scenario, the interactions between agents, the variability of the contextual conditions in, and the variability of the performance of the agents. As a result of these limitations the event tree approach lacks transparency of the development of the risk model, the quantification of the event probabilities, the risk results and the feedback to the design. In contrast, the MA-DRM uses direct representations and parameterization of the dynamics, agents' interactions, performance variability and contextual conditions, and as a result attains a better transparency for the development of the risk model, the quantification of its parameters, the explanation of its results and the feedback to design.

At a quantitative level, we showed that the differences in the event tree and MA-DRM based approaches gave rise to considerably lower estimates of the accident risk and considerably higher estimates of the risk reduction by the ATC alert system, as assessed by the event tree approach in comparison with the MA-DRM approach. In addition, the quantitative results of the MA-DRM based study made clear that the level of risk is not manifest from the performance of individual human operators and technical systems, nor from the sole relations between human operators and/or technical systems, but from the totality of the performance and interactions of all human operators and technical systems in the operational context considered. This was clearly illustrated for the performance of the air traffic controller in the MA-DRM based study of the runway incursion scenario. Even though the controller very often warns the pilots and frequently does so before the pilots detected the conflict themselves, still the risk increases to a small extent only if the controller would be out of the loop. The performance and interactions between the remaining agents in this case effectively compensate for the lack of controller warnings and restrict the risk increase. The quantitative accident results, which show that safety is the resultant of the totality of the performance and interactions between human operators and technical systems in the operational context considered, are well in line with similar qualitative arguments in the safety literature [1-4].

The emergent behaviour observed in the Monte Carlo simulations means that in assessing the contributions for prevailing accidents by interviewing single operators (pilots and controllers) and by judging their contributions, it is difficult to account for the dependencies of the interactions in conflict scenarios. For instance, based on controller interviews in the event tree based study it was assessed that the controller, when supported by an ATC alert system, would have a large effect on reducing the accident risk of the runway incursion scenario. However, for a controller it is not well possible to judge the probability that a controller warning reaches the pilots before they have detected the conflict independently. Even more difficult is the estimation by a controller of the accident risk reduction of a controller warning, since it supposes an evaluation of all other possibilities by other agents to detect and resolve the conflict scenario. It is well known that expert elicitation for properly accounting of dependencies in risk analysis is a major problem [58] and the event tree based results confirm this problem.

The contrast between the seemingly good performance of a human operator and the limited effect of this performance on the accident risk in a conflict scenario has been found by the Monte Carlo simulations of the MA-DRM based study only. This poses limitations on the safety conclusions that can be attained by other types of simulations. In the air traffic control domain, new concepts are regularly evaluated by human-in-the-loop simulations, in which the performance of (real) air traffic controllers is evaluated in a simulated environment. For operations on the airport this is done in tower simulators, where simulated aircraft movements on the aerodrome are projected in a 360 degrees view, the controllers are supported by their usual ATC systems (which may include alerts) and the controllers can communicate with pseudo-pilots who control the movements of the simulated aircraft. The numbers of aircraft handled in such simulations are similar to what can be achieved in reality, e.g. a runway controller may handle about 25 to 40 aircraft per hour. Human-in-the-loop simulation experiments typically last several days and often aim to evaluate several configurations, typically leading to some hundreds of aircraft handled in a particular configuration. In human-in-the-loop simulations occasionally conflict scenarios may be instantiated and the effectiveness of a controller to detect the conflict and warn pilots may be evaluated. However, the results of this paper indicate that performance measurement of human operators in real-time simulations say little about their contributions to reduction in the accident risk of air traffic scenarios. Consider, for instance, a hypothetical result of a human-in-the-loop simulation experiment that a controller is able to warn the pilots in conflict situations in the large majority of conflicts (say 95%). This might be interpreted as an indication that the controller is contributing considerably to avoiding

accidents, thus forming an important safety barrier. However, the presented results provide an example where the controller warns the pilots in 99% of the cases and still the accident risk would increase only slightly without any contributions of the controller due to the performance of the other agents in the operation. More in general, the results of this paper indicate that if the number of simulations is not sufficiently large to estimate the accident risk in a conflict scenario, it is hard to judge from the performance of individual agents what their effect on safety at the level of accident risk may be.

In conclusion, for the runway incursion scenario example, this paper has confirmed by quantitative accident risk results of the MA-DRM based approach, the in the literature developed qualitative argumentation that safety of a complex socio-technical system is the resultant of the totality of the performance and interactions between human operators and technical systems in the operational context considered. More specifically, it has been shown that the MA-DRM based approach resolves a considerable number of problems that are not well captured by an event sequence based approach. The findings also have significant ramifications for the evaluation and testing of novel air traffic operations: commonly applied analysis processes, such as human-in-the-loop simulations, model development, model validation and feedback to design, appear to have a serious lack in capturing the safety related impacts of interactions between the multiple agents involved in such novel operations.

Now that it has been shown that the MA-DRM safety assessment approach evaluates many safety effects in the runway incursion scenario considered that are not are not evaluated by an event sequence based approach, there is good reason to continue research and development of this approach. Main on-going research directions are:

- To further increase the power of ABM in modelling hazards, which now have to be assessed through bias and uncertainty assessment, e.g. hazards related to complacency, trust, organizational changes and negotiation processes [59, 60].
- To further develop methods for accelerating rare event Monte Carlo simulations [61-63].
- To continue the development of the TOPAZ toolset in support of conducting MA-DRM safety risk assessment for advanced air traffic scenarios, e.g. [64-66].
- To develop an informatics environment in support of the further development of novel TOPAZ toolsets for air traffic scenarios. This should include systematic support in SDCPN specification and Monte Carlo simulation code generation.

8 REFERENCES

[1] Turner BA. Man-made disasters. London, UK: Wykeham Science Press; 1978.

[2] Perrow C. Normal accidents: Living with high-risk technologies. New York, USA: Basic Books; 1984.

[3] Hollnagel E. Barriers and accident prevention. Aldershot, England: Ashgate; 2004.

[4] Dekker S. Drift into failure: From hunting broken components to understanding complex systems. Farnham, England: Ashgate; 2011.

[5] Zio E. Reliability engineering: Old problems and new challenges. Reliability Engineering & System Safety. 2009;94:125-41.

[6] Bedford T, Cooke RM. Probabilistic risk analysis: Foundations and methods. Cambridge, UK: Cambridge University Press; 2001.

[7] Leveson N. A new accident model for engineering safer systems. Safety Science. 2004;42:237-70.

[8] Dougherty EM. Human Reliability Analysis – Where shouldst thou turn? . Reliability Engineering and System Safety. 1990;29:283-99.

[9] ICAO. Manual on the prevention of runway incursions. International Civil Aviation Organization; 2007.

[10] Eurocontrol. European action plan for the prevention of runway incursions, release 1.1. Eurocontrol; 2004.

[11] De Jong HH, Tump RS, Blom HAP, Van Doorn BA, Karwal AK, Bloem EA. Qualitative safety risk assessment of a RIASS based operation at Schiphol airport including a quantitative model: Crossing of departures on 01L/19R under good visibility conditions. Amsterdam, The Netherlands: National Aerospace Laboratory NLR; 2001.

[12] Stroeve SH, Blom HAP, Bakker GJ. Multi-agent situation awareness error evolution in accident risk modelling. Proceedings 5th USA/Europe Air Traffic Management R&D Seminar, Budapest, Hungary, 2003.

[13] Stroeve SH, Blom HAP, Bakker GJ. Systemic accident risk assessment in air traffic by Monte Carlo simulation. Safety Science. 2009;47:238-49.

[14] Scholte JJ, Blom HAP, Van den Bos JC, Jansen RBHJ. Management of ATM performance in operational concept development and validation: a case study. Eight USA/Europe Air Traffic Management R&D Seminar, Napa, USA, 2009.

[15] Blom HAP, Stroeve SH, Scholte JJ, De Jong HH. Accident risk analysis benchmarking Monte Carlo simulation versus event sequences. Third

International Conference on Research in Air Transportation (ICRAT 2008), Fairfax (VA), USA, 2008.

[16] Stroeve SH, Blom HAP, Bakker GJ. Comparison of accident risk assessment by event sequence analysis versus Monte Carlo simulation. Eurocontrol Safety R&D Seminar, Southampton, UK, 2008.

[17] Stroeve SH, Blom HAP, Bakker GJ. Contrasting safety assessment of a runway incursion scenario by event trees and agent-based dynamic risk modelling. Ninth USA/Europe Air Traffic Management R&D Seminar, Berlin, Germany, 2011.

[18] Blom HAP, Bakker GJ, Blanker PJG, Daams J, Everdij MHC, Klompstra MB. Accident risk assessment for advanced air traffic management. In: Donohue GL, Zellweger AG, editors. Air Transport Systems Engineering: AIAA; 2001. p. 463-80.

[19] Blom HAP, Stroeve SH, De Jong HH. Safety risk assessment by Monte Carlo simulation of complex safety critical operations. In: Redmill F, Anderson T, editors. Developments in Risk-based Approaches to Safety: Proceedings of the Fourteenth Safety-citical Systems Symposium, Bristol, UK, 7-9 February 2006: Springer; 2006.

[20] Bonabeau E. Agent-based modeling: methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences of the USA. 2002;99:7280-7.

[21] Burmeister B, Haddadi A, Matylis G. Applications of multi-agent systems in traffic and transportation. IEE Transactions on Software Engineering. 1997;144:51-60.

[22] Shah AP, Pritchett AR, Feigh KM, Kalarev SA, Jadlav A, Corker KM, et al. Analyzing air traffic management systems using agent based modeling and simulation. Proceedings of the 6th USA/Europe Seminar on Air Traffic Management Research and Development, Baltimore (MD), USA, 2005.

[23] Endsley MR. Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society. 1995;37:32-64.

[24] Endsley MR. A taxonomy of situation awareness errors. In: Fuller R, Johnston N, McDonald N, editors. Human factors in aviation operations. Aldershot, UK: Ashgate Publishing; 1995. p. 287-92.

[25] Kurzhanski AB, Varaiya P. On reachability under uncertainty. SIAM Journal on Control and Optimization. 2002;41:181-216.

[26] Prandini M, Hu J. A stochastic approximation method for reachability computations. In: Blom HAP, Lygeros J, editors. Stochastic Hybrid Systems: Theory and Safety Critical Applications. Berlin, Germany: Springer; 2006.

[27] Labeau PE, C. S, Swaminathan S. Dynamic reliability: towards an integrated platform for probabilistic risk assessment. Reliability Engineering and System Safety. 2000;68:219-54.

[28] Davis MHA. Markov Models and Optimization. London, UK: Chapman & Hall; 1993.

[29] Bujorianu ML, Lygeros J. Reachability questions in piecewise deterministic Markov processes. In: Mahler O, Pnuelli A, editors. Proceedings of Hybrid Systems Computation and Control. 2623 ed. Berlin, Germany: Springer; 2003. p. 126-40.

[30] Pola G, Bujorianu ML, Di Benedetto MD. Stochastic hybrid models: an overview with applications to air traffic management. Proceedings of IFAC Conf Analysis and Design of Hybrid Systems (ADHS), Saint-Malo, Brittany, France, 2003.

[31] Bujorianu ML. Extended stochastic hybrid systems. In: Mahler O, Pnuelli A, editors. Proceedings of Hybrid Systems Computation and Control. 2993 ed. Berlin, Germany: Springer; 2004. p. 234-49.

[32] Bujorianu ML, Lygeros J. Towards a general theory of stochastic hybrid systems. In: Blom HAP, Lygeros J, editors. Stochastic Hybrid Systems: Theory and Safety Critical Applications. Berlin, Germany: Springer; 2006. p. 3-30.

[33] Krystul J, Blom HAP, Bagchi A. Stochastic hybrid processes as solutions to stochastic differential equations. In: Cassandras CG, Lygeros J, editors. Stochastic hybrid systems: Recent developments and research trends: CRC Press; 2007. p. 15-45.

[34] David R, Alla H. Petri Nets for the modeling of dynamic systems - A survey. Automatica. 1994;30:175-202.

[35] Sadou N, Demmou H. Reliability analysis of discrete event dynamic systems with Petri nets. Reliability Engineering & System Safety. 2009;94:1848-61.

[36] Kleyner A, Volovoi V. Application of Petri nets to reliability prediction of occupant safety systems with partial detection and repair. Reliability Engineering & System Safety. 2010;95:606-13.

[37] Bouali M, Barger P, Schon W. Backward reachability of Colored Petri Nets for systems diagnosis. Reliability Engineering & System Safety. 2012;99:1-14.

[38] Ghazel M. Using stochastic Petri nets for level-crossing collision risk assessments. IEEE Transactions on Intelligent Transportation Systems. 2009;10:668-77.

[39] Cassandras CG, Lafortune S. Introduction to discrete event systems: Second edition: Springer; 2008.

[40] Jensen K. Coloured Petri Nets: Basic Concepts, Analysis Methods and Practical Use. London, UK: Springer; 1992.

[41] Haas PJ. Stochastic Petri Nets, Modeling, Stability, Simulation. New York, USA: Springer-Verlag; 2002.

[42] Everdij MHC, Blom HAP. Petri nets and hybrid state Markov processes in a power-hierarchy of dependability models. Proceedings of IFAC Conference on Analysis and Design of Hybrid Systems, Saint-Malo Brittany, France, 2003.

[43] Everdij MHC, Blom HAP. Piecewise deterministic Markov processes represented by Dynamically Coloured Petri Nets. Stochastics. 2005;77:1-29.

[44] Everdij MHC, Blom HAP. Hybrid Petri Nets with diffusion that have into-mappings with generalised stochastic hybrid processes. In: Blom HAP, Lygeros J, editors. Stochastic Hybrid Systems: Theory and Safety Critical Applications. Berlin, Germany: Springer; 2006. p. 31-63.

[45] Everdij MHC, Blom HAP. Hybrid state Petri nets which have the analysis power of stochastic hybrid systems and the formal verification power of automata. In: Pawlewski P, editor. Petri Nets. Vienna, Austria: I-Tech Education and Publishing; 2010. p. 227-52.

[46] Everdij MHC, Klompstra MB, Blom HAP, Obbink BK. Compositional specification of a multi-agent system by stochastically and dynamically coloured Petri nets. In: Blom HAP, Lygeros J, editors. Stochastic hybrid systems: Theory and safety critical applications: Springer; 2006. p. 325-50.

[47] Blom HAP, Bakker GJ, Everdij MHC, Van der Park MNJ. Collision risk modelling in air traffic. IFAC European Control Conference 2003, Cambridge, UK, 2003.

[48] Blom HAP, Krystul J, Bakker GJ, Klompstra MB, Klein Obbink B. Free flight collision risk estimation by sequential Monte Carlo simulation. In: Cassandras CG, Lygeros J, editors. Stochastic hybrid systems; recent developments and research trends. Boca Raton, USA: CRC Press; 2007. p. 249-81.

[49] Cerou F, Del Moral P, Le Gland F. Limit theorems for the multilevel splitting algorithms in the simulation of rare events. Proceedings of Winter Simulation Conference, Orlando (FL), USA, 2005.

[50] Villen-Altamirano M, Villen-Altamirano J. Restart, a method for accelerating rare event Monte Carlo simulations. Proceedings 13th Int Teletraffic Congress, Copenhagen, Danmark, 1991.

[51] Everdij MHC, Blom HAP, Stroeve SH. Structured assessment of bias and uncertainty in Monte Carlo simulated accident risk. Proceedings 8th International Conference on Probabilistic Safety Assessment and Management, New Orleans, USA, 2006.

[52] Everdij MHC, Blom HAP. Bias and uncertainty modelling in accident risk assessment. Amsterdam, The Netherlands: National Aerospace Laboratory NLR; NLR-CR-2006-284; 2005.

[53] Everdij MHC, Scholte JJ, Blom HAP, Stroeve SH. An investigation of emergent behaviour viewpoints in literature, and their usability in air traffic management. First Complex World Annual Conference, Seville, Spain, 2011.

[54] Bedau MA. Weak emergence. In: Tomberlin J, editor. Philosophical Perspectives: Mind, Causation, and World. Malden (MA), USA: Blackwell; 1997. p. 375-99.

[55] Corning PA. The re-emergence of “emergence”: A venerable concept in search of a theory. Complexity. 2002;7:18-30.

[56] Chalmers DJ. Strong and weak emergence. In: Clayton P, Davies P, editors. The re-emergence of emergence: Oxford University Press; 2006.

[57] Cardosi K, Chase S, Eon D. Runway safety. Air Traffic Control Quarterly. 2010;18:303-28.

[58] Cooke RM, Goossens LLHJ. TU Delft expert judgment data base. Reliability Engineering & System Safety. 2008;93:657-74.

[59] Stroeve SH, Everdij MHC, Blom HAP. Hazards in ATM: Model constructs, coverage and human responses. SESAR Joint Undertaking; E.02.10-MAREA-D1.2; 2011.

[60] Stroeve SH, Blom HAP. How well are human-related hazards captured by multi-agent dynamic risk modelling? In: Landry SJ, editor. Advances in human aspects of aviation. Boca Raton (FL), USA: CRC Press; 2012. p. 462-71.

[61] Blom HAP, Bakker GJ, Krystul J. Rare event estimation for a large-scale stochastic hybrid system with air traffic application. In: Rubino G, Tuffin B, editors. Rare event simulation using Monte Carlo methods: Wiley; 2009. p. 193-214.

[62] Prandini M, Blom HAP, Bakker GJ. Air traffic complexity and the interacting particle system method: An integrated approach for collision risk estimation. Proc American Control Conf, ACC 2011. San Francisco, CA, USA2011. p. 2154-9.

[63] Bakker GJ, Goswami A, Blom HAP. Sequential Monte Carlo simulation using periodic boundary condition. Proc 9th Int Workshop on Rare Event simulation (RESIM2012), Trondheim, Norway, 2012.

[64] Blom HAP, Klein Obbink B, Bakker GJ. Simulated safety risk of an uncoordinated airborne self separation concept of operation. ATC Quarterly. 2009;17:63-93.

[65] Everdij MHC, Blom HAP, Bakker GJ, Zmarrou H. Agent-Based Safety Risk Analysis of Time Based Operation in Future TMA. Proceedings 3rd ATOS Conference, Delft, The Netherlands, 2011.

[66] Blom HAP, Bakker GJ. Safety of advanced airborne self separation under very high en-route traffic demand. Proceedings SESAR Innovation Days, Toulouse, France, 2011.