Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)

Contrasting Safety Assessments of a Runway Incursion Scenario by Event Sequence Analysis versus Multi-Agent Dynamic Risk Modelling

Sybert H. Stroeve, Henk A.P. Blom, G.J. (Bert) Bakker
Air Transport Safety Institute, National Aerospace Laboratory NLR
Anthony Fokkerweg 2, 1059 CM Amsterdam, The Netherlands
E-mail: stroeve@nlr.nl, blom@nlr.nl, bakker@nlr.nl

***Abstract* — Recently we compared safety analyses for a runway incursion scenario based on an event sequence analysis, as a key exponent of a traditional risk assessment technique, versus one based on an agent-based dynamic risk model (DRM), as an exponent of new techniques based on system complexity and variability-based accident models. We found that lower accident risk levels were assessed in the event sequence analysis and we compared various factors contributing to these differences. As the reasons of these differences were not completely understood, this paper sets forth additional analyses towards a better understanding of the relations between conflict recognition and resolution events that may occur in the runway incursion scenario and their relation to accident risk. To this end, such events were recorded in additional Monte Carlo simulations of the DRM and a broader set of conditions was considered with agents being in or out of monitoring/control loops. The results show that the accident risk can be very elastic for changes in the operation. The level of this risk elasticity 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 only from the totality of the performance and interactions of all human operators and technical systems in the operational context considered. Implications for real-time simulations, expert judgement and feedback to design are discussed.**

Keywords- Event tree; dynamic risk model; runway incursion; air traffic control; accident risk; human performance; multi-agent

I. INTRODUCTION

In complex and distributed socio-technical organizations the level of safety depends on the interactions between many entities of various types at multiple locations. 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, each of which would be unlikely, singly, to defeat the established safety systems. Also Perrow [2] describes accidents as the consequence of complex interactions and tight couplings in socio-technical systems in his Normal Accident theory, stressing that given such system characteristics, multiple and unexpected interactions of failure conditions are inevitable. Building forward on the notion of normal accidents, Hollnagel [3] argues that performance in complex systems is necessarily variable due to the performance variability of its entities and the complexity of their interactions. Reasons for variability in the performance of humans include the dependency on contextual conditions, the efficiency-thoroughness trade-off in their performance and the intrinsic variability of perceptual and cognitive functions. Accidents may occur as a result of the interactions, performance variability, failures and contextual conditions of the socio-technical system.

A detailed account of complex interactions and performance variability is typically lacking in probabilistic risk assessments (PRA) of socio-technical systems by traditional event sequence-based techniques such as fault trees (FTs) and event trees (ETs). FTs represent relations between events and conditions leading to a safety-relevant situation and ETs represent relations between possible events following such a situation and the resulting consequences (e.g. accidents). They are pictorial representations of Boolean logic relations between events and they use event probabilities in PRA. The probabilities of the end events can thus be calculated straightforwardly and these end results are qualities of the same kind as the data used to obtain them: both are event probabilities. FT and ETs have been applied extensively for safety assessment in various fields, including air traffic [4][5]. An advantage of these techniques is that their structure is transparent and easy to understand. Their limitations include the difficultness to represent varieties of interdependencies between organizational entities and their dynamics, as well as the restricted evaluation of human performance by human error and conflict resolution probabilities. As such their use for risk assessment of complex socio-technical systems tends to be problematic [3][6].

In recognition of the limitations of event sequence-based techniques and in an effort to more directly address performance variability in complex socio-technical systems and the therein emergent safety risks, various methods have been developed. These developments include FRAM [3],

which pursues a qualitative analysis of safety-critical interdependencies in a functional model of an operation, STAMP [6], which uses system theoretic modelling of control loops and processes to obtain quantitative results on safety-related process variables, and TOPAZ [7], which uses agent-based dynamic risk models (DRM) to obtain accident risk probabilities of air traffic scenarios. In agent-based DRM accident risk is an emergent property [8][9] that is obtained by simulation of the dynamics of interacting elements in safety relevant scenarios and which uses data of these dynamics that is of a completely different nature than the accident risk. Although system complexity and performance variability-based safety assessment methods are not yet part of the standard repertoire of techniques and are being further developed, they have already been applied in several practical safety assessments, such as assessment of NASA's safety culture by STAMP [10] or risk assessment of operations of the ANSP in the Netherlands (LVNL) by TOPAZ.

To relate these two ways of thinking about the development of accidents, we performed a benchmark study for safety analyses of a particular runway incursion scenario [11][12]. In these papers we compared the results of an event sequence-based analysis with those of an assessment using an agent-based DRM. We found that lower accident risk levels were assessed in the event sequence analysis and we compared various factors contributing to these differences. As the reasons of these differences were not completely understood, this paper goes beyond benchmarking by running additional Monte Carlo simulations in order to gain a better understanding of the relations between conflict recognition and resolution events that may occur in the runway incursion scenario and their relation to accident risk. Furthermore, this paper sets forth to contrast the probability of agents' conflict recognition and conflict resolution events with the risk effects of a broader set of changes in the operation, where agents are in or out of monitoring/control loops. In this context we adopt the notion of risk elasticity to argue about changes in risk in response to changes in the socio-technical system; the risk elasticity is high if the changes have only a small effect on the level of risk.

This paper is organized as follows. Section II introduces the runway-incursion related safety studies. Section III describes the methods and results of the event sequence-based safety assessment. Section IV describes the methods and results of the DRM-based safety assessment. Section V defines additional events in the MC simulations of the DRM and the results achieved. Section VI presents the risk elasticity results for scenarios with agents being in or out of the monitoring and control loops. Section VII discusses the results of the event sequence and DRM approaches and their implications. Section VIII presents the conclusion of this research.

II. RUNWAY INCURSION-RELATED SAFETY STUDIES

A. *Runway incursion*

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" [13]. Within air traffic, the risk of runway incursion is recognised as an important safety issue. Safety programmes such as [13][14] promote procedures and training to reduce runway incursion risk, such as following ICAO compliant procedures and naming, applying standard radiotelephony (R/T) phraseology, pilot training on aerodrome signage and markings, using standard taxi routes, etc. In addition, technology is being used and developed to reduce the likelihood and consequences of runway incursions, such as alerting systems and traffic displays. Assessment of runway incursion risk and of the potential effect of runway incursion risk reducing measures and technologies are demanding tasks, given the large number of human operators, aircraft and supporting technical systems that closely interact on the aerodrome. This complexity makes runway incursion-related safety assessments suitable candidates for comparison of the two types of accident models.

B. *Safety assessments in support of taxiing operations at Amsterdam airport*

The two safety assessments were done for an active runway crossing operation in good visibility conditions. The operation was proposed for crossing of runway 18C/36C at Amsterdam airport for traffic coming from and going to a new parallel runway 18R/36L. During the development of infrastructure and operational concepts for taxiing to the new runway, various risk assessment studies were done; their history is described in detail in [15]. These studies included the use of event sequences for the assessment of the risk of various safety relevant scenarios of the active runway crossing operation [16]. Having recognized the complexity of some of these scenarios, this led to the development of a DRM for a scenario of the active runway crossing operation [17][18]. Since this DRM was developed for the same operation and considered the same set of hazards contributing to the safety relevant scenario, the models and results of these two studies provide a suitable basis for the comparison of event sequence and DRM-based risk assessment approaches.

C. *Active runway crossing operation*

As the focus in this study is not on the specific results for Amsterdam airport obtained in the safety assessments, but rather on the followed lines of reasoning, in the remainder of the paper we discuss the operation and its context in generic terms. The runway considered is used for departures and has a taxiway that crosses the runway at a distance of 1000 m from the runway threshold.

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.

Aircraft may taxi across the active runway via the following procedure. First, the control over the taxiing aircraft is

transferred from a ground controller to the runway controller. The runway controller specifies a crossing clearance to the taxiing aircraft and switches off the remotely controlled stopbar. The crew of the taxiing aircraft acknowledges the clearance, initiates taxiing across the runway and reports when the taxiing aircraft has vacated the runway. After passage of the stopbar, it is automatically switched on again.

Standard communication, navigation and surveillance systems are used: communication between controllers and crews is by R/T systems, the pilots use their knowledge on the aerodrome layout and 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 runway incursion alert for the situation that an aircraft is on the runway in front of an aircraft that has initiated to take off; (2) a stopbar violation alert for the situation that an aircraft crosses an active stopbar in the direction of the runway.

III. EVENT SEQUENCE-BASED SAFETY STUDY

In the safety assessment of [16] several safety relevant scenarios were considered for the active runway crossing operation. For the purpose of the comparison in this study, we focus on a scenario that an aircraft is taking off and a taxiing aircraft is crossing the runway while it should not; thus a runway incursion is due to the taxiing aircraft. In particular, the event sequence-based study considers that the pilot of the taxiing aircraft starts crossing without contacting the runway controller (e.g. by misunderstanding the ground controller). The ET of the runway incursion scenario, given the taxiing aircraft is crossing while it should not, considers 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 ET differentiate between early, medium and late recognition of the conflict by the pilots and the runway controller. 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. The outcomes of the ET specify the timing of the resolution of the conflict (early/medium/late) or the inability to timely resolve it (accident).

The parameter values of the ET are the probabilities of event occurrences. In the event sequence-based assessment, lower and upper bounds of the event probabilities were estimated by expert (controller and pilot) elicitation. 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. These probabilities must be interpreted as conditional probabilities in the ET.

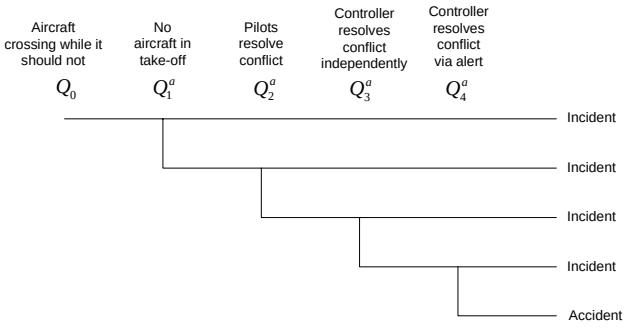

Figure 1: Aggregated ET for runway incursion scenario.

The discussed ET can be condensed in a simpler aggregated ET shown in Figure 1, which neglects the resolution stage (early/medium/ late) and focuses on the contributions of no aircraft in take-off during the crossing (event Q_1^a), direct conflict recognition and resolution by the pilots (event Q_2^a), conflict recognition by the controller independently from the alert system that leads to effective warning of the pilots and resolution of the conflict by the pilots (event Q_3^a), and conflict recognition by the controller as result of an alert that leads to effective warning of the pilots and resolution of the conflict by the pilots (event Q_4^a).

TABLE 1: EVENT PROBABILITIES OF THE AGGREGATED ET.

Event		Event probability		
		Lower bound	***Geometric mean***	***Upper bound***
Q_1^a	No aircraft in take-off	0.75	0.75	0.75
Q_2^a	Pilots resolve conflict	0.995	0.99961	0.99997
Q_3^a	Controller resolves conflict independently	0.38	0.52	0.71
Q_4^a	Controller resolves conflict via alert	0.906	0.938	0.970
Accident (given aircraft crossing while it should not)		6.5E-8	2.2E-6	7.5E-5

The probabilities of the events in the aggregated ET are shown in Table 1. These data include the accident risk and they reveal that in the event sequence-based safety assessment it has been assumed that the pilots have a large contribution to avoiding a collision for the runway incursion scenario (about 99.96% of the cases), the controller can only add to this independently in about half of the cases, and the controller can effectively add to the collision avoidance after an ATC alert in about 94% of the cases. An explanation of the small contribution of the controller independent from the alert system is, that as the pilot of 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. In contrast, the effectiveness of the alert system is assessed to be high as it reduces the risk by a factor 16.

IV. DRM-BASED SAFETY STUDY

A. Dynamic risk model

The multi-agent DRM of the runway incursion scenario is specified by a stochastic dynamic extension of the Petri net formalism [19] and is discussed in more detail in [18]. The main agents are the aircraft taking-off and taxiing, the pilots flying of the aircraft, the runway controller and the ATC system. Key aspects of the models of these agents are highlighted next.

Taking-off Aircraft (AC-TO): The model of the taking-off aircraft represents the ground run, airborne transition and airborne climb-out phases during take-off and includes the possibility of a rejected take-off. The aircraft initiates take-off from a position near the runway threshold and it may be medium-weight or heavy-weight.

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 and its entrance time is uniformly distributed around the take-off time of AC-TO. The aircraft may be medium-weight or heavy-weight.

Surveillance (ATC subsystem): 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.

Alerts (ATC subsystem)**:** 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.

R/T (ATC subsystem): 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.

Pilot flying of the Taking-off Aircraft (PF -TO): The model for the performance of PF-TO 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. 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.

Pilot Flying of Taxiing Aircraft (PF-TX): The model structure of PF-TX is similar to that of PF-TO. 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 at stochastically distributed times. PF-TX may detect a conflict if AC-TX is within a critical 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. Following conflict detection, PF-TX starts a collision avoiding 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.

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 the 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.

B. Risk assessment results

A key result of the Monte Carlo simulations is the probability of collision between the aircraft taxiing and taking-off. Since collision risks considered in air traffic are small, simulation speed-up by risk decomposition has been applied. Results presented earlier [17][18] indicate that a wrong intent situation awareness of the pilot flying of the taxiing aircraft is a condition with a strong effect on the accident risk. For the comparison with the risk results of the ET 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 crosses the runway without contacting the runway controller, which is the condition considered in the event sequence-based risk assessment.

Although the multi-agent DRM considers a considerable amount of performance and interaction aspects of the agents, as any model it differs from reality. To identify these differences and evaluate their effect at the level of risk, a bias and uncertainty assessment method is an integrated part of the TOPAZ risk assessment methodology [20]. Results of a bias and uncertainty assessment are reported in [18] and they reveal that lack of knowledge on pilot performance contributes mostly to uncertainty in the risk, and lack of knowledge on controller performance and ATC systems hardly contributes to uncertainty in the risk.

Overall, the following results were obtained for the conditional collision risk given the particular runway incursion scenario: a point estimate of 1.8E-4 per take-off and a 95% uncertainty interval of [4.1E-6, 7.3E-4] per take-off. These risk levels are considerably higher than those found in the event sequence-based assessment (Table 1), as already recognized in [11][12].

V. EVENTS IN THE MC SIMULATIONS

A. Definition of events

To improve the insight in the performance of the agents in the DRM, the relation of this performance with the accident risk and to support the comparison with the event sequence-based analysis, we defined and recorded event occurrences in the Monte Carlo simulations of the agent-based DRM. As an onset for the analysis of event occurrences in the Monte Carlo simulations of the agent-based DRM, Figure 2 presents events for conflict recognition and collision avoidance actions by the agents as well as relations between these events. For instance, Figure 2 indicates that an active stopbar violation alert (event E8) may result in conflict detection by ATCo-R (event E5) and this event, on its turn, may result in warnings specified by ATCo-R towards the PFs of both aircraft (events E6 and E7). The times of first occurrence of most events were recorded in the MC simulations. The occurrence of events E1', E3' and E5' was inferred from the occurrence of related events.

B. Results of event occurrences

A total of 10 million Monte Carlo simulation runs were performed for the condition that the PF TX 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 1.8E-4 for this condition found earlier. Table 2 shows the probabilities of the defined events and the conditional probabilities of these events given a collision. Key observations and explanations of the results in Table 2 are discussed next.

The pilots of both aircraft detect the conflict in more than 99% of the simulated conflict scenarios (events E1, E3). In the cases resulting in a collision, the probability of detection of the conflict is even higher for the PF-TO (99.7%), but it is considerably lower for the PF-TX (91.3%).

The controller detects the conflict in 99.3% of the simulated conflict scenarios in general and in 99.9% of the cases resulting in a collision (event E5).

The stopbar violation alert is active in 94% of the scenarios in general and in 99.9% of the cases resulting in a collision (event E8). Mostly, it is not activated in the cases that AC-TX stops close after the stopbar, such that the alert threshold has not yet been passed. In the critical cases it has almost always been active.

The runway incursion alert is active in 34% of the scenarios in general and in 99.9% of the cases resulting in a collision (event E9). It is not activated in the cases 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. In the critical cases it has almost always been active.

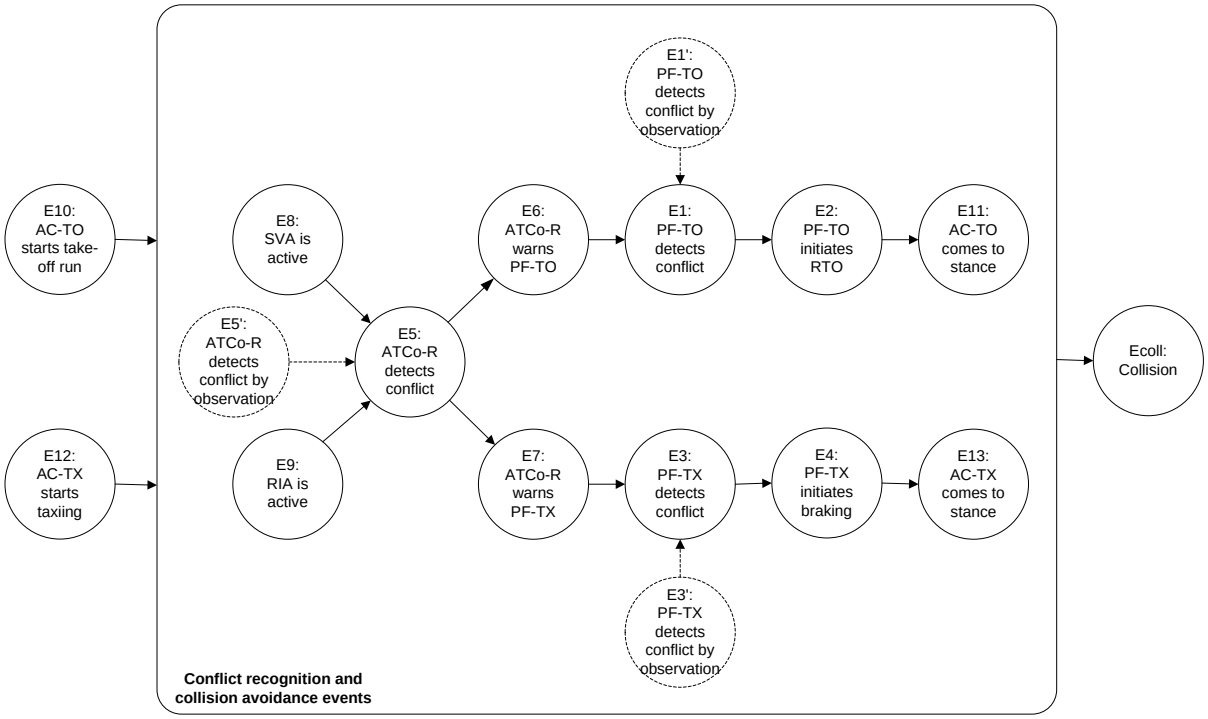

Figure 2: Relations between events in the MC simulations of the DRM. Events in solid circles are recorded in the MC simulations, events in dashed circles are inferred from the relative timing of recorded events.

TABLE 2: MC SIMULATION RESULTS FOR THE DEFINED EVENTS: EVENT PROBABILITY AND CONDITIONAL EVENT PROBABILITY GIVEN A COLLISION.

Agent	Event		Probabilities	
	ID	*Description*	*P(Eq)*	*P(Eq|Ecoll)*
PF-TO	E1	Detects conflict	9.92E-01	9.97E-01
PF-TO	E1'	Detects conflict by own observation	4.18E-02	5.89E-01
PF-TO	E2	Initiates rejected take-off	5.66E-01	2.39E-01
PF-TX	E3	Detects conflict	9.98E-01	9.13E-01
PF-TX	E3'	Detects conflict by own observation	2.21E-01	7.51E-01
PF-TX	E4	Initiates braking	6.88E-01	7.13E-01
ATCo-R	E5	Detects conflict	9.93E-01	9.99E-01
ATCo-R	E5'	Detects conflict by own observation	3.93E-01	2.28E-01
ATCo-R	E6	Warns PF-TO	9.93E-01	9.54E-01
ATCo-R	E7	Warns PF-TX	9.93E-01	5.69E-01
ATC System	E8	Stopbar violation alert is active	9.40E-01	9.99E-01
ATC System	E9	Runway incursion alert is active	3.41E-01	9.99E-01
AC-TO	E10	Start take-off run	1	1
AC-TO	E11	Come to stance	5.66E-01	0.00E00
AC-TX	E12	Start taxiing	1	1
AC-TX	E13	Come to stance	6.87E-01	2.95E-01
AC-TO AC-TX	Ecoll	Collision	1.81E-04	1

In general, the controller warns the pilots of both aircraft in 99.3% of the simulated conflict scenarios (events E6, E7), which is equal to the detection rate by the controller (event E5). However, in the cases resulting in a collision, the probability of a warning is decreased to 95.5% 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 the ATCo-R, leading to a delay in the communication such that the probability increases that PF-TX cannot be warned before a collision occurs.

Overall, ATCo-R warns PF-TO before this PF has detected the conflict independently in 95.8% of the cases (event E1'). 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. In this context, ATCo-R can very often effectively warn PF-TO. In the cases resulting in a collision, the fraction of effective warnings is decreased considerably to from 95.8% to 41.1% (event E1').

Overall, ATCo-R warns PF-TX before this PF has detected the conflict independently in 77.9% of the cases (event E3'). 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 AC-TX has passed the stopbar. In this context, ATCo-R can quite often effectively warn PF-TX. In the cases resulting in a collision, the fraction of effective warnings is decreased considerably from 77.9% to 24.9% (event E3').

Additional results (not shown) indicate that overall, ATCo-R warns at least one of the pilots before they detected the conflict independently in 97.7% of the cases. In the cases resulting in a collision this fraction is reduced to 51.4%.

In general, the ATC alerting systems are able to warn ATCo-R in 60.7% of the cases before ATCo-R has detected the conflict independently (event E5'). In the cases resulting in a collision, the fraction of effective warnings is increased to 77.2%, which is in line with the earlier observed larger probability of an alert in these most dangerous cases.

PF-TO initiates a rejected take-off (RTO) in 56.6% of the cases (event E2) and also in 56.6% of the cases AC-TO comes to stance (event E11). In the cases resulting 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. The simulation results show that situations in which AC-TO comes to stance at the runway crossing position and AC-TX then collides with it did not occur. The results also indicate that although the RTO initiation almost always results in a stop of AC-TO, there are some cases where it did not succeed in stopping before the crossing.

PF-TX initiates braking in 68.8% of the cases (event E4) and in 68.7% of the cases (event E13) AC-TX comes to stance. In the cases that resulted in a collision, braking was initiated in 71.3% of the cases and the aircraft came to stance in 29.5% of the cases.

Other detailed results (not shown) indicate that AC-TO is predominantly well within the first 500 m of the runway when the conflict is detected by either of the agents (PF-TO, PF-TX, ATCo-R, ATC System) or when the agents take action to prevent an accident. In contrast, for the cases resulting in a collision, these events often occur when AC-TO is between 500 m and 1000 m; only for the detection of the conflict by the controller and the ATC alerts a considerable part of the PDF is below 500 m.

Similar results (not shown) for AC-TX show that overall the front-wheel of AC-TX is predominantly within 100 m from the runway centre-line when the conflict is detected by PF-TO or PF-TX, and when they start their collision avoiding actions. Overall, the controller detects the conflict at an earlier stage, predominantly when the AC-TX is between 150 and 100 m, and this range overlaps with that of the stopbar violation alert. However, at the time that the controller has warned the pilots, AC-TX is predominantly already within 100 m from the runway centre-line. There is a considerable overlap between the cores of the PDFs of the position of AC-TX in general and given the occurrence of a collision.

VI. SIMULATION OF CHANGES IN THE OPERATION

A. Agents in/out of the monitoring and control loops

The results of the analysis in last section provided insight in the performance of the various agents in the runway incursion scenario and its relation with collision risk. To obtain insight in the change in risk due to changes in the operation, we also performed Monte Carlo simulations in which we placed one or more agents in or out of the monitoring or control loop. 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 are:

- PF-TX does not actively monitor the traffic situation visually, such that PF-TX may only detect a conflict via a call of ATCo-R;
- PF-TO does not actively monitor the traffic situation visually, such that PF-TO may only detect a conflict via a call of ATCo-R;
- ATCo-R cannot communicate with the pilots;
- ATC System does not specify alerts.

For all relevant combinations of agents in or out of the monitoring or control loop, the conditional collision risk of the runway incursion scenario considered in this paper was determined by Monte Carlo simulation. This gives rise to 12 relevant combinations of conditions, which are shown in Table 3. Note that for conditions where ATCo-R is out of the control loop, it does not matter whether or not the ATC alerts are included in the control loop, as these can only be effective via ATCo-R. The runway incursion scenario considered earlier is case T1. For convenience Table 3 includes risk factors with respect to the lowest risk as obtained for case T1.

TABLE 3: CONDITIONAL COLLISION RISK RESULTS FOR VARIOUS CONDITIONS WITH AGENTS OUT OF THE MONITORING/CONTROL LOOP.

Case	PF-TX	PF-TO	ATCo-R	Alerts	Risk	Risk factor
T1	yes	yes	yes	yes	1.8E-4	1
T2	no	yes	yes	yes	1.0E-2	56.6
T3	yes	no	yes	yes	3.4E-4	1.89
T4	yes	yes	no	yes/no	2.2E-4	1.22
T5	no	no	yes	yes	1.7E-2	94.4
T6	no	yes	no	yes/no	1.7E-2	94.4
T7	yes	no	no	yes/no	1.9E-2	106
T8	no	no	no	yes/no	9.4E-2	522
T9	yes	yes	yes	no	1.9E-4	1.06
T10	no	yes	yes	no	1.2E-2	66.7
T11	yes	no	yes	no	2.1E-3	11.7
T12	no	no	yes	no	3.4E-2	189

B. Risk elasticity for changes in the operation

Next we discuss key results of Table 3 in relation with the earlier presented results on events in the MC simulations of the runway incursion scenario (Table 2).

The collision risk of the runway incursion scenario is increased only by 6% if the ATC alert systems are not available (T9); thus the risk elasticity of the operation is considerable for exclusion of the alert system. This can be contrasted with the earlier found results showing that the alerting systems are able to effectively warn the controller. Although the ATC alert system is effective in this way, the other agents can well cope without the alerting system and keep the risk increase very modest.

The collision risk of the runway incursion scenario is increased only by 22% if the controller would always be out-of-the-loop (T4); thus the risk elasticity with respect to the controller performance is high. This result is quite surprising, given the earlier presented result that the controller warns the pilots flying of the taking-off and taxiing aircraft in 96% and 78% of the cases before they have detected the conflict independently. Notwithstanding this performance, the pilots can mostly detect the conflict themselves and react timely, such that the risk increase is modest if the controller is placed out of the control loop.

The collision risk of the runway incursion scenario is increased by a factor 522 if none of the agents would be actively monitoring the traffic situation (T8). 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. The accident risk of case T8 thus forms an upper bound for this particular runway incursion scenario.

The collision risk of the runway incursion scenario is increased by 89% in the (hypothetical) case that PF-TO would not be actively monitoring the traffic situation, but might still be warned by ATCo-R (T3). In comparison with the risk increase for the case without an actively monitoring PF-TX (T2) this is still a modest risk increase. Explanations of this modest increase are: (1) Already in the nominal case T1 ATCo-R effectively warns PF-TO very frequently (96%); (2) ATCo-R can often effectively warn PF-TO if AC-TO is in the first 500 m of the runway and ATCo-R; (3) ATCo-R can detect the conflict often before PF-TO, since ATCo-R considers AC-TX to be conflicting as soon as it has passed the stopbar, while the PF-TO considers it conflicting as soon as AC-TX is within 90 m from the runway centre-line.

The collision risk of the runway incursion scenario is increased by a factor 57 in the (hypothetical) case that PF-TX would not be actively monitoring the traffic situation, but might still be warned by ATCo-R (T2). This is a large risk increase and it is considerably larger than the risk increases for the situations without effective monitoring by ATCo-R (T4) or by PF-TO (T3). Explanations of this larger increase are: (1) Already in the nominal case T1, ATCo-R can effectively warn PF-TX to a smaller extent than PF-TO; (2) There is a large overlap between the timing and position of AC-TX for the warning of PF-TX by ATCo-R in the cases with a collision versus all cases. Thus a 'too late' warning by ATCo-R in the case of a collision can be considered to be quite normal. Thus, active monitoring of the traffic situation is more important for PF-TX than for PF-TO, as ATCo-R can more effectively warn PF-TO.

In the cases that only one of the involved human operators is actively monitoring the traffic situation (T5, T6, T7) similar conditional collision risks are attained. These risk levels are about a factor 100 above the risk with all humans in the monitoring loop and about a factor 5 below the risk without any of the humans in the loop. These results indicate that the risk elasticity is low for the hypothetical case that only one human would be monitoring actively.

Cases T10, T11, T12 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. T2, T3 and T5, respectively) that the

effect of the non-availability of the ATC alert systems differs considerably. In the case without active monitoring by PF-TX (T10 versus T2) the risk increase is quite limited (20%), indicating that in situations that may result in conflicts the alerts are often too late to warn the PF-TX. In the case without active monitoring by PF-TO (T11 versus T3) the risk increase is considerable (600%), indicating that in this context the ATC alerts often warn ATCo-R such that ATCo-R can timely warn PF-TO.

VII. Discussion

A. Risk levels and contributions

The risk levels obtained in both risk assessment studies differ. The conditional accident probability given the runway incursion scenario is 2.2E-6 (geometric mean) / 7.5E-5 (upper bound) in the event sequence-based safety assessment, whereas the conditional accident probability is 1.8E-4 (point estimate) / 7.3E-4 (upper bound) in the DRM-based safety assessment. The accident risk was thus assessed to be considerably lower by the event sequence-based assessment in comparison with the DRM-based assessment.

With regard to the risk reduction contributions of involved agents, easy understandable results could be derived from the ET. These indicate that pilots can avoid a collision in about 99.96% of the cases, the controller can add to collision risk reduction independently in about half of the cases and ATC alerts can further reduce the risk with 94%. The DRM-based approach provides a more complex view on the risk reducing potential of events and agents in the runway incursion scenario. On the one hand, control loops such as the ATC alerts that warn the controller and the controller that warns the pilots are working quite effectively with respect to the provision of new information. On the other hand, removing the alert systems or the controller out of the control loops lead to risk increases of only 6% and 22%, respectively. Thus the effectiveness of the introduction of an ATC alert system to reduce the accident risk was assessed much higher (by a factor 15) by the event sequence-based risk assessment versus the DRM-based risk assessment.

B. Events in ET and DRM

In this study we showed a variety of events and their probabilities in the ET- and DRM-based safety assessments. With respect to the values of the event probabilities, a key difference between the approaches is that in the ET-based analysis they are mostly input, whereas in the DRM-based analysis they are output. In particular, in the ET-based assessment the event probabilities were based on interviews with operational experts, who expressed their opinion on the possibilities to recognize and resolve conflicts at a particular stage. Only for the incident and accident events the ET-based assessment provides probability values as output. In contrast, in the DRM-based assessment the probability values of the shown events are all outcomes emerging from the MC simulations of the DRM, whether they refer to events for conflict recognition and/or resolution by agents or to aircraft collisions. In particular we obtained the event probabilities by evaluating a large number of MC simulation runs of the runway incursion scenario, with the variability in the performance of the agents as specified in the DRM. The thus obtained event probabilities could be related to the occurrence of collisions and to variables of agents (e.g. aircraft positions), and a variety of event combinations could be evaluated. As such a considerably more diverse overview of relations between events and collision risk could be obtained by the DRM-based approach.

In general, the values of the event probabilities in an ET-based approach depend on the specific context of the conflict scenario and on the data source chosen, such as expert interviews, incident databases or general literature. The values of the event probabilities obtained in the DRM-based approach also depend on the specific context of the conflict scenario and on the assumptions adopted in the DRM, such as model structural assumptions and parameter values. For the evaluation of the probability of the collision event, a method was applied in [18] to assess the sensitivity and uncertainty in the collision risk stemming from the assumptions adopted in the DRM. In principle, such methods [20] might also be applied for other events emerging in the Monte Carlo simulations of the DRM, such that a better insight would be obtained in the variance of the event probabilities in relation with the assumptions adopted in the DRM.

C. Risk elasticity for changes analysed by the DRM

The DRM-based approach supports analysis of the elasticity of the accident risk given changes in the operation, such as including/excluding particular agents (e.g. ATC alerts, controller). Key insights of the DRM results include:

1. The accident risk can be very elastic for changes in the organization of the operation;
2. The level of this risk elasticity 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 only from the totality of the performance and interactions of all human operators and technical systems in the operational context considered.

The first insight is well illustrated by the accident risk results of the DRM, which show that the risk would increase only by 6% without an ATC alert system specifying stopbar violation and runway incursion alerts, and the risk would only increase by 22% if the runway controller would be out of the loop and the conflict recognition and resolution is done only by the pilots of the involved aircraft. These small risk increases can be contrasted with large increases of about a factor 100 for organizations with one agent in the monitoring loop or a factor 500 without any agent in the monitoring loop, showing the potential for a large risk increase in the runway incursion scenario.

The second insight is well illustrated by the event analysis in the DRM. It follows that in about 94% of the runway incursion scenarios at least one of the alert types is active and in 61% of the scenarios the alert system warns the controller before (s)he has detected the conflict independently. Nevertheless, the risk increases by only 6% without an ATC alert system. The reasoning is even stronger for the

contribution of the controller. The model results indicate that the controller detects the conflict and warns the pilots in 99.3% of the cases and that in 96% and 78% of the cases the controller is able to warn the pilots flying of the taking-off or taxiing aircraft, respectively, before they have detected the conflict independently. In spite of this laudable performance of the controller in the model, the accident risk would only increase by 22% if the controller would not play a role at all in the resolution of the runway incursion scenario. It is only by considering the totality of the interactions between the agents and the variability in their performance in huge numbers of simulations that reveals such levels of accident risk elasticity for changes in the organization of the operation.

D. Risk elasticity for changes analysed by the ET

Insight in the risk elasticity for changes in the operation was not be obtained by the ET. The ET provides a representation of the reduction of the accident risk by various events and by agents related to these events. For instance, it follows from the presented results that the risk is reduced by a factor 2.1 due to own observation by the controller, by an additional factor 16 due to the ATC alert system and by a factor 33 (being the multiplication of the former two factors) due to the alert-supported controller. In the ET a change in the operation, such as leaving out an ATC alert system, would imply that alert-related events cannot occur. Assuming that the other event probabilities remain the same, this would lead to a risk increase by a factor 16, which would indicate a low risk elasticity for leaving out the alert system. However, the assumption that the other event probabilities remain the same must be characterized as a false one, since event probabilities are conditional upon earlier events in the ET. Nevertheless, having recognized the conditionality of the event probabilities, the ET approach as such does not provide ways to (re)assess their values. Thus the ET does not support analysis of the risk elasticity for changes in the organization of an operation; the best it can do is to represent events and event probabilities arrived at by a more effective analysis method.

E. Implications for real-time simulations

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, as was illustrated for the runway controller above, poses limitations on the safety conclusions that can be attained by limited series of simulations. In particular, the results indicate that if the numbers of simulations are not sufficient to estimate the accident risk of 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 the air traffic control domain, new concepts are regularly evaluated by real-time 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 real-time simulations are similar to what can be achieved in reality, e.g. a runway controller may handle about 25 to 40 aircraft per hour. Real-time 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 such real-time simulations occasionally conflict scenarios may be instantiated and the effectiveness of a controller to detect the conflict and warn pilots may be evaluated. Whereas it manifest that the numbers of conflict scenarios that can be evaluated in real-time simulations are far too small to evaluate safety up to the level of accident risk, the results of this study moreover indicate that results on the performance of human operators in such simulations say little about their contributions to safety. Consider, for instance, a hypothetical result of a real-time 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 present 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 considerable elasticity of the risk for the controller's performance.

As a way forward for using real-time simulations in safety assessment, aspects of the performance variability of human operators in safety relevant scenarios may be measured and such measurement results may be used to support the development of appropriate agent models in a DRM. Detailed discussion of such coupling of real-time simulations and dynamic risk modelling is out of the scope of this paper.

F. Implications for expert judgement

As a result of the conclusion that the level of safety 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, it also follows that assessing the contributions for prevailing accidents by interviewing single operators (pilots and controllers) and by judging their contributions, does not well account for the complexity of the interactions in conflict scenarios and thereby may well lead to inaccurate safety assessment results.

G. Implications for Resilience Engineering

Recent research in Resilience Engineering is in line with the notion of performance variability in complex socio-technical systems. Hollnagel [21] defines a resilient system as "A resilient system is able effectively to adjust its functioning prior to, during, or following changes and disturbances, so that it can continue to perform as required after a disruption or a major mishap, and in the presence of continuous stresses." Hollnagel stresses that the key term is adjust and that resilience is more than the ability to continue functioning in the presence of disturbances. We argue that as a basis for the engineering of such active adjustment processes, we should understand the risk elasticity for changes in a system or organization. In this paper we showed that an agent-based DRM provides an effective basis for risk elasticity analysis of an operation.

H. Implications for feedback to design

In general, a safety assessment aims to evaluate the level of safety of an operation, to relate this to a target level of safety and to provide feedback to the designers of the operation and its technical systems about potential safety-critical issues in the design and requirements for aspects in the design. The most prominent feedback to the design provided by event sequence analyses is the specification of requirements on the performance of components in the overall system, such that a target level of safety can be attained. Given the ingredients of an event sequence analysis, being events, their interrelations and their probabilities, the resulting types of requirements typically are minimum or maximum values for event probabilities, e.g. the probability of a non-availability of an alert system should be below a required maximum, or the probability that a controller reacts to an alert should be above a required minimum. Given the earlier discussed limitations of an event sequence-based analysis, the appropriateness of such requirements may however be questionable.

The feedback to the design provided by a DRM-based safety assessment can be more diverse than that of an event sequence analysis. As the DRM represents the dynamic performance of interacting agents rather than merely the occurrence of events, the potential feedback to the design reflects this larger variety of performance aspects. For instance, the performance of an alert system may include alert threshold and system availability settings, the performance of a pilot may include the timing of visual monitoring and the timing of reactions to controller calls. If the risk levels found by a DRM-based risk analysis are above a target level of safety, a sensitivity analysis may reveal the critical factors in the DRM that lead to the high risk levels, such that the design may be adapted in an effort to reduce the risk. If the risk levels are below a target level of safety, assumptions on the performance of agents are a suitable basis for the formulation of requirements on the performance of human operators and technical systems in the operational context.

VIII. Conclusion

In conclusion, considerably different results were obtained in the accident risk assessments of the complex socio-technical system involved in the runway incursion scenario. The Monte Carlo simulations of the DRM reveal that the accident risk can be very elastic for changes in the organization of the operation. The level of this risk elasticity 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 only from the totality of the performance and interactions of all human operators and technical systems in the operational context considered. These findings imply that judging the contributions of single human operators or technical systems for prevailing accidents may neglect the complexity of the interactions in socio-technical systems and thereby lead to inaccurate safety assessment results.

References

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

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

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

[4] Eurocontrol. Air navigation system safety assessment methodology. SAF.ET1.ST03.1000-MAN-01, edition 2.0, 2004

[5] EUROCAE. ED78A Guidelines for approval of the provision and use of ATS supported by data communication, 2000

[6] Leveson N. A new accident model for engineering safer systems. Safety Science 42:237-270, 2004

[7] Blom HAP, Bakker GJ, Blanker PJG, Daams J, Everdij MHC, Klompstra MB. Accident risk assessment for advanced air traffic management. In: Donohue GL and Zellweger AG (eds.), Air Transport Systems Engineering, AIAA, pp. 463-480, 2001

[8] Bedau MA. Weak emergence. In Tomberlin J (ed.), Philosophical Perspectives: Mind, Causation, and World, Vol 11, pp. 375-399, Blackwell, Malden (MA), USA, 1997

[9] Corning PA. The re-emergence of "emergence": A venerable concept in search of a theory. Complexity 7(6):18-30, 2002

[10] Dulac N, Owens B, Leveson N et al. Demonstration of a new dynamic approach to risk analysis for NASA's constellation program. MIT SCRL Report, March 2007

[11] Blom HAP, Stroeve SH, Scholte JJ, De Jong HH. Accident risk analysis benchmarking Monte Carlo simulation versus event sequences. Proceedings Third International Conference on Research in Air Transportation (ICRAT 2008), Fairfax (VA), USA, 1-4 June 2008

[12] Stroeve SH, Blom HAP, De Jong HH, Scholte JJ. Comparison of accident risk assessment by event sequence analysis versus Monte Carlo simulation. Eurocontrol Safety R&D Seminar, Southampton, UK, 22-24 October 2008

[13] ICAO. Manual on the prevention of runway incursions. International Civil Aviation Organization, Doc 9870 AN/463, first edition, 2007

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

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

[16] 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 10L/19R under good visibility conditions. National Aerospace Laboratory NLR, memorandum LL-2001-017, Amsterdam, The Netherlands, 2001

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

[18] Stroeve SH, Blom HAP, Bakker GJ. Systemic accident risk assessment in air traffic by Monte Carlo simulation. Safety Science 47:238-249, 2009

[19] Everdij MHC, Klompstra MB, Blom HAP, Klein Obbink B. Compositional specification of a multi-agent system by stochastically and dynamically coloured Petri nets. In: Blom HAP, Lygeros J (eds.), Stochastic Hybrid Systems, LNCIS 337, Springer-Verlag, pp. 325-350, 2006

[20] Everdij MHC, Blom HAP, Stroeve SH. Structured assessment of bias and uncertainty in Monte Carlo simulated accident risk. Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management, May 14-18 2006, New Orleans, USA, 2006

[21] Hollnagel E. The four cornerstones of resilience engineering. In Nemeth CP, Hollnagel E, Dekker S (eds.). Resilience Engineering Perspectives, Volume 2: Preparation and restoration, pp. 117-134, Ashgate, Aldershot, UK, 2009