A Conceptual and Computational Framework for Identifying and Predicting the Performance of Novel Airspace Concept of Operations

Final Report

Contract Number: NNX07AO62A Submitted to

NASA Ames Research Center
Moffett Field, CA 94035

Technical Monitor: Paul F. Borchers
NASA/FAA NTX Research Station
13800 FAA Road Fort Worth, Texas 76155-2104

for the period
November 1, 2007 - September 30, 2011

Vitali Volovoi, Ph.D., Principal Investigator
School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0150
(Tel) 404-894-9811
(Fax) 404-894-2760
vitali.volovoi@ae.gatech.edu

December 29, 2011

SUMMARY OF ACCOMPLISHMENTS

Improving the total performance of the air traffic management (ATM) system in terms of capacity, safety, efficiency, and flexibility rely on dramatic system-wide transformations as well as improvements in the performance of individual communication, navigation, and surveillance (CNS) systems. The evaluation of specific performance levels of ATM system requires a robust structural modeling and simulation framework that can evaluate emergent system-wide performance arising from the behavior of individual system components including human operators. This project facilitates establishing a conceptual and computational framework to identify and predict the usefulness of specific level and potential groupings of applications and system performance capabilities in important Next Generation Air Transportation System (NextGen) scenarios.

In particular, the focus is on understanding the safety implications of any changes to the operations of National Airspace, which presents a host of challenges due to the highly complex and coupled nature of this system. The system's current state might not be the most efficient, but its safety features are well established and grounded in years of relatively successful experience. Technological advances in both aircraft and the supporting infrastructure promise great improvements in efficiency, but their successful implementation necessitates appropriate procedural changes. As a result, a comprehensive assessment of the hazards associated with the introduction of new technologies must involve modeling of interactions among aircraft behavior, supporting infrastructure, and the operational procedures. The traditional approach to risk assessment focuses on the occurrence of relevant events regardless of their relative timing. The likelihoods of those events are computed externally, usually by means of physics-based simulations, which, while increasingly realistic in capturing physical phenomena, are limited to describing only few relevant interactions to keep the overall complexity tractable. In contrast, the present approach relies on an intermediate layer of analysis that has enough fidelity to capture time-dependent coupling among relevant entities of the system, while being compact enough to track a large number of those relevant entities simultaneously. The utility of Stochastic Petri Nets (SPNs) in the role of this intermediate layer has been demonstrated using the application of the nested analysis to the conflict resolution between the merging flows of air traffic that uses an optimized profile descent approach. SPNs are coupled with agent-based simulation, and the efficiency of the merging procedures and their sensitivity to wind conditions and the traffic patterns are analyzed.

In addition, a novel analytical procedure is developed for evaluating the risks associated with collision avoidance systems. This procedure relies on augmentation of the state space, so that discrete (non-homogeneous) Markov chains can be constructed and solved in closed form to evaluate system reliability. The presented method provides an efficient means for modeling dependent subsystems without explicit state-space representation of individual components by semi-inverting Markov chains for non-repairable portions of the model and using the obtained transition rates in the full Markov model. The developed procedure is general, but a specific application for the Advanced Airspace Concept (AAC) is used to demonstrate the method's capabilities and to compare the results with those of published Monte Carlo simulations. The advantage of the presented method is not only computational efficiency and higher precision, but increased transparency of the contributing risk factors, which is particularly beneficial given the uncertainty about the input parameters and the associated need of sensitivity studies.

1 Introduction

The joint research effort of NASA and FAA's Next Generation, NextGen, aims at a total overhaul of the National Airspace System (NAS), which would enable it to meet the expected increase in air travel demand for the coming decades. Critical to the success of the envisioned overhaul is the ability to maintain safety at levels that are at least as good as the current ones, which implies the need for an improved ability to recover from unforeseen disturbances and to mitigate their large-scale impact or propagation by means of a more graceful degradation. From a safety perspective, the proposed changes to the NAS can be generally grouped into the following three categories:

1. New concepts and technologies pertaining to the vehicles themselves, which usually improve performance and/or reduce costs, but might also provide fundamentally new functionalities *e.g.*, Unmanned Aerial Vehicles (UAVs);

2. New capabilities for supporting infrastructure, including Communication, Navigation, and Surveillance (CNS);

3. Novel air-traffic procedures aimed generally at increasing the throughput (*e.g.*, spiral landing for CESTOL aircraft [8]), but also at reducing environmental impact and increasing the cost-efficiency of operations (*e.g.*, optimized-profile descent, also referred to as continuous descent approach [10]), respectively.

These changes are tightly coupled, as the procedures should accommodate new vehicles as well as take advantage of the supporting infrastructure. One of the recognized top-level changes envisioned as a part of NextGen is an increase in the level of decentralization within the NAS. This trend is consistent with the developments in other complex distributed network systems (*e.g.*, the Internet and the National Power Grid). One of the advantages of the increasingly autonomous concepts of operation that rely on peer-to-peer coordination is the enhanced system robustness with respect to disturbances (*e.g.*, extended bad weather conditions) and other off-nominal operations. However, these changes require several technological advancements in terms of automation, communication, and coordination among the various systems and entities comprising the NAS. These changes will lead to a more sophisticated infrastructure entailing more complexity at the component and overall-system levels. The resulting impact on systems' reliability and traffic safety is far from clear, and requires thorough examination. In fact, higher complexity is generally linked to a larger number of possible failure modes or safety hazards, whose resolution and mitigation are often pursued via redundancy that in turn can lead to yet more complexity as a part of a vicious cycle[30].

In the presence of such a large-scale technological and procedural restructuring envisioned for the NAS, the assessment of traffic safety for the new framework can become a rather daunting task. While the procedures for assessing the risk impact of a single new vehicle technology or concept is relatively well developed, the investigation of combined effects of changes to the NAS and their interactions in the presence of multiple uncertainty sources requires new analytical approaches capable of addressing the problems of escalating complexity, increasing coupling and decreasing knowledge [7]. Unexpected failure modes and hazardous traffic conditions may originate from flight procedures being executed in the presence of under-predicted disturbances within the NAS, or may be caused by a specific vehicle's malfunction affecting its neighboring air traffic and not being attended to promptly.

The traditional approach to risk assessment involves a fault-tree based representation of hazards [15]. Fault trees rely on Boolean (static) logic and evaluate the probability of the occurrence of relevant events regardless of their relative timing. Event trees are a vital part of probabilistic risk assessment due to the fact that they account for the consequences of the relative order of events [33]. However, they do not provide any means for evaluating the likelihood of the events occurring in the particular order, which as a result must be obtained from external sources, usually by means of physics-based simulations. Physics-based simulations (including agent-based simulations) are increasingly realistic in capturing particular physical phenomena, but the depth of the analysis comes at the expense of its breadth, and so they are limited to only few relevant interactions. In this context, this research study is aimed at providing a scheme for risk assessment for highly coupled changes within the NAS in general, and at the identification and quantification of traffic hazards associated with novel procedural concepts in particular, as the latter problem exhibits such coupled behavior. Specifically, two novel techniques are developed and tested: First, a hierarchical (nested) approach is investigated as a means to perform risk analysis; the methodology is illustrated on the application to the Optimized Profile Descent (OPD) procedure as implemented at the Los Angeles airport. Second, a hybrid

method for combined use of fault trees and Markov chains is developed for systems that have both renewable and non-renewable components. This method relies on using fault trees for evaluating the states dynamics of the non-renewable portion of the system, evaluating the corresponding state transition rates, and, finally utilizing those transition rates in the Markov model of the whole system. This hybrid method has been successfully applied to Advanced Airspace Concept (AAC).

2 Multi-layered risk analysis

The risk-based analysis for air traffic investigated in this work consists of a two-step bottom-up process. The first step focuses on the exploration, generation, and simulation of hazardous scenarios potentially leading to unsafe air travel conditions that could arise as a consequence of the adoption of new flight patterns. The second step of the methodology entails statistical processing of the simulation data aimed at quantifying the level of risk associated with those hazards.

2.1 Agent-Based Modeling

The NAS is a rather complex and heterogeneous system, in which different types of airborne and ground--based entities need to interact and collaborate efficiently, often under tight schedules. Each of those entities (*e.g.*, pilots, air traffic controllers, airlines, dispatchers, etc.) follows its own set of behavioral rules and is characterized by its own dynamics (*i.e.*, action/reaction times and decision-making processes for human agents, and kinematic/physics-based behavior for machine agents), and operates in a highly dimensional and constantly changing environment. Hence, system dynamic modeling in a classical sense may quickly become infeasible for such large and diversified architectures. The use of agent-based modeling and simulation has been found to be well suited to describe these types of systems, as this approach takes advantage of the concept of decentralization by focusing on each agent's microscopic behavior rather than attempting to macroscopically model the entire framework's dynamics, which instead is left to emerge from the agent-to--agent interactions. The literature offers several examples of its usage, in the engineering domain [27, 21] as well as in other research areas such as biology and sociology [6, 28]. Specific to the NAS, simulation tools like IMPACT (Intelligent agent-based Model for Policy Analysis of Collaborative Traffic flow management), SAMPLE (Situation Assessment Model of Pilot-in-the-Loop Evaluation), FACET (Future Air Traffic Management Concepts Evaluation Tool), or ACES (Airspace Concept Evaluation System) have been developed to evaluate various aspects of the air traffic ranging from human operator performance to the flow of NAS flights and response to external disturbances (*e.g.*, weather) to negotiation schemes for conflict resolution in the context of free flight [9, 14, 24].

The agent-based approach is used to model a certain portion of the NAS and its elements, namely aircraft trajectories and the actions of pilots and ground operators, so that hazard scenarios can be explored and significant metrics extracted as a function of relevant traffic parameters and conditions (*e.g.*, aircraft separation or weather disturbances). This information is then utilized to construct a statistical abstraction of the simulated scenarios via Stochastic Petri Nets (SPNs). While the considered scenarios can be fully implemented using agent-based simulation, SPNs provide a higher level of abstraction, as well as the visual means to explore a wider range of interactions and scenarios. Preliminary comparison of the results from SPNs and agent-based simulation are presented, and the sources of the differences will be investigated in the follow up study to ensure that adequate accuracy is maintained.

2.2 Stochastic Petri Nets

In the proposed risk-assessment methodology, the outputs from agent-based simulations, along with any available information about the expected likelihoods of the events triggering the explored scenarios (*e.g.*, failures of hardware or weather/traffic patterns), are meant as inputs into system-level hybrid risk models that combine discrete events related to failures, blunders, and other relevant events, with continuous parameters related to temporal-spatial representations of aircraft. To this end, SPNs enable the structured combination of the effects of discrete logic relationships and the descriptions of the timing of relevant events. A particular extension of Petri Nets has been designed for modeling system risk and reliability through special objects called aging tokens that facilitate the compression of continuous parameters present at the lower level of analysis into "age" (effectively related to the timing of events, including the history of previous events and their timing). SPN-based modeling permits the abstraction of the simulation environment via a more

compressed representation of continuous space while still including a wide range of system and human errors associated with hazardous scenarios. Petri Nets were introduced by Anton Petri in 1962[31]. The SPN framework focuses on modeling the components' states that comprise the system, so that the state of the overall system can be inferred from the states of its components.

In Petri Nets, a component (denoted by a small filled circle called token) can be in any of its possible states (denoted with larger hollow circles). The advantage of such a representation is that it allows representing the whole system implicitly, thus potentially mitigating the state-space dimensional explosion. To enable component-level descriptions of the state-space of the whole system, the interdependence among the possible states of individual components is described within a network using unidirectional transitions, where a change of state is called "firing" of a transition. Importantly for describing the dependence among the components behavior, the firing of a transition can only occur when it is "enabled," *i.e.* certain conditions ought to be satisfied. The original Petri net has not included the concept of time, so that an enabled transition may only fire immediately. SPNs represent extensions of the the original Petri nets. SPNs belong to a subset of so-called non-autonomous Petri Nets [11]; they were introduced about twenty years later [34, 26, 25] and are of particular relevance to the modeling of system reliability. SPNs introduce delays between the enabling and firing of a transition, where those delays are transition attributes that can be either absent (an immediate transition), deterministic, or sampled from a given distribution (stochastic). If the transitions are memory-less (*e.g.*, following exponential distribution) SPNs can be used as a pre-processor for Markov chain analysis, but a discrete event (*e.g.*, Monte Carlo) simulation can also be used to solve SPNs directly without any restriction on the type of the delays used[12]. Effective system modeling using SPNs involves its decomposition into a set of relevant entities, where each entity does not necessarily represent a physical component of the system, as it might, for example, describe a phase of operation or an environmental condition. Due to their flexibility in terms of modeling both discrete logic and continuous states, SPNs have been considered particularly useful in the context of modeling air transportation systems[4]. However, in the proposed work the emphasis is on the use of SPNs that are decoupled from 4-D continuous representation. Such a decoupling provides a critical advantage in analyzing the main factors in safety assessment by keeping the models significantly more compact, and the input information portable, thus enabling the use of heterogeneous sources of inputs. It must be noted that the use of Agent-Based Modeling to provide a probabilistic representation of individual events that might lead (or on the contrary, prevent) hazardous situations is analogous to the use of physics-of-failure models [29] that are increasingly popular in the reliability field.

2.3 Agent-Based Implementation

Depicted in Figures 1-2 are simulation results from the agent-based model, developed using the freeware software NETLOGO [37], where two air traffic flows are taken into consideration as they merge into the same final approach path. Included within the simulation together with the aircraft are also pilots and air traffic controller (ATC). As the aircraft's schedules to the metering points are disrupted, the air traffic controller monitors their relative distances and issues any necessary command should there be a separation violation. The ATC's commands are executed by the pilots according to some response time delay aimed at emulating their reaction.

The agent-based model represents a two-stream air-traffic and relies on the following assumptions:

- **Environment**
  The air traffic being modeled consists of a portion of the East Feeder Sector for LAX airport, where only the air traffic fluxes going through the feeder points "GRAMM" and "LAADY" were taken into consideration (Fig. 3). Furthermore, each airplane is assumed to fly an Optimized Profile Descent (OPD) computed via the approach developed by Ren *et. al* [32].

  Traffic uncertainty lies in the time it takes each single vehicle to reach the metering point in the presence of wind. In the absence of wind, all aircraft starting from the same point will reach the runway along the same path in the same amount of time; that is not the case if different wind conditions are experienced by each aircraft. Even though the wind is assumed to be fully compensated, aircraft will be off-course with respect to their no-wind schedule, which then introduces uncertainty in terms of their relative spacing.

  Separation in the presence of multiple traffic flows is estimated by projecting each of them onto a common vertical plane (thus reducing the dimensionality of the analysis) where its computation consists in estimating the time of arrival at the merging point. In case of conflict, two evasive maneuvers are

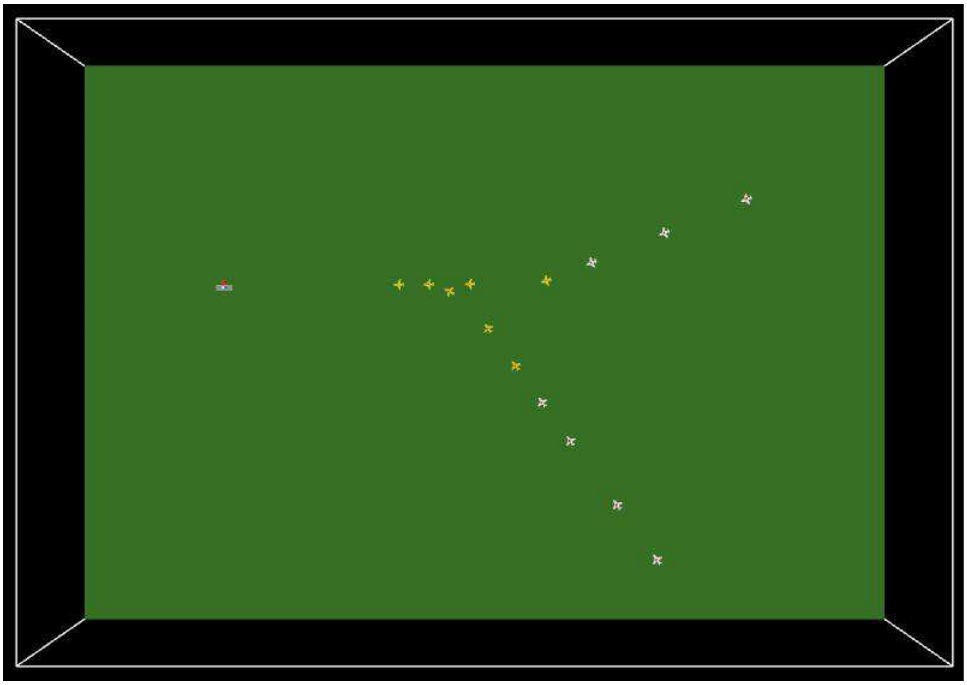


Figure 1: Merging of two flows of aircraft performing OPD procedures.

possible to reestablish the proper separation: speed change, and a re-vectoring of the vehicle from one traffic flow to the other. This research focuses on the effectiveness of the speed-changing maneuvers.

The model includes two main sources of uncertainty, namely associated with wind conditions and human behavior, even though failure modes or disruptions of other nature could be incorporated, at least in their first-order approximation (*e.g.*, an avionics malfunction causing a shift in flight schedule). As the traffic volume increases, the negative impact upon the safety of these unknowns is expected to increase as the available window for corrective intervention may reduce in the presence of more intense operations. On the one hand, the use of agent-based modeling permits the stressing of the airspace to its limit to unveil critical or irrecoverable scenarios; on the other hand, the usage of stochastic Petri Nets is intended as a statistical means capable of summarizing the occurrence of the off-nominal observed events (*e.g.*, violation of separation, and success of recovery from an unsafe situation) in terms of likelihoods and probabilities of interest through which safety-critical conditions can be assessed and their risk level quantified.

- **Management of Space Violation and Conflict Resolution**
  An underlying assumption for the feeder sector is that the feeder points and the merging point for the various traffic streams are placed in space in such a way that airplanes flying from their respective feeder points will arrive at the merging point at the same time if moving with the same inertial velocity profile. Therefore, in nominal conditions, no conflict should arise if the aircraft are properly spaced within each traffic stream and the streams are properly synchronized, even though that is often not the case due to the differences in the environmental conditions and the kinematics of each aircraft.
  Conflicts are checked by ATC for every vehicle arriving at a feeder point; at that time any aircraft trailing behind, either on the same traffic stream or on another one, is issued a -20 kts or -40 kts velocity-change if its horizontal distance from its feeder point is within 5 nm or 2.5 nm, respectively.

  This model captures only a portion of the actual procedures that ensure the proper spacing upon merging. A simple conflict resolution logic consisting of one single maneuver being issued per conflict was investigated with only two maneuvers able to be executed. In reality, vectoring, vehicle acceleration, creating "holes" for aircraft arriving from other directions, and complex combination of several maneuvers are employed. However, the model can still provide useful insights as shown below, and can be considered as a building block in constructing a more complete representation of the procedures.

  The two traffic streams are considered to be totally independent, but within each flux an appropriate

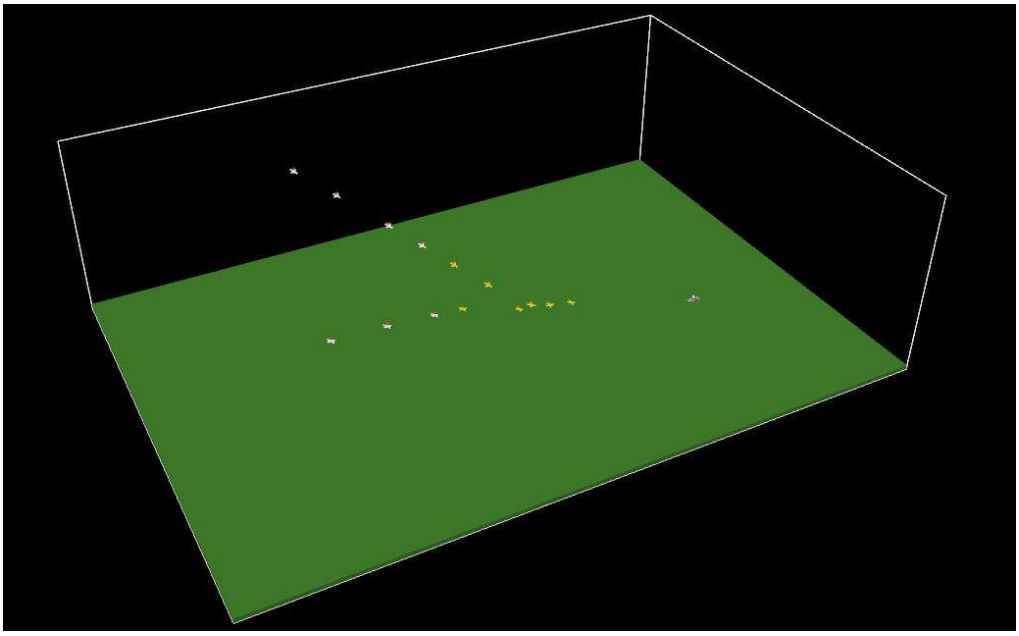


Figure 2: Merging of two flows of aircraft performing OPD procedures.

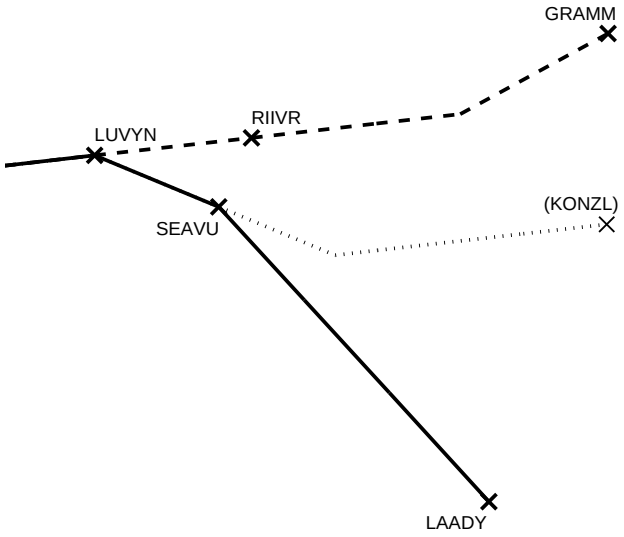


Figure 3: Layout of the East Feeder Sector at LAX.

space separation is maintained, so that a new vehicle is introduced at least Δt_o seconds after the preceding aircraft's appearance within the simulation. This separation $\Delta t_{in-flux\,separation}$ was chosen to follow an exponential distribution:

$$\Delta t_{in-flux\,separation} \sim \Delta t_o + E(\mu_E) \;\; [sec] \tag{1}$$

where the value Δt_o is to be assigned based on the aircraft density in each traffic stream, and whose effectiveness will depend on the kinematics of each vehicle and its neighbors.
Lastly, due to the nature of the OPD trajectory profiles (namely altitude as a function of time), vertical separation was not used as part of the metric describing space violation. Instead, conflicts were characterized solely by means of the horizontal separation between airplanes or between an airplane and the feeder point of its traffic stream (although the developed agent-based model provides means to evaluate the vertical separation as well).

- **Modeling of Personnel**
  Two types of personnel were modeled, namely the air traffic controllers (ATC) and the pilots, where their time-varying interactions introduces additional uncertainty, the impact of which will be more or less significant depending on the traffic volume. The nature of the OPD approach is such that it requires little communication between the vehicles and the ground operators in normal (nominal) conditions; hence interaction has been limited only to the issuing of corrective commands.

  – The ATC is in charge of monitoring the air traffic and issuing corrective-maneuver commands when necessary, as mentioned above. At this stage, monitoring for conflicts was assumed to be continuous in time, hence the issuing of commands to the pilots was supposed to occur without any human delay.

  – The pilots' performance, instead, was modeled by means of an execution delay τ_p described via a log-normal distribution $L(\mu_L = 7, \sigma_L = 3)$, not to exceed 12 seconds with respect to the issuing/reception of the command from ATC.

- **Modeling of the Wind Impact on Vehicle's Motion**
  Given the nature of the OPD procedures (*i.e.*, based on similar unconstrained geometric trajectories, but traveling with a different flight/travel time by each vehicle depending on its inertial velocity profile) and the layout of the feeder sector, conflicts (especially those occurring at the merging point of various traffic streams) may be characterized primarily in terms of the vehicles' travel time $t_{_F}$ to fly between two locations. Hence, instead of modeling a space-time wind profile $\mathbf{W}(x, y, z, t)$ and account locally for the wind, a macroscopic effect was incorporated based on $t_{_F}$ under various wind conditions. The relationship between flight time and wind was established by generating 36 OPD trajectories under different wind scenarios $\mathbf{W}(x, y, z, t)_j$ and by computing the corresponding t_{F_j} $(j = 1, \ldots, 36)$ to fly from any of the feeder points to the merging point. Depicted in Figure 4 are the data set as well as the cubic interpolation being constructed between t_{F_j} and the wind along the trajectory, evaluated at the feeder points themselves. Due to the interest primarily in tail/head wind effects, this approximation proved to be sufficiently accurate and more straightforward than direct manipulation of multi-variate random variables. All aircraft on a specific traffic flux were assumed to fly along the same geometric trajectory (with given nominal $\bar{t}_{_F}$), but with velocity profiles scaled through the ratio $\tilde{t}_{_F}/\bar{t}_{_F}$ between a randomly generated travel time $\tilde{t}_{_F}$, corresponding to a random wind, and the nominal flight time $\bar{t}_{_F}$ for that trajectory. It is to be noted that correlation between the winds experienced by neighboring vehicles was neglected at this stage of the analysis. Both time- and space-based wind correlation can be introduced into the model, but since this correlation is expected to be positive, the use of uncorrelated wind provides a conservative approach to the estimation of the wind impact on vehicle separation.

  Finally, random wind conditions were generated using the "NCEP Reanalysis 1" data for the years from 2007 to 2009 [16], illustrated in Figure 5 together with the fitting log-logistic probability density functions.

The setup described above was used to generate all the necessary statistical information for the event time parameters that are used as inputs into the SPNs. The reasoning behind certain assumptions consists in capturing the first-order limitations and effectiveness of a conflict-resolution logic focused on reduced intervention by the ATC's, *i.e.*, aimed at minimizing the number of corrective maneuvers to be ordered to

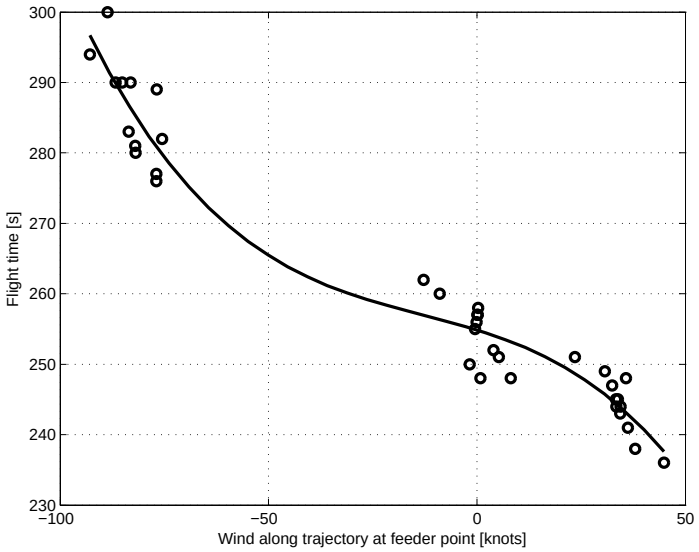
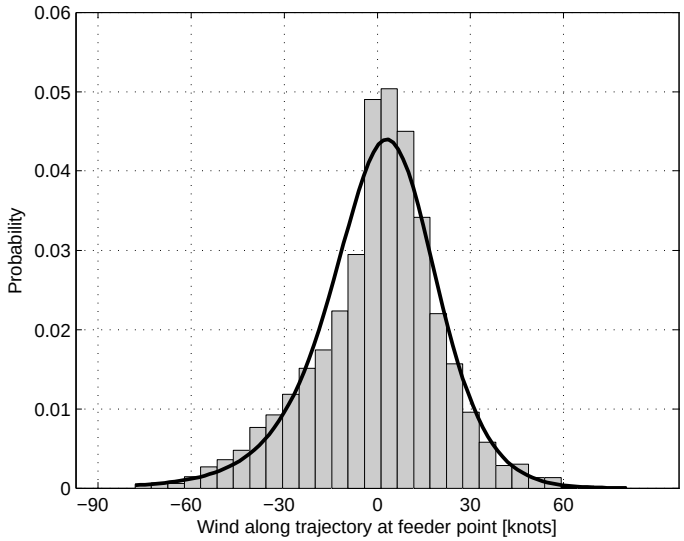
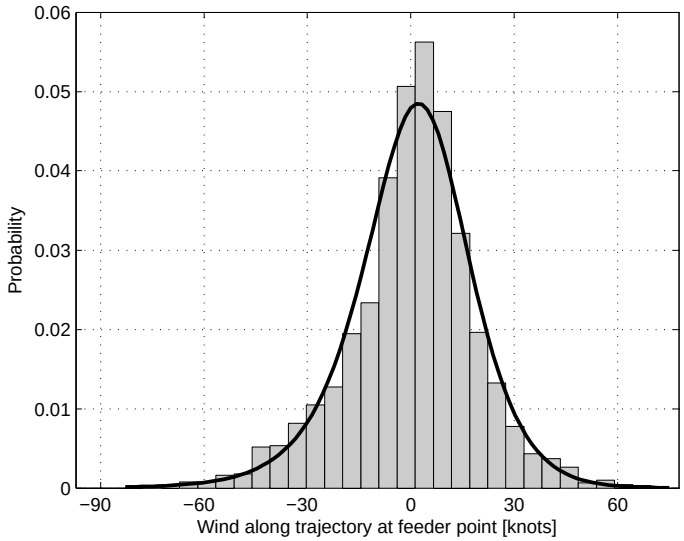


Figure 4: Flight time from the feeder point (GRAMM or LAADY) to the merging point (LUVYN).

the aircraft to address and resolve their conflicts. As the air traffic density is forecast to increase due to higher demand, human-operator performance and workload as well as system automation will incur limitations. On the one hand, the higher density of air traffic will cause the windows of opportunity for ATC's and pilots to reduce, thus requesting more precision the first time around as well as more assistance from *ad hoc* automation systems. On the other hand, system automation may suffer a drastic nonlinear escalation in its own complexity, which could be necessary to handle a much larger number of conflict scenarios, either currently occurring and being tackled, or potentially arising at a later time because of disruptions in the nominal traffic, either associated with ordered corrective maneuvers or with its inherent uncertainty (*e.g.*, weather and varying wind conditions, and differences in velocity profiles among the aircraft).



(a) At feeder point GRAMM



(b) At feeder point LAADY

Figure 5: "NCEP Reanalysis 1" wind data and fitting log-logistic probability density functions. [16]

2.4 SPN model implementation

The corresponding schematics of the SPN model are shown in Figure 6. Places (hollow circles) depict possible states (positions) of an aircraft, while the aircraft are represented by tokens (small solid circles). Transitions (rectangles) govern the timing of "firing" or moving tokens between places (*i.e.*, how the aircraft changes its states). Here the state representation is selected to be as small as possible while fully capturing the relevant behavior and dependencies. The top of the model represents flux 1 (moving from the right to the left) passing through the GRAMM feeder point, while the bottom corresponds to LAADY flux 2. Each token representing an aircraft is assigned one of the three colors (0 – green, 1 – yellow, 2 – red) based on whether ATC issues a command for this aircraft to slow down. The delays associated with the transitions "travel 1" and "travel 2" are color-dependent with the corresponding distributions provided by the agent-based simulation. The colors are assigned in the following manner: when a token in GRAMM place (that is, an aircraft that has just reached the feeder point GRAMM as depicted in Figure 6), two transitions are enabled for the LAADY flux by means of so-called "enablers" (which are opposite to the more commonly used SPN inhibitors and denoted by an arc originating in the place and terminating at the transition with a solid circle).

As a result, if there is a token in the input place for one of those transitions (in Figure 6, this is true for place "-20 knots," *e.g.*, an aircraft is located somewhere between 2.5 miles and 5 miles away from fix LAADY), then this token is fired through the transition "+1," which deposits the token into the same place but changes its color from green to yellow (note that there is a fixed small delay ϵ associated with this transition while the token in place GRAMM stays for 2ϵ ensuring that this change occurs). The color-changing transition is color-sensitive as well (the firing occurs only for the green color). Similarly (not depicted), if there is a token in place "-40 knots" (*e.g.*, an aircraft is located less than 2.5 miles from fix LAADY) then this token is fired through the transition "+2," which deposits the token into the same place while changing its color from green to red. The reciprocal policies are implemented for the aircraft that need to be slowed down within the GRAMM flux based on their position when another aircraft passes the LAADY feeder point. The delay associated with the "spacing" transition determine the degree of spacing violation: statistics are collected of the frequency of two tokens located in "Merging" place at the same time, so the longer delay of "spacing" transition corresponds to the larger spacing. Effectively, here the spacing is evaluated in the time domain, which is obtained from the space domain based on the speed profile of the vehicles.

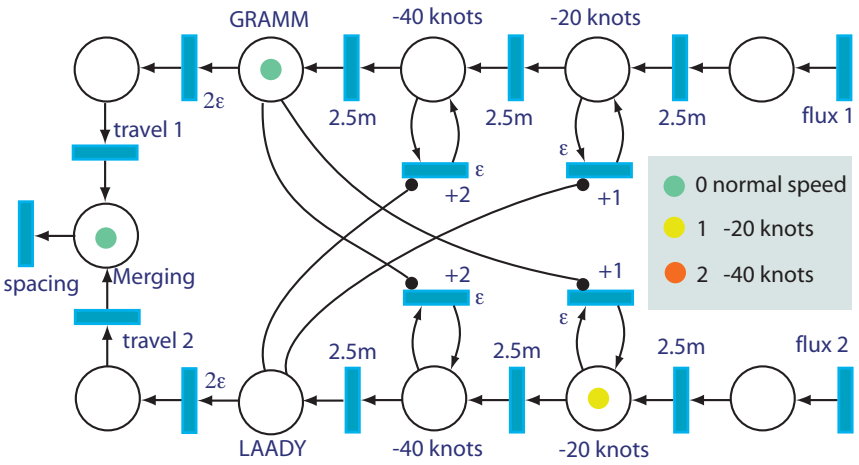


Figure 6: SPN model describing the merging two fluxes of aircraft (GRAMM and LAADY).

2.5 Test scenario: OPD-based traffic at LAX

Only a portion of the actual procedures at LAX has been modeled that related to the merging of two fluxes and coordinating the spacing based on three possible actions ("do nothing," "slow down by 20 knots," and "slow down by 40 knots"). However, even this simplified model provides some interesting insights into the efficiency of the resulting spacing, with the SPN model representing a convenient and fast means to explore the sensitivity to various parameters. The results presented are developed based on running 10,000 Monte Carlo simulations for 10,000 seconds of operation for each set of parameters, which takes about 15 seconds of total simulation using a MacAir laptop. The mean value of time separation between aircraft for each flux was considered to be fixed at 180 seconds, so that each simulation involved about 110 airplanes. This value was selected to provide a representative overall density of traffic given the fact that the actual stream of traffic at LAX airport involves the merging of more than two fluxes.

While the mean value was unchanged, the minimum spacing within each flux was varied to investigate the "domino" effect, *i.e.*, when the slowing down of an aircraft aimed at improving the spacing upon merging leads to a separation conflict with the following aircraft from the same flux. Figure 7 shows the dependence of the frequency of spacing conflicts as a function of minimum separation within each flux for 2.5 nm and 3 nm separation thresholds. No wind uncertainty is considered in this figure. One can observe that coupling or "domino" effects emerge when the spacing in each stream falls below 95-100 seconds. As a reference point, it might be noted that the traveling time for the last 5 miles prior to the tie point is about 45 seconds, so the threshold for the emergence of the "domino" effect is slightly more than 10 miles. One can also note that delaying the subsequent aircraft by introducing additional slow-down commands if the preceding aircraft was ordered to slow down does not constitute a viable solution to this problem. Indeed, despite the increased complexity those additional commands can provide only a partial mitigation of the situation, as the effectiveness of coordinating the conflicts between the fluxes will be reduced: if the matching aircraft from another flux is in conflict, then the speed adjustment to that aircraft will be less productive. Therefore, the threshold value will be likely to stay the same, while the number of conflicts might be slightly reduced. To put the number of observed conflicts in the context of the overall conflict resolution efficiency, it is useful to note that with no coordinated maneuvers executed (*i.e.*, when the "natural" merging takes place) there are about 16.5 and 20 spacing conflicts per 100 aircraft for 2.5 nm and 3 nm thresholds, respectively, and those values practically do not change within the range of considered minimal spacing for individual aircraft.

Next, the impact of the wind uncertainty on the frequency of spacing conflicts is investigated when aircraft are sufficiently separated within each stream to avoid the domino effect. For that purpose, the minimum separation in each flux is considered to be 120 seconds. Figures 8, 9 show the results of SPN simulation for

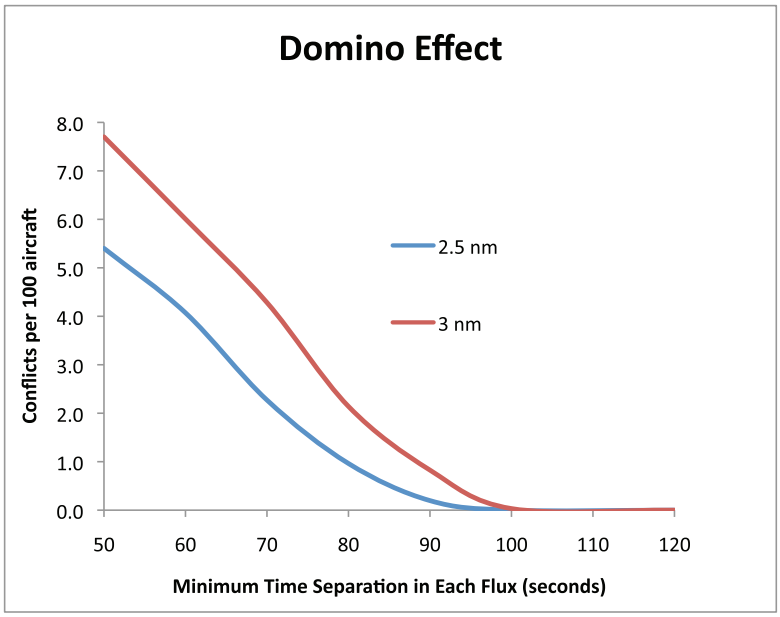


Figure 7: Results of SPN simulation showing the frequency of spacing conflicts as a function of minimum separation within each flux. 2.5 nm and 3 nm separation are shown; no wind uncertainty is considered.

frequency of spacing conflicts as a function of the uncertainty associated with the travel time from the tie point to the merging point that is measured in terms of a standard deviation in seconds. Curves correspond to the travel times that follow lognormal distributions (with the standard deviation on the horizontal axis calculated as an average of six travel time distributions (two fix points and three speed regimes). Diamonds denote the values obtained from the "global" agent-based simulation (that is, the whole procedure is implemented in the agent-based simulation). It can be observed that agent-based simulation predicts a slightly higher number of conflicts than SPN and the difference is statistically significant, and the source of the difference is currently investigated. Since slowing down by each 20 knots "buys" about 20 seconds of extra travel time, it is clear that those maneuvers could not assure a separation of more than 3 nm (which correspond to about 40 seconds of traveling time right after merging), and the results shown support this intuitive observation. As the uncertainty of the traveling increases, the efficiency of the maneuvers to ensure 2.5 nm and 3 nm starts to decrease when the value of the standard deviation reaches a threshold of about 4 seconds, which is larger than the uncertainty estimated based on the observed wind conditions. If (positive) spatial and temporal wind correlation is taken into account, the value of the threshold is expected to increase, so the presented model can be considered conservative. Therefore, for the considered set of parameters one can conclude that the uncertainty of the traveling time between the fix point and the merging point due to the wind variability does not preclude successful conflict resolution within the range of its application (*i.e.*, if vectoring or other maneuvers are not required).

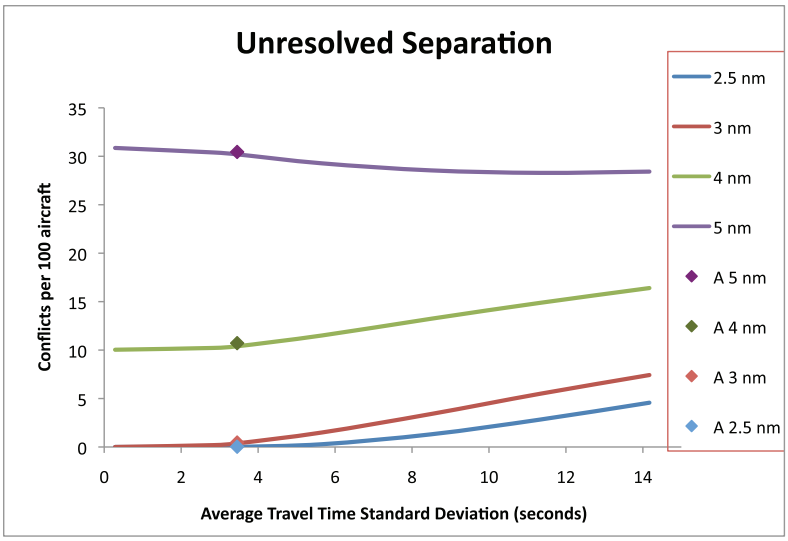


Figure 8: Results of SPN simulation showing the frequency of spacing conflicts as a function of the uncertainty associated with the travel time from the feeder point to the merging point that is measured in terms of a standard deviation in seconds. Diamonds represent the results from the "global" agent-based simulation.

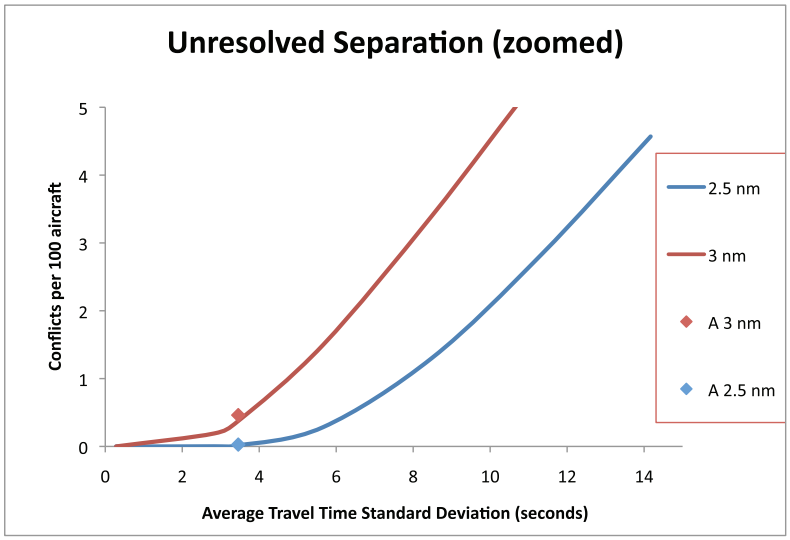


Figure 9: Results of SPN simulation showing the frequency of spacing conflicts as a function of the uncertainty associated with the travel time from the feeder point to the merging point that is measured in terms of a standard deviation in seconds. Diamonds represent the results from the "global" agent-based simulation. Only 2.5 nm and 3 nm spacing are shown.

3 Safety of Collision avoidance systems

Collision avoidance systems are critical for the safety of airspace, especially given the expectations of a significant increase in operation density in the future. Another related but distinct factor that increases the importance of collision avoidance systems is the issue of integration of Unmanned Air Vehicles (UAVs) into the national airspace. The targeted levels of safety as a function of air traffic density have been used to provide a baseline characterization (*i.e.,* without taking into account any mitigation action) of the collision risk [36], thus providing clear motivation for mitigation strategies in UAV operations. Quantification of the overall (system-level) safety impact of collision-avoidance systems requires understanding of the relevant interactions among the various layers of protection against collisions, as well as the frequencies and patterns of encounters that can lead to collisions. The latter (collision encounter problem) recently has been extensively investigated [18, 17], but the challenges of the former problem are also rather significant, and this paper is devoted to addressing those challenges. One can break the overall problem of estimating the risk of collision into three steps:

1. Determining the conflict frequency;
2. Given the conflict, determining the chances of resolving it by a deployed collision avoidance system (the focus of this this paper);
3. Determining collision chances, given Near Mid-Air Collision (NMAC), which is the failure of the collision avoidance system to resolve a conflict.

It is a common and generally a reasonable assumption that the calculations involved in these three steps are mutually independent (at least in the first approximation).

From the system reliability and safety modeling standpoint, a collision-avoidance system relies on time redundancy, as there are several consecutive attempts to detect and resolve a conflict. This time redundancy is supplemented by functional redundancy as well, as the time before the conflict is separated into distinct phases (layers) where the conflict resolution task is assigned to distinct subsystems. This functional separation is motivated by the increased urgency of the task combined with less uncertainty about the conflict. So, as a general rule, as time progresses, conflict resolution should be simpler (less complex) in order to facilitate reliability, and can be simpler, as it deals with less uncertainty. In addition, increasing the diversity of the protective layers provides some protection against common-cause failures that can defeat the intended redundancy. Combining structural and time redundancy is not unique to collision avoidance, and is well recognized as providing the more efficient means of protection than each type of redundancy alone in other applications, such as in designing fault-tolerant computer systems that would negate the effects of transient faults [19]. There are several generic methods for modeling the system reliability of time-redundant systems using semi-Markov processes [22] or universal generating function technique [23]. However, neither method is directly applicable to the modeling of automated collision avoidance systems, as neither allows the presence of accumulated permanent faults. Indeed, permanent faults violate the basic assumption of semi-Markov processes: that the transition from a state is fully determined by the current state and the holding time in that state.

Fault-tree analysis (i.e., analysis based on static boolean algebra) can provide important initial insights [2, 15], but this approach is insufficient for capturing the dynamic interactions that are critical for a more detailed analysis. In general terms, those interactions stem from periodic detection of potential conflicts. On the one hand, this detection becomes more efficient as time progresses (as the uncertainty about the trajectories decreases); on the other hand, there is a potential for accumulation of failures that hinder successful detection and resolution of those conflicts. Fault trees are based on Boolean algebra operations of probabilities and rely on the assumption of independence of basic events. In the case of sequential attempts of conflict resolution, this assumption does not hold, as the basic events are conditioned by the events that occurred at the previous steps. In particular, the probabilities of failure to resolve the conflict at each attempt cannot be simply multiplied, as this would lead to meaningless results.

If these dynamic interactions are confined to a single layer of protection, then a decoupled (hierarchical) analysis is possible, as advocated in the context of sense-and-avoid systems [1]: an inner loop that includes a collision encounter model and relies on Monte Carlo simulation combined with an "outer loop" analysis based on fault trees. If different layers share common failure modes, neglecting this coupling in the fault-tree analysis can lead to nonconservative risk estimates. In order to account for this coupling the scope of Monte Carlo simulation can be extended to encompass several layers of conflict avoidance. However, this leads to

to the increased level of complexity of the simulation models and simultaneously increases the demand for the number of the simulation runs needed to capture rare events. While importance sampling can provide an increase in the convergence rate [3], this improvement is problem-dependent, with the most improvements obtained where the rare event is structured as a combination of several (less rare) events.

The proposed approach relies on providing a structured dynamic model, so that the time dependency is explicitly captured in a (nonuniform) Markov state-space representation. In what follows, the developed procedure is applied to a specific application, Advanced Airspace Concept (AAC) [13], with the issues regarding the generality of the procedure addressed as appropriate. The selection of the application is motivated by the availability of the detailed description of an automated avoidance system safety model that has been conducted using Monte Carlo simulation [5, 35]. The goal of this paper *is not* to evaluate the assumptions made for AAC [5, 35], but to demonstrate an analytical method that can successfully capture the dynamic interactions without the need of lengthy Monte Carlo simulations for systems with the mixture of non-repairable or Markov (when state transitions only depend on the current state and time) components. This approach can be contrasted to the standard analytical reliability methods, where either a non-repairable or Markov assumption is made for all the components (and no mixture of the two is allowed). The developed procedure provides a mapping between the system-level risks and the relevant parameters of a collision-avoidance system, thus enabling sensitivity studies that are important due to the uncertainty about those input parameters.

3.1 General formulation using Markov Discrete Space

Successful conflict resolution requires, on the one hand, that the appropriate equipment is operational, and on the other hand, that trajectory generation and conflict detection is successful as well. There are several layers of conflict avoidance, each invoked in sequence as time progresses. Importantly, there is an overlap in terms of the equipment used by each layer, so that the permanent failures of layers are not independent. Furthermore, within each layer, several attempts are made to resolve the conflict. In the case of AAC, there are three such layers: Autoresolver (AR), Tactical Separation-Assured Flight Environment (TSAFE), and Traffic alert Collision Avoidance System (TCAS), which are engaged in sequence: first AR is engaged (from 8-20 minutes until 3 minutes before the conflict), followed by TSAFE (from 3 minutes until 1 minute), and TCAS (at 1 minute before the conflict). There is an additional (final) level of safety (visual avoidance by pilots) that is applied last, and its efficiency is provided by the fraction of conflicts that were unresolved by the first three layers but resolved by the fourth layer (so its evaluation is decoupled from the evaluation of the first three layers).

In order to minimize the complexity of the system reliability analysis, it is important to identify subsystems (modules) that are as large as possible without obscuring the coupling among the subsystems. Specifically, the common components that make the performance of the three layers dependent require a separate treatment. In the case of AAC [5], the following coupling mechanisms are identified:

1. Mode S transponder on each aircraft. Its functionality is critical to all three conflict-avoidance systems, and we denote its probability of failure as $F_T(t)$ when a transponder system on either of the two aircraft fails. Here T denotes the transponder subsystem (corresponding to the transponders of both aircraft), and t is time elapsed from the moment AR engaged. Based on the assumptions made in Ref. [5], when such a failure occurs, the entire collision avoidance system fails.

2. Resolution Delivery (RD). There are shared components between AR and TSAFE contributing to RD functionaility; however, the chances of the loss of RD functionality during AR phase are negligible. Indeed, the chances of RD functionality in AR configuration for the whole flight can be calculated as follows (see figure Fig. 12 and Appendix for the values of involved components):

$$P_{AR\_RD} = [1 - (1 - VDL2)(1 - RR)(1 - FMS)]^2 \, VC^2 \approx 5.76 \times 10^{-15} \tag{2}$$

Noting that the overall risk of system failure is on the order of $\sim 1 \times 10^{-6}$ — 1×10^{-9}, the failure of this branch of the fault tree can be neglected. Qualitatively, this is explained by the fact that RD functionality has a quadruple redundancy at the component level for AR phase. As a result, the failure of RD functionality needs to be considered only for TSAFE, and therefore this source of coupling between different phases can be neglected. Such prescreening of the contributing risk factors (at the fault-tree rather than at the component level) is important to simplify the analysis.

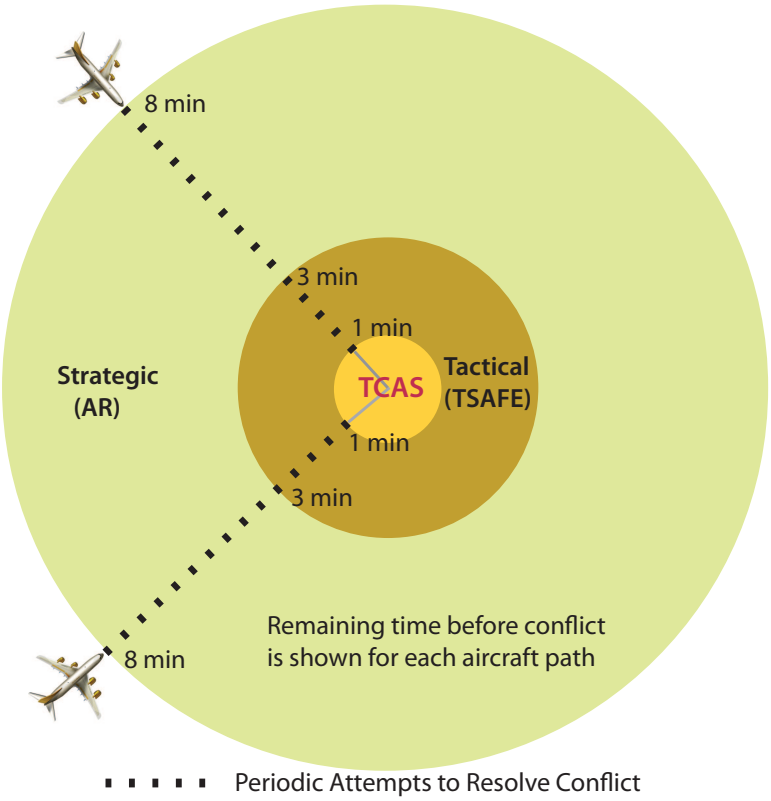


Figure 10: Three layers of conflict resolution in AAC (the fourth layer, visual avoidance, is not depicted).

3. The speaker that announces the resolution to the pilot in both the TSAFE and TCAS systems. We will denote this subsystem as K, so that its failure occurs when a speaker system on either of the two aircraft fails (here the letter S is not used for this system to avoid confusion with the success state that is introduced below). The purpose of separating the functionality of this subsystem stems from the fact that while TSAFE can operate with one of the speakers down, TCAS cannot (here we follow the assumptions made in Ref. [5] for consistency, although an argument can be made that TCAS can facilitate collision avoidance even if only one of the aircraft reacts). In order to make this distinction, we introduce separate states during TSAFE operation (B_k if both speakers are operating, and E_k otherwise).

4. In order for either AR or TSAFE to operate, both aircraft need to be located. Effectively, this functionality is common to AR and TSAFE. The mode S transponder is part of this subsystem, but since we treat it separately, we introduce the probability of failure to locate both aircraft due to failure of components other than the mode S transponder, $F_L(t)$. The fault tree that corresponds to subsystem L is shown in Fig. 11. The subsystem of AR that delivers functionality, which is related to neither location nor transponder, is denoted as A and the corresponding probability of failure as $F_A(t)$. Similarly, the subsystem of TSAFE that delivers functionality, which is related to neither location nor transponder, is denoted as B and the corresponding probability of failure as $F_B(t)$.

After evaluating the fault tree for AR, taking into account that subsystems L and T are treated separately, and noting that branch shown in Fig. 12 is of negligible importance (the chances of failure of this branch for the whole flight is 5.76×10^{-15}; see above discussion of the common modes), we can conclude that subsystem A effectively consists only of ACR (see Appendix for the definitions of individual components). As a result, the following distinct subsystems are considered: A, B, K, L, T. This is the required level of granularity to capture all the coupling from the equipment perspective.

At the beginning, we assume that at $t_1 = 0$ there is the first possibility of identifying the conflict (this is the initial state A_1) - the state is $T = 8$ min away from the conflict. Furthermore, at that point there are no failures in any of the subsystems [5, 35]. At every time step t_k, $k = 1 \ldots n$, an attempt is made to resolve

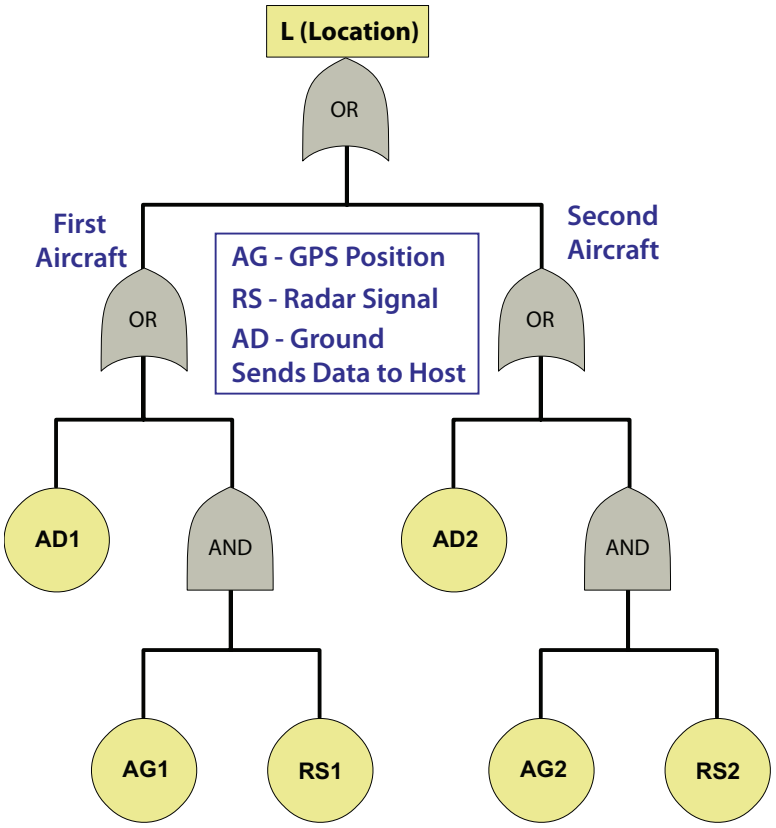


Figure 11: Fault tree for location function of both aircraft (subsystem L)

the conflict. In the considered example, $n = 15$ and $t_{k+1} - t_k = 0.5$ min for $k = 1 \dots n-1$. In general n corresponds to the total number of opportunities to resolve the conflict for all layers (phases). As a result, we consider the following sets of intermediate states $A_1 \dots A_{10}$ for AR, $B_1 \dots B_4, E_1 \dots E_4$ for TSAFE, and C for TCAS. In addition, there are states F and S for failed and successful conflict conflict resolution, respectively. Therefore, we have the following total enumeration of the states:

$$X = [A_1 \dots A_{10}, B_{11} \dots B_{15}, E_{11} \dots E_{15}, C, S, F] \,, \tag{3}$$

Next, individual layers (phases) are modeled.

3.2 Modeling the first phase (AR)

A successful conflict resolution from A_1 (that is transition to S) implies that the following three conditions are met:

1. Trajectories for both aircraft has been successfully generated (for each aircraft the probability is $1 - P_{FPT}$, see Appendix for definition of this and other component probabilities)
2. Conflict detection occurs with the probability $P_D(t)$
3. AR functions properly

As a result, the corresponding transition probability can be calculated as follows:

$$\alpha(k) = (1 - P_{FPT})^2 P_D(k) \tag{4}$$

Note that this transition probability is applicable for all time steps k, since being in state A_k (or B_k) implies that all the relevant subsystems operate at time t_k. Full state transitions for A_k are depicted in Figure. 14. In order to calculate all transition rates, we need to consider the following order of priorities: if no successful

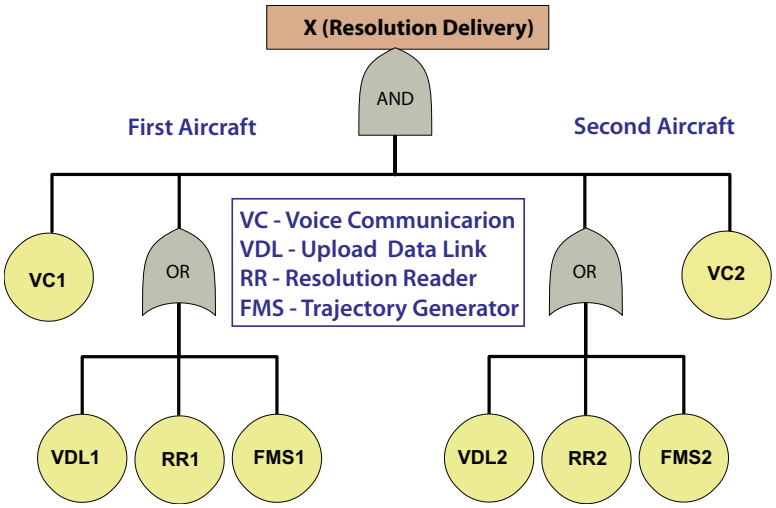


Figure 12: Fault Tree for Resolution Delivery (RD) of AR that has a negligible failure probability

resolution of the conflict takes place at time t_k, and transponder T fails by the time t_{k+1}, then (regardless of failures of other subsystems) the transition to state F occurs. The corresponding chances are determined by the discrete version of the hazard rate:

$$h_T(k) = \frac{F_T(k+1) - F_T(k)}{1 - F_T(k)} \tag{5}$$

Hazard rates for other sub-systems are defined in the same fashion. Only if the transponder failure has not occurred can other transitions take place: first the transition to state C (if subsystem L failed); next, to state B_1 if subsystem A failed. Finally, the transition to state A_{k+1} is complementary to all other transitions (if none of the failures occurred):

$$\gamma(k) = (1 - \alpha(k))h_T(k) \tag{6}$$
$$\lambda(k) = (1 - \alpha(k))(1 - h_T(k))h_L(k) \tag{7}$$
$$\mu(k) = (1 - \alpha(k))(1 - h_T(k))(1 - h_L(k))h_A(k) \tag{8}$$
$$\beta(k) = (1 - \alpha(k))(1 - h_T(k))(1 - h_L(k))(1 - h_A(k)) \tag{9}$$

With individual subsystems, failures are calculated as follows:

$$F_T(k) = 2F_{AB}(t_k) - F^2_{AB}(t_k)) \tag{10}$$
$$F_{L1}(k) = F_{AG}(t_k)F_{RS}(t_k) + F_{AD} - F_{AD}(t_k)F_{AG}(t_k)F_{RS}(t_k) \tag{11}$$
$$F_L = 2F_{L1}(k) - F^2_{L1}(k) \tag{12}$$
$$F_A(k) = F_{ACR}(t_k) \tag{13}$$

It must be noted that during the first phase (i.e., the operation of AR), additional failures can occur that will cause the system to transition to a more degraded configuration. The following transitions are therefore provided for B_1 and C:

$$\pi(k) = h_T(k) \tag{14}$$
$$\sigma(k) = (1 - \pi(k))h_L(k) \tag{15}$$

Note that the same transition probability $\pi(k)$ is used for both states (at this point, we do not evaluate the failure of subsystem B). At this phase all considered subsystems are independent. However, this is not the case for the following phases, as described next.

3.3 Modeling dependent subsystems

In order to model the transition between the two layers of ACC, one must consider dependent subsystems, since subsystems K and B are not mutually independent (the former is a part of the latter). As a result, we

have to resort to conditional probabilities to properly evaluate transitions. Since both of those systems have not been previously evaluated, any failures of those two subsystems between time $t = 0$ and $t = 5$ minutes must be included. Effectively, we need to consider four disjoint (mutually exclusive) events separately:

1. At time $t_{11} = 5$ minutes, both speakers subsystem K and subsystem B failed $P_{B \bigcap K}(11)$. The system should transition to state F.

2. At the time subsystem B has failed, but the speakers are intact $P_{B \bigcap \bar{K}}(11)$ (the bar over the K indicates that the failure has not occurred), the system should transition to state C.

3. The speakers subsystem K has failed, but system B is operational: $P_{\bar{B} \bigcap K}(11)$. The system transitions to state E_1 (if no other failures occurred).

4. Neither subsystem has failed $P_{\bar{B} \bigcap \bar{K}}(11)$. The system transitions to state B_1 (if no other failures occurred).

We can note that the combined probability of those four events is unity. A brute-force approach would involve explicit consideration of the possible states of all components of these two subsystems. Since there are five components involved, one would have to deal with $2^5 = 32$ states. When both systems are down, the occurrences of further failures are irrelevant, so there are actually fewer distinct states, and symmetry considerations can be used to further reduce the state space. Still, this approach is obviously prone to state-space explosion, so it has poor scalability for problems with a larger number of components. Instead, a compressed state-space representation of conditional states for two subsystems is employed, as described next.

The key consideration (that to the best of the authors' knowledge has not been previously described in the literature) is the procedure of inverting the Markov sub-model that corresponds to permanent failures in order to obtain the transition rates that are used in the full model. In the traditional setup of Markov processes, the transition rates are known, and the probability of being in a particular state is evaluated as a function of time. In contrast, in the current procedure, Boolean algebra is used to calculate the probabilities of relevant states for those subsystems that are subject to permanent failures (the subsystems are non-repairable). Next, this information is used to infer the transition rates for this subsystem, and finally, those transition rates are utilized to construct the transition rates for the whole system (this procedure is demonstrated below for the TSAFE phase).

The fault tree for subsystem B is shown in Figure 13 (for brevity, the dependence on t is omitted on the right-hand side, but all expressions are function of t). Using this fault tree, we can calculate the following probabilities:

$$P_{\bar{B} \bigcap \bar{K}}(t) = (1 - F_{TCR})(1 - F_S)^2(1 - F_{RR}^2) \tag{16}$$

$$P_{B \bigcap \bar{K}}(t) = (1 - F_S)^2 \left[F_{TCR} + (1 - F_{TCR})F_{RR}^2\right] \tag{17}$$

$$P_{B \bigcap K}(t) = F_S^2 + 2F_S(1 - F_S)\left[F_{TCR} + (1 - F_{TCR})F_{RR}\right] \tag{18}$$

$$P_{\bar{B} \bigcap K}(t) = 2F_S(1 - F_S)(1 - F_{TCR})(1 - F_{RR}) \tag{19}$$

It can be checked that these four probabilities sum up to unity.

The corresponding transitions are shown in Fig. 15, while the probabilities of those transitions are derived next. First we derive the transition to the failed state:

$$\gamma(10) = (1 - \alpha(10))(h_T(10) + (1 - h_T(10))\left[P_{B \bigcap K}(11) + P_{\bar{B} \bigcap K}(11)h_L(10)\right]) \tag{20}$$

Here the γ transition includes an additional term corresponding to the failures of both B and K ($P_{B \bigcap K}$); indeed the failure of B implies that TSAFE is not operational, while the speakers' failure (subsystem K) implies that TCAS will not be operational either. Note that the failure accumulation occurs up to the next step (in the case of incremental change, like in all other transitions, the hazard rate also contains the value from the next step, but the values from all previous steps are subtracted).

Next we evaluate the transition to state C:

$$\begin{aligned}\lambda(10) = (1 - \alpha(10))\{&(1 - h_T(10))[P_{B \bigcap \bar{K}}(11) + P_{\bar{B} \bigcap \bar{K}}(11)h_L(10) + \\ &+ P_{SOAR}\left(1 - h_L(10)\right)\left(P_{\bar{B} \bigcap K}(11) + P_{\bar{B} \bigcap \bar{K}}(11)\right)]\}\end{aligned} \tag{21}$$

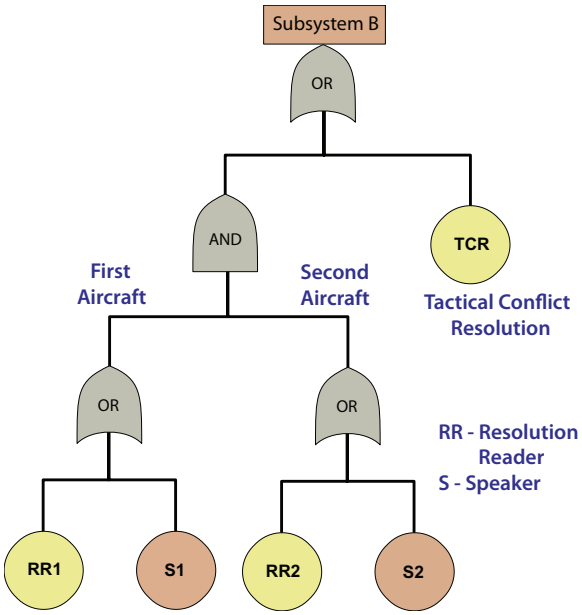


Figure 13: Fault tree for subsystem B that provides functionality for TSAFE

In comparison to the previous steps, we need to take into account the possibility of failed speakers that would preclude operation of TCAS. Note also that we take into account the possibility that the override from AR to TSAFE will fail with the probability P_{SOAR}, even if all the systems function properly, so the corresponding terms are added to $\lambda(10)$.

The remaining two transitions from the state A_{10} have the following expressions:

$$\mu(10) = (1 - P_{SOAR})\,(1 - \alpha(10))\,(1 - h_T(10))\,P_{\bar{B}\bigcap K}(11)\,(1 - h_L(10)) \tag{22}$$

$$\beta(10) = (1 - P_{SOAR})\,(1 - \alpha(10))\,(1 - h_T(10))\,P_{\bar{B}\bigcap \bar{K}}(11)\,(1 - h_L(10)) \tag{23}$$

Similarly, we can assess the transitions for B_1 and C (unlike A_{10}, the probability of remaining in those states is not zero, but arcs that point to the same state are omitted for clarity):

$$\pi(10) = \frac{\gamma(10)}{(1 - \alpha(10))} \tag{24}$$

$$\sigma(10) = \frac{\lambda(10)}{(1 - \alpha(10))} \tag{25}$$

$$\rho(10) = \frac{\mu(10)}{(1 - \alpha(10))} \tag{26}$$

$$\phi(10) = h_T(10) + (1 - h_T(10))\left[P_{B\bigcap K}(11) + P_{\bar{B}\bigcap K}(11)\right] \tag{27}$$

Finally, transitions during the operation of TSAFE can be evaluated (see Figure 16). Here, the evaluation of discrete transition rates that describe the dependent states of subsystems B and K requires some special attention. Transitions among these four states can be described using a (nonuniform in time) discrete Markov chain (Figure 17). The following three balance equations can be written (noting that the fourth one is redundant):

$$P_{B\bigcap \bar{K}}(t_{k+1}) = h_1(k)P_{\bar{B}\bigcap \bar{K}}(t_k) + (1 - h_5(k))\,P_{B\bigcap \bar{K}}(t_k) \tag{28}$$

$$P_{\bar{B}\bigcap K}(t_{k+1}) = h_2(k)P_{\bar{B}\bigcap \bar{K}}(t_k) + (1 - h_4(k))\,P_{\bar{B}\bigcap K}(t_k) \tag{29}$$

$$P_{B\bigcap K}(t_{k+1}) = P_{B\bigcap K}(t_k) + h_3(k)P_{\bar{B}\bigcap \bar{K}}(t_k) + h_4(k)P_{\bar{B}\bigcap K}(t_k) + h_5(k)P_{B\bigcap \bar{K}}(t_k) \tag{30}$$

These equations could be used to calculate the discrete transition rates, but there are five unknowns, $h_1(k) \ldots h_5(k)$, and only three equations. However, $h_4(k)$ and $h_5(k)$ can be calculated in a relatively straightforward fashion: indeed, for the former, we can note that in the state $\bar{B}\bigcap K$, all redundancy is depleted,

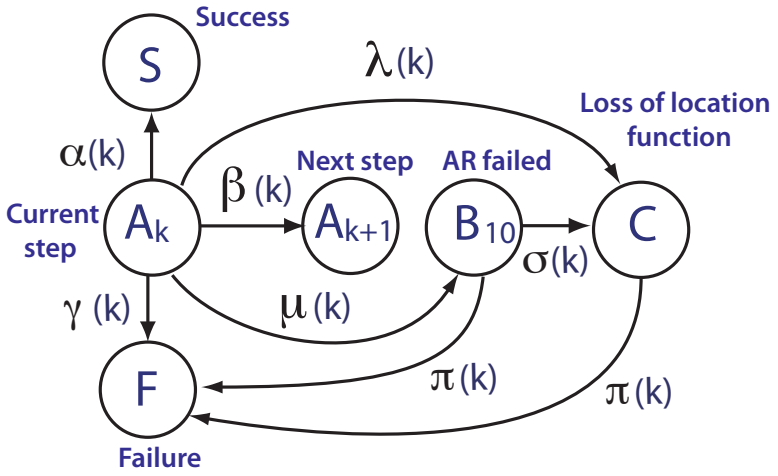


Figure 14: State transition for AR phase $k = 1, \ldots 9$

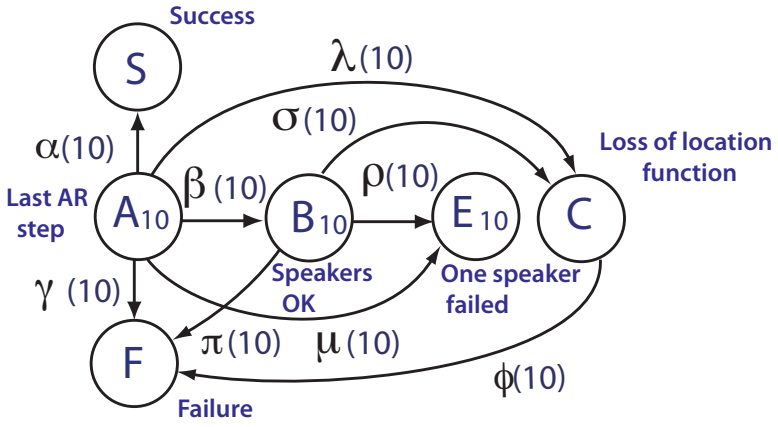


Figure 15: State transition for AR phase $k = 10$ (the last time when AR is invoked)

and there is exactly one speaker and one RR operational. Similarly, for the latter, we note that the transition from the state $\bar{B} \bigcap K$ depends only on the state of the speakers (both of which are operational). The corresponding cumulative distribution functions can be therefore expressed as follows (dependence on t_k is implied on the right hand side for brevity):

$$F_4(k) = F_{TCR} + (1 - F_{TCR})\,[F_S + (1 - F_S)F_{RR}] \tag{31}$$

$$F_5(k) = 2F_S - F_S^2 \tag{32}$$

In general, there might be more than one possible configuration for the degraded state, and in this case the total transition rates have to be weighted in accordance with the probability of each degraded configuration.

Now, using Eq. 5 for expressions Eqs. 31 and 32, we can obtain expressions for $h_4(k)$ and $h_5(k)$. Finally, we can substitute those expressions into Eqs. 28, 29, and 30 to obtain $h_1(k)$, $h_2(k)$, and $h_3(k)$, respectively.

Next, we note that the transition rate to success from states B_k and E_k are still given by $\alpha(k)$ (see Eq. 4) $k = 11 \ldots 14$. The rest of the transitions for state B_k are described next.

$$\gamma(k) = (1 - \alpha(k))\,[h_T(k) + (1 - h_T(k))(h_3(k) + h_2(k)h_L(k))] \tag{33}$$

$$\lambda(k) = (1 - \alpha(k))(1 - h_T(k))\,[h_1(k) + h_L(k)\,(1 - h_1(k) - h_2(k) - h_3(k))] \tag{34}$$

$$\mu(k) = (1 - \alpha(k))(1 - h_T(k))h_2(k)\,(1 - h_L(k)) \tag{35}$$

$$\beta(k) = (1 - \alpha(k))(1 - h_T(k))\,(1 - h_L(k))\,[1 - h_1(k) - h_2(k) - h_3(k)] \tag{36}$$

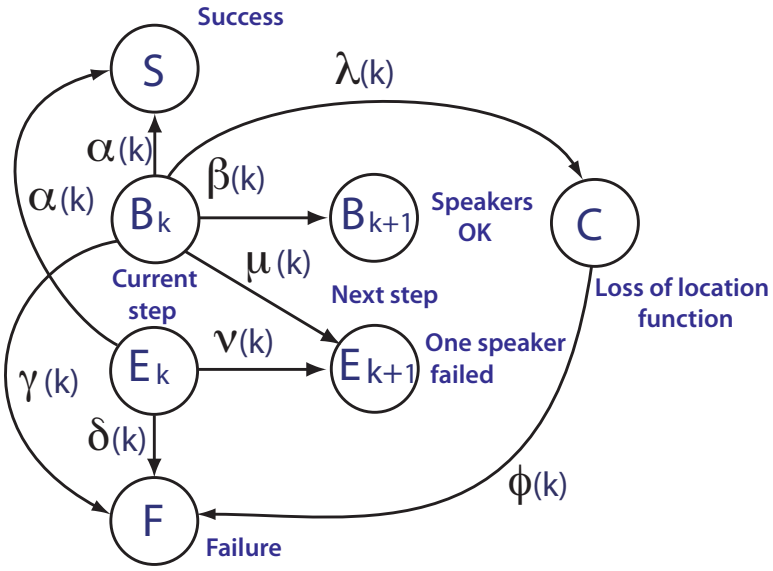


Figure 16: State transition for TSAFE phase $k = 11, 15$

Similarly, the rest of the transitions are provided for states E_k and C:

$$\delta(k) = (1 - \alpha(k)) \left[h_T(k) + (1 - h_T(k))(h_L(k) + (1 - h_L(k))h_4(k))\right] \tag{37}$$

$$\nu(k) = 1 - \alpha(k) - \delta(k) \tag{38}$$

$$\phi(k) = h_T(k) + (1 - h_T(k))h_5(k) \tag{39}$$

At the last step of TSAFE, we can note that E_{15} implies failure (and so the corresponding probability needs to be added to the failed state; see below), while B_{15} is the same as C. After this addition, the $P(C)$ represents the chances that TCAS will be required (and both speakers and both transponders are operational). Therefore the final expression for the probability of failure of AAC is

$$P_{fail} = P(F) + P(E_{15}) + [P(C) + P(B_{15})] \left[P_{TCOTS} + (1 - P_{TCOTS})(2P_{OTHER} - P_{OTHER}^2)\right] \tag{40}$$

3.4 Additional considerations

As mentioned in the introduction, the system safety structure and parameters of the AAC model [5] were used as an illustration and a reference point for constructing the corresponding analytical model, so this work should not be considered as an endorsement of that model (and as a result, the endorsement of the corresponding risk estimation). However, this can be considered a reasonable starting point for constructing meaningful models, which can be effectively used for building safety cases for particular collision avoidance implementation, along with the requirements for the performance characteristics of the individual components. This goes beyond the scope of the present paper, but several initial observations can be pointed out:

1. Probability of detection: in the paper, the probability of detection is based on squaring the probability of not deviating by half of the distance [5] . Based on purely geometric considerations, the probability of detection is actually significantly higher. In the extreme case of a head-on collision, one can derive an analytical formula using normal cross-track error distribution and the Euclidean difference. It can be ascertained that the currently used formulae provide a conservative estimate. In addition, there is a possibility of a correlation between the errors (common bias) both in time and between the two aircraft in conflict. The former would lead to the decrease in the probability of successful detection. Importantly, introducing different functions of probability detection as a function of time would not interfere with the structure of the system-level model, and the procedure described in the paper should be still applicable.

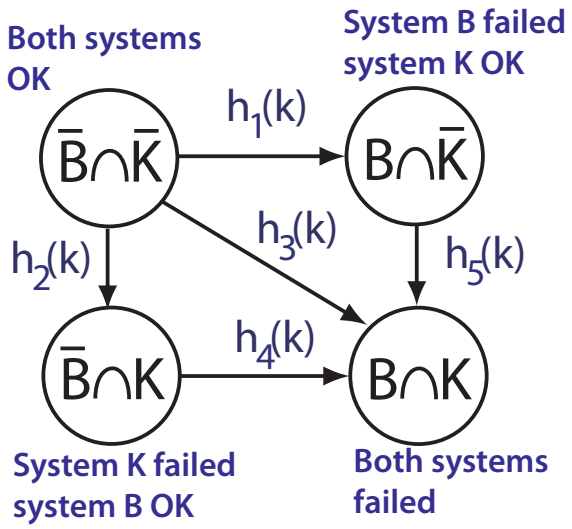


Figure 17: State transition for subsystems B and K

2. Commission error: for example, a "false positive" situation where the system assumes that there is a conflict, although there is actually is none; this is potentially an important consideration due to the reduction of the time available for correcting the error. Similarly, resolution of an existing conflict can be executed incorrectly. The issue is related to so-called "Byzantine fault tolerance"[20]. A more detailed modeling of conflict resolution is needed to estimate the associated probabilities, and this is a different failure mode that must be modeled separately.

3. At time $t_1 = 0$, all systems are assumed to function properly; this needs to be revisited, as the risks during the recovery process need to be estimated. This recovery process also requires a separate model.

4. The motivation behind the use of exponential transformation TCAS [5] is questionable (see the discussion in the Appendix regarding P_{OTHER}). This consideration will only impact the values of the used parameter and not the structure of the model.

5. The failure rate of ADS-B Mode S transponder, AB (see Appendix) is obtained from Ref. [15], with the source citing value that is one order of magnitude higher, and is based on an exposure of 20 minutes (and not two hours). Apparently, this reduction of the failure rate is due to a credit for some redundancy. Due to the importance of this parameter on the overall failure of the system, this issue should be further investigated. However, this consideration will only impact the values of the used parameter and not the structure of the model (unless the redundancy of the transponders has to be modeled explicitly).

4 Research Conclusions

The application of Stochastic Petri Nets as an intermediate fidelity-level analysis for safety assessment is described. SPNs use discrete state space, but maintain continuous-time representation thus providing a means to capture time-dependent interactions among various entities of the National Airspace System. Inputs to SPN models consist of the appropriate time distribution of relevant events that can be obtained from physics-based simulations. In the considered example, agent-based simulations have been used to obtain the needed distributions. Specifically, the uncertainty of traveling time between the feeder points and merging point was taken into account to evaluate the efficiency of the coordination of two fluxes of aircraft that follow optimized profile descents. The frequency of space violations are estimated as a function of relevant parameters. Traditional risk assessments that rely on the use of fault trees are not capable of capturing the timing of events, so this type of analysis would have to be conducted within a single physics-based simulation, or in this context, an agent-based simulation. Preliminary validation of the results of SPN against the agent-based simulation that includes all the relevant logic (and is orders of magnitude slower and much more difficult

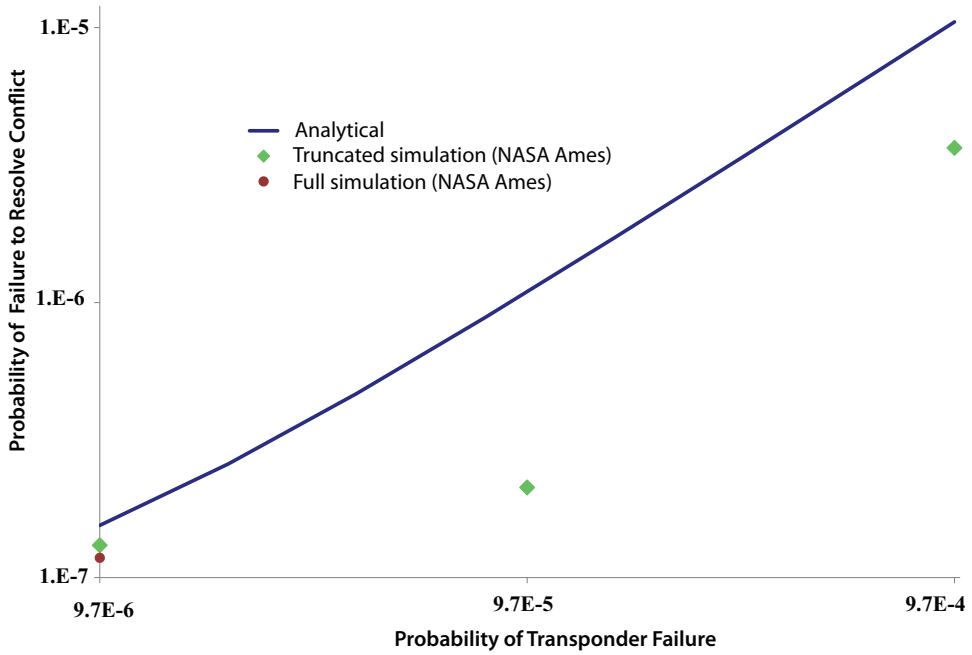


Figure 18: Sensitivity of the probability of ACC failure to the failure of AC Mode S transponder; log-log scale is used.

to create) demonstrate a very good accuracy that is well within the sampling accuracy of the agent-based simulation (see Figures 8,9).

In addition, a fully analytical procedure has been developed that evaluates the reliability of several layers of collision avoidance. One of the distinct features of the developed procedure is its ability to model dependent subsystems (see the discussion of modeling subsystems K and B) by employing a novel semi-inversion of the Markov model. Specifically, a submodel that corresponds to the nonrenewable part of the system is evaluated using Boolean algebra, and the expressions obtained for subsystem state probabilities are used to infer the transition rates among those states. Finally, those transition rates are supplemented by the inclusion of recurrent events resulting in a complete state-space representation.

In accordance with Monte Carlo simulation [5], out of 10 billion runs there were 1180 cases where three layers of AAC failed (829 of those case were resolved using the last fourth layer). This translates into the chances of failure of all three systems to be 1.180×10^{-7}.

This can be compared with the numerical results obtained analytically using the developed procedure. If only the transponder is allowed to fail (and all the other subsystems cannot fail), then the total probability of failure is 1.117×10^{-7}, which is very close to the result reported in Ref. [5]. As expected, transponders dominate the overall failure rate (due to any lack of considered redundancy). Upon the completion of the first phase (9 steps of AR), the AAC system will fail with the probability 1.04315×10^{-7}, while the chances that the system will transition to TSAFE is 7.0173×10^{-4}. After all three phases are completed and TCAS has been engaged, the chance of failure of AAC is 1.548×10^{-7}. As mentioned previously, this number is obtained using uncorrected $P_{OTHER} = 0.10549$, while in Ref. [5], a correction is made and $P_{OTHER-CORR} = 0.100112$ is considered instead. As described in the Appendix, the authors of the current paper question the motivation behind this correction. However, if it is used, that the final number slightly changes to 1.524×10^{-7}. Figure 18 depicts the sensitivity of the probability of ACC failure with respect to the probability of AC Mode S transponder failure, which is the main driver of the system failure. The latter is varied over the two orders of magnitude, and depicted using the log-log scale with the results from Ref. [5, 35] shown for comparison as well.

Appendix

All the values used are described in more detail in Ref. [5]. The values are either obtained from prior sources, including Ref. [15], or assumed where no data was available. There are two types of probabilities specified:

1. The failure probabilities (per flight) for AAC components as shown in Table 1. Please note that all quantities provided in the table (except ACR) are per aircraft (so indices 1 and 2 are used in fault trees to indicate each aircraft). Each of these component failures can be used to evaluate the probability of such a component being in failed state at time t (using the exponential distribution assumption and the fact that no failures are considered prior to $t = 0$):

$$F_{AG}(t) = 1 - \exp\left[-P_{AG}\frac{t}{T}\right] \tag{41}$$

   Here T is the total duration of a flight (in the context of AAC, we consider $T = 120$ min)[5].

2. The probabilities of failure on demand are given in Table 2. The probability of detection for an aircraft (both AR and TSAFE) depends on the time to conflict (this probability is related to the uncertainty of the trajectory generation, and is approximated by the following curve fit):

$$P_D(t) = \left(-6.2 \times 10^{-12}\right) t^4 + \left(1.2 \times 10^{-8}\right) t^3 + \left(-6.3 \times 10^{-6}\right) t^2 + \left(-3.0 \times 10^{-4}\right) t + 1.0 \tag{42}$$

   where t is measured in seconds. Note that P_{OTHER} corresponds to the failure of TCAS to resolve conflict per aircraft, assuming that both speakers and the transponders are operating properly. The numerical value for $P_{OTHER} = 0.10549$ is calculated in Ref.[5] by removing assumed failure probabilities of transponders and speakers from the historically observed probability of TCAS failure. This value is used in the current paper; however, in contrast to Ref.[5], no exponential correction is used in parametric calculation, as the number already corresponds to per-demand failure (and is not directly associated with the accumulation of failure with time).

Component	Functionality Description	Failure Probability
AG	ADS-B: AC gets position via GPS	0.0005
AB	ADS-B: AC Mode S transponder	0.0000097
AD	ADS-B: Ground station sends ACs data to Host	0.00002
RS	Radar: Ground signals AC to reply	0.00682
ACR=TCR	Conflict resolution module	0.000001
VDL2	VDL2: Resolution upload via data link (per AC)	0.00004
VC	AC data sent to host comps via voice communication	0.00055
RR	Onboard resolution reader (per AC)	0.000001
FMS	Onboard resolution trajectory generator (per AC)	0.000097
S	Onboard speaker (per AC)	0.000001

Table 1: Failure probability for components (per flight)

Component	Functionality Description	Failure Probability
P_{FPT}	Generate AR 4-D flightplan trajectory (per AC)	0.000001
P_{DRT}	Generate TSAFE 4-D flightplan trajectory (per AC)	0.000001
P_{TSOAR}	Successful override of AR by TSAFE	0.000001
P_{TCOTS}	Successful override of TSAFE by TCAS	0.000001
$P_D(t)$	Successful detection (both AC)	Varies (see Eq. 42)
P_{OTHER}	Successful TCAS (per AC)	0.10549

Table 2: Probabilities of failure on demand

References

[1] Sense and avoid for unmanned aircraft systems. final report of faa sponsored sense and avoid workshop. Technical report, Federal Aviation Administration, 9 October 2009.

[2] J.W. Andrews, H. Erzberger, and J.D. Welch. Safety analysis for advanced separation concepts. *Air Traffic Control Quarterly*, 14(1):5–24, 2006.

[3] Henk A P Blom, Bart Klein Obbink, and G J Bakker. Simulated safety risk of an uncoordinated airborne self separation concept of operation. *Air Traffic Control Quarterly*, 17(1):63–93, 2009.

[4] Henk A. P. Blom, Bart Klein Obbink, and G.J. Bakker. Safety risk simulation of an airborne self separation concept of operation. In *Collection of Technical Papers - 7th AIAA Aviation Technology, Integration, and Operations Conference*, volume 1, pages 331–339, 2007.

[5] David M. Blum, David Thipphavong, Tamika L. Rentas, Ye He, Xi Wang, and M. Elisabeth Pate-Cornell. Safety analysis of the advanced airspace concept using monte carlo simulation. In *AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario, Canada, Aug. 2-5 2010.

[6] E. Bonabeau and C. Meyer. Swarm intelligence: a whole new way to think about business. *Harvard Business Review*, pages 107–114, May 2001.

[7] D. Braha, A. A. Minai, and Y. Bar-Yam, editors. *Complex Engineered Systems: Science meets Technology*. Springer, Berlin, Germany, 2006.

[8] G. Calanni Fraccone, V. Volovoi, A. Colón, M. Hedrick, and R. Kelley. Agent-based simulation of off--nominal conditions during a spiral descent (nextgen vehicle nra). In *Proceedings of the 9^{th} Aviation Technology, Integration, and Operations (ATIO) Conference*, Hilton Head, SC, USA, September, 21-23 2009.

[9] K. Campbell, W. Cooper, D. Greenbaum, and L. Wojcik. Modeling distributed human decision-making in traffic flow management operations. In *Third USA/Europe Air Traffic Management Research and Development Seminar*, Naples, Italy, June, 13-16 2000.

[10] J.-P. Clarke, N. Ho, and L. Ren. Continuous descent approach: Design and flight test for louisville international airport. *Journal of Aircraft*, 41(4):1054–1066, September-October 2004.

[11] R. David and Hassane Alla. *Discrete, Continuous, and Hybrid Petri Nets*. Springer-Verlag, Berlin, Heidelberg, Germany, 2005.

[12] Y. Dutuit, E. Châtelet, J.-P. Signoret, and P. Thomas. Dependability modelling and evaluation by using stochastic Petri nets: Application to two test cases. *Reliability Engineering and System Safety*, 55:117–124, 1997.

[13] H. Erzberger and R.A. Paielli. Concept for next generation air traffic control system. *Air Traffic Control Quarterly*, 10(4):355–378, 2002.

[14] K. Harper, S. Guarino, A. White, and M. Hanson. An agent-based approach to aircraft conflict resolution with constraints. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Monterey, CA, USA, August, 5-8 2002. Paper No. AIAA 2002-4552.

[15] Robert Hemm and Andrew Busick. Safety analysis of the separation assurance function in today's national airspace system. In *9th AIAA Aviation Technology, Integration, and Operations Conference*, Hilton Head, South Carolina, Sep. 21-23 2009.

[16] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, Roy Jenne, and Dennis Joseph. The ncep/ncar 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3):437–471, 1996.

[17] M.J. Kochenderfer, M.W.M. Edwards, L.P. Espindle, J.K. Kuchar, and J.D. Griffith. Airspace encounter models for estimating collision risk. *J. Guid. Control Dyn. (USA)*, 33(2):487 – 99, 2010.

[18] M.J. Kochenderfer, L.P. Espindle, J.K. Kuchar, and J.D. Griffith. A comprehensive aircraft encounter model of the national airspace system. *Linc. Lab. J. (USA)*, 17(2):41 – 53, 2008.

[19] C.M. Krishna and A.D. Singh. Reliability of checkpointed real-time systems using time redundancy. *Reliability, IEEE Transactions on*, 42(3):427 –435, September 1993.

[20] Leslie Lamport, Robert Shostak, and Marshall Pease. The byzantine generals problem. *ACM Trans. Program. Lang. Syst.*, 4:382–401, July 1982.

[21] S. Lee, A. Pritchett, and D. Goldsman. Hybrid agent-based simulation for analyzing the national airspace system. In *Proceedings of the 33^{rd} Winter Simulation Conference*, pages 1029–1036, Arlington, VA, USA, December, 9-12 2001.

[22] A. Lisnianski and A. Jeager. Time-redundant system reliability under randomly constrained time resources. *Reliability Engineering and System Safety*, 70(2):157 – 166, 2000.

[23] A. Lisnianski, G. Levitin, and H. Ben-Haim. Structure optimization of multi-state system with time redundancy. *Reliability Engineering and System Safety*, 67(2):103 – 112, 2000.

[24] L. Meyn, T. Romer, K. Roth, L. Bjarke, and S. Hinton. Preliminary assessment of future operational concepts using the airspace concept evaluation system. In *4^{th} AIAA Aviation Technology, Integration, and Operations (ATIO) Forum*, Chicago, IL, USA, September, 20-22 2004.

[25] M. K. Molloy. *On the Integration of Delay and Throughput Measures in Distributed Processing Models*. PhD thesis, Department of Computer Science, University of California, Los Angelos, CA, USA, 1981.

[26] S. Natkin. *Les réseaux de Petri Stochastiques et leur Application a l'Evaluation des Systemes Informatiques*. PhD thesis, Conservatoure National des Arts et Metier, Paris, France, 1980.

[27] W. Niedringhaus. An agent-based model of the airline industry. The Mitre Corporation, 2000. McLean, VA, USA.

[28] M. Parker. What is ascape and why should you care? *Journal of Artificial Societies and Social Simulation*, 4(1), 2001.

[29] M. Pecht and J. Gu. Physics-of-failure-based prognostics for electronic products. *Transactions of the Institute of Measurement and Control*, 31:309–322, 2009.

[30] C. Perrow. *Normal Accidents: Living with High-Risk Technologies*. Princeton University Press, Princeton, NJ, USA, 1999.

[31] A. Petri. *Kommunikation mit Automaten*. PhD thesis, Institut für Instrumentelle Mathematik, Schriften des IIM, 1962.

[32] L. Ren, N. T. Ho, and J.-P. Clarke. Workstation-based fast-time aircraft simulator for noise abatement approach procedure study. In *4^{th} AIAA Aviation Technology, Integration, and Operations (ATIO) Forum*, Chicago, IL, USA, September, 20-22 2004.

[33] R. Shepherd, R. Cassell, R. Thapa, and D. Lee. A reduced aircraft separation risk assessment model. In *AIAA Guidance, Navigation, and Control Conference*, New Orleans, LA,, Aug. 11-13 1997.

[34] F. J. W. Symons. *Modelling and Analysis of Communication Protocols Using Numerical Petri Nets*. PhD thesis, Department of Electrical Engineering Science University of Essex, Essex, England, 1978.

[35] David Thipphavong. Accelerated monte carlo simulation for safety analysis of the advanced airspace concept. In *10th AIAA Aviation Technology, Integration, and Operations Conference*, Fort Worth, Texas, Sep. 13-15 2010.

[36] Roland E. Weibel and R. John Hansman Jr. Safety considerations for operation of different classes of uavs in the nas. volume 1, pages 341 – 351, Chicago, IL, United states, 2004.

[37] U. Wilensky. Netlogo. http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, USA, 1999.