Safety performance functions to predict separation minima infringements in en-route airspace

Raquel Delgado-Aguilera Jurado, Victor Fernando Gómez Comendador, María Zamarreño Suárez and Francisco Pérez Moreno

Department of Aerospace Systems, Air Transport and Airports, Universidad Politécnica de Madrid (UPM), Madrid, Spain

Christian Eduardo Verdonk Gallego

Air Traffic Management Research and Development Reference Centre (CRIDA), Madrid, Spain, and

Rosa María Arnaldo Valdes

Department of Aerospace Systems, Air Transport and Airports, Universidad Politécnica de Madrid (UPM), Madrid, Spain

Abstract

Purpose – The purpose of this study is to establish a systematic framework to characterise the safety of air routes, in terms of separation minima infringements (SMIs) between en-route aircraft, based on the definition of models known as safety performance functions.

Design/methodology/approach – Techniques with high predictive capability were selected that enable both expert knowledge and data to be harnessed: Bayesian networks. It was necessary to establish a conceptual framework that integrates the knowledge currently available on the causality and precursors of SMIs with the hindsight derived from the analysis of the type of data available. To translate the conceptual framework into a set of causal subnets, the concepts of air traffic management (ATM) barrier model and event trees have been incorporated.

Findings – The model combines analytics and insights, as well as predictive capability, to answer the question of how airspace separation infringements are produced and what their frequency of occurrence will be. The main outputs of the network are the predicted probability of success for the ATM barriers and the predicted probability distribution of the vertical and horizontal separation of an aircraft in its closest point of approach.

Originality/value – The main contribution of this work is that, by virtue of the calculation capacity obtained, the network can be used to draw conclusions about the impact that a modification of the airspace and of the traffic, or operational conditions, would have on the effectiveness of the barriers and on the final distributions of distance between aircraft in the CPA, thereby estimating the probability of SMI.

Keywords Bayesian networks, ATM safety barriers, Predictive capability, Safety performance functions, Separation minima infringement

1. Introduction

One of Europe’s challenges is to achieve an efficient, smart, sustainable and safe air transport system (Rodríguez-Sanz et al., 2020).

© Raquel Delgado-Aguilera Jurado, Víctor Fernando Gómez Comendador, María Zamarreño Suárez, Francisco Pérez Moreno, Christian Eduardo Verdonk Gallego and Rosa María Arnaldo Valdes. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial & non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licenses/by/4.0/legalcode

This project has received funding from the SESAR Joint Undertaking under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 892542.

Received 4 November 2021
Revised 1 February 2022
Accepted 28 February 2022

Definitions, acronyms and abbreviations

ATCo = air traffic controller;
ATM = air traffic management;
ATS = air traffic service;
BN = Bayesian network;
CPA = closest point of approach;
dCPA = distance closest point of approach;
ETA = event tree analysis;
FL = flight level;
ft = feet;
LoS = loss of separation;
NM = nautical miles;
SESAR = single European sky ATM research;
SMI = separation minima infringement;
SPF = safety performance function;

STCA = short-term conflict alert; and
TLC = time of last clearance.
The term was coined in the field of road transport, with the development of models to predict the occurrence of traffic accidents, and over the years, their use has been expanded to other areas such as railways and aviation.

The term is used in many sectors to refer to, in general, mathematical models that have the ability to explain and predict the occurrence of safety events.

In accordance with this approach, some projects have extended the concept of SPFs to ATM to develop models capable of explaining and predicting the occurrence of SMIs by considering different precursors.

In Arnaldo Valdes et al. (2019), a frequentist statistical approach is used to characterise SMIs between aircraft as count data with an excess of zeros and over dispersion. Subsequently, the relationships between the number of conflicts between aircraft in a certain route segment and the airspace design and traffic flow characteristics are modelled, using zero-inflated models (Lang and Li, 2010).

Based on the characteristics of the route segment, the distribution that best matches the observations of the number of conflicts in the airspace segments is a zero-inflated negative binomial probability distribution. It also takes into account the large number of null values that characterise aviation safety events. However, this first approximation does not fully exploit the potential of Bayesian network (BN) technology to use causality inference and prediction in ATM.

A more complete alternative is to use empirical Bayesian models. These models allow two common problems associated with predictive security models to be addressed (Hauer, 2002): 1) regression to the mean (RTM) is considered; and 2) the lack of data when there is an insufficient historical period or a very low number of occurrences.

RTM is a common bias when evaluating a network in terms of accident rate or safety, since a point or element of the network can have high frequencies of occurrence one year, and the following year can have a lower and more characteristic frequency of occurrence.

Traditional statistical approaches are not useful in this case due to the small number of SMIs (Arnaldo Valdes and Gomez Comendador, 2021). With a very small number of SMIs in the total flight sample, conventional approaches will not be statistically relevant. The decision was made to use other techniques with high predictive capability that enable both expert knowledge and data to be harnessed and which have proven useful in estimating low probability events: BNs.

The BN provides a better estimate of the safety of a part of the air transport system, taking into account not only the number of safety occurrences in that location but also the occurrences observed in similar environments, naturally incorporating the knowledge of experts on the causes that could have produced them (Hauer, 2002).

However, although there are examples of situations where BN have been skillfully applied in ATM (Neil et al., 2003; Gomez Comendador et al., 2019; Bujor and Ranieri, 2016; F. and Y., 2012), its potential for explaining and predicting the occurrence of safety events, such as SMI, has not yet been evaluated.

To better understand the potential of BN and support the complex model proposed in this document, the next section provides an initial overview of the main concepts involved in the construction of a BN and how its results can be exploited.
posterior probabilities. BNs can be created using expert knowledge only, directly from the data or as a combination of both, as in this case. The process involves data processing, structure learning and parametric learning (Mora Villazán, 2017).

Conditional probability is the probability of an event A occurring, knowing that an event B occurs. It is given by the following expression:

$$P(A_i | B) = \frac{P(B | A_i) \cdot P(A_i)}{P(B)} = \sum_{i=1}^{n} P(B | A_i) \cdot P(A_i)$$

(1)

where $P(A_i)$ is the a priori probability; $P(B | A_i)$ is the conditional probability; $P(B)$ is the probability of observing $B$, the marginal probability; and $P(A_i | B)$ is the posterior probability.

1.1.1 Methods of constructing Bayesian networks
A BN can be built based on expert knowledge, real data or a mixture of both. This section explains how a network should be created based on the method used (Cooper and Herskovits, 1992).

- **Expert knowledge**: The experts must decide which variables to include in the model and establish the causal relationships between them. They also have to complete the conditional probability tables for each child node.
- **Real data**: When constructing BNs directly from data, the causal relationships, and conditional probability tables, will be extracted from the data.
- **Mixed case**: In this case, the causal relationships created directly from the data can be modified by adding, removing or changing the directions of the arcs.

In the present study, the mixed case was chosen.

1.1.2 Construction of Bayesian networks
The construction of a Bayesian network using a database and software comprises three steps as shown in Figure 1:

- **Data pre-processing**: In this step, the variables that will be used to model the problem are selected. Data cleansing is performed to detect potential failures and correct them, if necessary. This step is essential, since if the data is not correct or has a fault, the learning of the network will not be optimal, and therefore, the results obtained from it will not be valid.
- **Structure learning**: In this step, the structure of the network is determined; in other words, the dependent and independent relationships between the variables are established. This learning can be done directly from the data provided or, as in this study, prior knowledge can be introduced into the model.
- **Parametric learning**: The last stage consists of obtaining the necessary a priori and conditional probabilities, given the previously defined structure. These probabilities are obtained from the observed frequency of the data.

2. Methodology
It is not easy to develop a BN model for a problem as complex as the prediction of SMIs. It was necessary to establish a conceptual framework that integrates the knowledge currently available on the causality and precursors of SMIs with the hindsight derived from the analysis of the type of data available in the project, in particular those that reflect the interventions of ATCo (Netjasov et al., 2019).

- initial approach to SPF focus on the CPA;
- ATM barrier model;
- event tree analysis (ETA); and
- definition of the BN.

2.1 Initial approach to safety performance function: focus on closest point of approach
The conceptual framework underpinning the proposed BN model considers the general scenario in which aircraft evolves and focuses on analysis of the CPA (Katta and Madani, 2017), for any possible pair of aircraft in an air traffic sector, and in understanding and quantifying the process that leads to the aforementioned CPA.

The CPA is an essential factor in aircraft safety, especially when the aircraft is avoiding loss of separation or avoiding collision. The CPA is an estimated point at which the distance between the aircraft itself and another aircraft (viewed as another target or intruder) will reach the minimum value. In a conflict, the distance between the aircraft will first decrease, and then, if the collision does not occur, it will start to increase again. The CPA is the location of the aircraft in three-dimensional airspace where, based on information on the current state, the two aircraft have the least separation. If no action is taken, the location is always on the current flight path (Errico and Di Vito, 2019).

The three main elements of the conceptual framework are considered.

Figure 2 shows the interaction between two pairs of aircraft in a sector and their respective CPAs, which are represented by a red circle. The actual final CPA between a pair of aircraft can be interpreted as the result of a process in which the predicted trajectories of the aircraft are modified as a result of ATCo intervention. The CPA between a pair of aircraft may then be considered to be the shortest actual distance between these two aircraft, expressed as vertical separation and horizontal separation. This distance will be known as the final CPA. It is also possible to calculate what the CPA would have been between this pair of aircraft, had both of them followed their intended trajectory without any modification or intervention by ATCo. This distance will be known as the prior CPA. The difference between both distances, final CPA and prior CPA, is
attributed to alterations in the predicted trajectories due to ATCo intervention.

Based on this approach, the framework considers three main elements in the interaction of each pair of aircraft within the area of responsibility of a specific controller:

- **Distance closest point of approach prior (dCPA prior):** Distance at which two aircraft would cross considering only their predicted trajectories, measured horizontally in NM and vertically in feet.

- **Time of last clearance (TLC):** Time elapsed since the last ATCo clearance to either of the pair of aircraft and the instant in which the CPA occurs, measured in seconds (Oktal and Yaman, 2011).

- **Distance closest point of approach final (dCPA final):** Shortest final distance crossed by the two aircraft after receiving clearance from the controller. The same units as dCPA prior.

Figure 3 gives a diagram with the concepts explained above. The two blue points in the diagram correspond to a pair of aircraft, A_i and A_j. In Figure 3, dCPA prior (the distance at which the two aircraft would cross based on their expected trajectories) is shown in red.

The controller is responsible for detecting whether this dCPA prior could constitute a possible SMI between this pair of aircraft and, if so, acting upon them to avoid it. The distance between the aircraft at the time the controller issues the last clearance is indicated in the diagram as d at TLC. The aircraft will eventually cross at a distance called dCPA final, which is represented in Figure 3 with a green line.

Each of these three elements can be modelled as a subnet consisting of the precursors and the causal relationships that influence them (see Figure 4). The subnet represented by the CPA prior bubble will, based on a set of selected precursors, estimate the probability distribution of the vertical and horizontal separation between any pair of aircraft in its CPA prior. By comparing this separation distribution with the applicable separation minima, the probability of potential conflicts can be determined. The concept behind the subnet indicated by the CPA final bubble is similar. The sub-network indicated by the TLC bubble accounts for ATCo clearance and its precursors.

To translate the conceptual framework into a set of causal subnets, we incorporate the concepts of ATM barrier model and event trees.

### 2.2 Air traffic management barrier model: an abstraction of aircraft separation provision function

Based on the *Swiss cheese model*, the ATM barrier model explicitly shows the progression of the incident and may be used as a “live” model to prevent future separation infringements or to intervene in an incident to halt its development (Pejovic et al., 2020). It divides the process of controlling the aircraft into different stages in which the safety barriers are identified.

The barrier model represented in Figure 5 is an adaptation of that proposed by Eurocontrol (Perrin and Kirwan, 2007), taking into account the knowledge- and data-based approaches followed in the saFety And Resilience guidelines for aviatiOn (FARO) [1] project, safety and resilience guidelines for aviation, funded by the European Union in the SESAR 2020 programme. It specifically reflects the stages and obstacles in...
the progress of an SMI that could be studied and analysed with the data available in the project.

A process is carried out to identify the sequence of barriers that arise before a loss of separation occurs (see Figure 5):

- **Interaction identification:** The first step is to find out if the two aircraft constitute an interaction. An interaction is defined as two aircraft that are within 20 NM of each other. The concept of a 20NM horizontal boundary is simply a filter to optimise the process of monitoring the CPA between aircraft. It makes little sense to constantly monitor the CPA of two aircraft which, although they are in the same sector, are known to have a minimum separation of more than 20NM since they will never constitute an SMI.

When two aircraft constitute an interaction, they are considered to be a pair. Due to the nature of the data, the scope of the FARO safety model is aircraft pairs.

- **Assessment of potential conflict:** If two aircraft constitute a pair, the probability that the pair constitutes a potential conflict under the conditions of the situation existing at that moment, i.e. without the action of the controller, must be evaluated.

- **Conflict identification:** Again, in the case of a potential conflict, it will be necessary to assess whether it is detected by ATCo.

- **Conflict resolution in conflicts identified:** The likelihood of conflict resolution by ATCo of those cases that are potential conflicts and are detected by ATCo.

- **Short-term conflict alert (STCA) alert:** The probability of activation of the STCA is evaluated for conflicts not identified by ATCo.

- **Conflict resolution:** Finally, for conflicts detected but not resolved, the probability of resolution of the conflict following activation of the STCA alert.

### 2.3 Event tree analysis

The principles of ETA are also used to effectively translate this barrier model into a representation of a causal network. The event tree provides a top-down logic modelling technique for success and failure that explores responses via a single initiating event. It establishes a mechanism for evaluating the probabilities of the results and global analysis of the system (Flammini et al., 2006). The outcome of each node represents Boolean logic.

This model provides a visual approach to cause–effect relationships, as well as enabling us to explore all of the possibilities. Furthermore, it allows complex models to be simplified and addressed in a more understandable way. The analyst must identify the initial challenge. The probabilities of success or failure are often difficult to determine.

The results of this integrated model appear in Figure 6. The upper bar gives the sequence of the ATM barriers. The probabilities of each event occurring (or not) are shown, as are the bifurcation lines of each decision and the consequences of each of the branches. Furthermore, the likelihood of each of the results can be expressed as a conditional probability of each of the branches of the tree, as indicated in Figure 7.

### 2.4 Definition of Bayesian network

Now that the conceptual framework and its causal model have been developed at a more detailed level of granularity, the barrier model can be implemented using BN. Based on the barrier model described above, each of these barriers will be modelled using a BN, as shown in Figure 8, which gives an overview of the proposed network.

Two types of subnets can be identified: those designed to estimate the probability of an event occurring and those designed to estimate the vertical and horizontal distance in the CPA between aircraft.

The first type of subnet explains the modelling of ATM barriers and their effectiveness. They are depicted in Figure 8 by a box containing blue and yellow bubbles. The second type explains the modelling of the results of the event tree and the probability distribution of the vertical and horizontal separation between the aircraft pair included in each result. These subnets are represented in Figure 8 by a red and blue histogram. It should be noted that the entire model comprises 24 subnets integrated into a large BN.

The entry conditions of a subnet consist of precursors, traffic states and scenario conditions. Arising from the data, the input variables are modelled as a probability distribution. Using the data, the subnets will learn the table of conditional probabilities related to all the variables. The subnets will, therefore, capture the causal relationship between them (Sacchi, 2015). Finally, the subnet will enable the probability distribution of the outcome variables to be predicted. One of the most relevant characteristics of this BN is that it will infer ATCo behaviour based on the recorded actions of ATCo.

The output of the first type of subnet is a binomial variable representing the probability of failure or success of each of the
barriers in the model, for example, the probability of conflict detection, the probability of conflict resolution, etc. The output of the second type of subnet is the probability distribution of the vertical and horizontal separation between aircraft at the CPA. This probability distribution is given as a weighted average of the branches of the event tree, as indicated in Figure 9.

The model considers normal operation and uses a probabilistic model to identify the ATCo actions that could lead to a STCA or an SMI. Furthermore, the model takes account of how that SMI is managed under different conditions, for example, whether or not a previous STCA has occurred.

The sequence of barriers is intended to reflect the timeline in which the barriers occur.

3. Analysis of results: Use case

To complete the presentation of the methodology and to illustrate the results obtainable from the network, the principles of each type of subnetwork will be explained, as well as the main analyses that can be carried out to refine and adjust each of these networks.

The data used in the model is ATC data obtained from the ENAIRE radar traces. This was provided to the authors following processing and validation by the research company, ATM research and development reference centre.

3.1 Subnet type 1: Assessment of air traffic management safety barriers

The main characteristic of this type of network is that the result is always a binary variable/node that represents the effectiveness of an ATM safety barrier. To illustrate the operation of these subnets, the aircraft interaction evaluation barrier is examined below (Figure 10).

The objective of this subnet is to identify the probability that an aircraft will have an interaction with another aircraft before ATCo acts on either of them.

For each of these pairs of aircraft, the vertical (ft), horizontal (NM) and 3D (NM) separations are provided, as well as
the latitude, longitude and flight level (FL) of each aircraft, every 5 s.

The general conditions of the sector or the traffic conditions when aircraft enter the sector form the input data to this subnet. These are all the variables that are actually known at the time the aircraft enters the sector and which are considered to have a direct influence on the probability that two aircraft will interact.

According to the event tree model, this subnet will give rise to two different branches “p” and “1-p”. In this way, the probability “p” that a pair of aircraft does not interact can be obtained, as well as the probability “1-p” that a pair of aircraft does interact. Aircraft in the first branch “p” will be analysed later on subnet 1 A, while aircraft in branch “1-p” will be analysed in subnet 2.

Therefore, the construction of the subnet is based on:

- inputs to the subnet: input data on the macroscopic conditions of the sector;
- training data: for each aircraft, the model checks the pairs file to see if it has formed a pair with another aircraft; and
- outputs from the subnet: probability of interaction. This node has two states: yes/no.

**3.2 Subnet type 2: Assessment of distribution of vertical and horizontal separations**

Part of the subnets of the model are used to capture the distribution of vertical and horizontal separations for each branch of the probabilistic event tree. The main characteristic of this type of network is that the results are always a discrete probability distribution that represents the vertical or horizontal separation at the CPA for the pairs of aircraft belonging to each branch of the event tree.

This concept is illustrated in **Figure 12**. The vertical and horizontal components have been grouped into two different subnets. The subnet that estimates the distribution of the vertical separation primarily considers the variables related to the vertical component of the aircraft’s motion. The network that estimates the distribution of horizontal separation primarily considers variables that are related to the horizontal component of the aircraft’s motion.

The distribution of horizontal separations is estimated only for those pairs of aircraft whose vertical separation at the CPA is less than 1,000 feet. Horizontal separation follows a uniform distribution for those pairs of aircraft whose vertical separation at the CPA is 1,000 ft or greater.

**Figure 13** gives further information about each of the subnets. It shows the subnets, their input variables and the discretisation of the output variable.

As previously mentioned, the distribution of vertical distances will be obtained and those that are less than 1,000 feet will be fed into the horizontal separation subnet to also give this distribution.
The horizontal subnet takes the following parameters into consideration: difference in horizontal speed between the aircraft entering the sector, difference in cruising speed, distance and time of entry of the aircraft into the sector and, finally, difference in direction.

Figure 14 depicts the entries and results of the model developed thus far. The left-hand side summarises the input variables. These are characterised as probability distributions and relate to the general conditions characterising the sector and its traffic. The figure includes several examples of input variables, such as the distribution of incoming aircraft to the sector, the distribution of FL at the point of entry of aircraft to the sector and the distribution of aircraft speed.

The outcomes of the network will be characterised as probability distributions and are summarised in the centre and right-hand side of the figure. For a given state of the input variables, the central part of the triptych shows the predicted probability of success of the ATM barriers, for example, the probability of interaction between aircraft, the probability of potential conflict, the probability of detecting the conflict, the probability of resolving the conflict and so on. For a given state of the input variables, the right-hand side of the triptych shows the predicted probability distribution of the vertical and horizontal separation of the aircraft at the CPA.

4. Conclusions

The framework makes it possible to estimate the impact of changes in the conditions of entry to the network on the effectiveness of the barriers and on the final distributions of distance between the aircraft at the CPA, thereby giving the probability that an SMI will occur.

The overall BN permits both forward and backward analysis. In backward analysis, the model will be used to provide a particular configuration (combination) of the input variables leading to an SMI, by adjusting the output variables to a target value. In forward analysis, the model is used to predict the output variables by establishing the probability distribution of the input nodes.

It should be noted that the model must use data relating to a specific combination of ATC sectors. This means that the conditional probability tables of the model are specific for each use case and have to be learned from the data of the specific sector analysed. Adapting the generic model to the characteristics of the sector analysed in each use case implies:

- gathering and processing all the data of the sector/traffic being analysed in the use case. Processing implies the
discretisation of the data for each variable of the model. The discretisation scheme depends on knowledge; it cannot be automated and may be different for each sector because it depends on the characteristics of the sector and the traffic profile:

- parametric learning, i.e. obtaining the necessary a priori and conditional probability tables from the frequency observed directly from the data;
- sensitivity analysis to refine the network and identify the most influential variables for each particular sector; and
- backward and forward analysis to define thresholds for the variables that could affect safety performance in a scenario, where applicable.

Once the above work has been completed, the BN will be usable for a specific use case and validation activities can be carried out. Thus, the model developed combines analysis and knowledge, as well as the ability to predict, to respond to the question of how airspace separation infringements occur and what their frequency of occurrence will be. The main outcomes of the network are as follows:

- the predicted probability of success of the ATM barriers, for example, the probability of interaction between aircraft, the probability of potential conflict, the probability of detection of the conflict, the probability of resolution of the conflict, etc.; and
- the predicted probability distribution of the vertical and horizontal separation of aircraft at their CPAs.

With this capacity for calculation, the network makes it possible to draw conclusions about the impact that a modification of the airspace and of the traffic, or operational conditions, would have on the effectiveness of the barriers and on the final distributions of distance between aircraft in the CPA, thus estimating the probability of SMI.

**Note**

1. FARO project: [www.sesarju.eu/projects/faro](http://www.sesarju.eu/projects/faro)

**References**


Further reading


Corresponding author

Raquel Delgado-Aguilera Jurado can be contacted at: raquel.djurado@alumnos.upm.es