A civil aviation safety assessment model using a Bayesian belief network (BBN)

R. Greenberg, S. C. Cook and D. Harris
Systems Engineering and Evaluation Centre,
University of South Australia, Mawson Lakes Campus
Adelaide
Australia

ABSTRACT

In this paper we present a Bayesian belief network (BBN) socio-technical model for investigating the accident rate for multi-crew civil airline aircraft. The model emphasises the influence of airline policy and societal behaviour patterns on the pilots within the piloting system. The main claim of this paper is that a BBN can be used to bring most aviation safety-critical elements into a common quantitative safety assessment despite the unique problems posed by the very low probability of accidents. We support this claim by replicating certain phenomena such as the low accident rate, the difference between the 'more' and 'less' safe airlines and other statistical factors of civil aviation. In particular, the model succeeds in explaining the large gap of six to seven orders of magnitude between in-flight measurements of pilots' error and accident rate.

1. INTRODUCTION

The problem of safety assessment and prevention of low probability, high consequence (LPHC) accidents in complex systems involves:

- Dealing with a large number of variables and many different potential sources of evidence. Many of these variables are interrelated by dependencies that are not well understood formally.
- The need to model human performance as part of the system. Such modelling has been known as one of the most difficult problems. Operator performance is influenced by a combination of personal traits, social beliefs, and organisational policies.
- The fact that 60%-90% of LPHC accidents can be attributed to human operators either as a root cause or a contributor[4,29,34,39,7,23].

Civil aviation accidents are good example of LPHC accidents. This is because, firstly, fatal accidents are very rare and may cause the loss of many lives. Secondly, airliners include highly sophisticated automated systems that are interrelated by dependencies, which are not well understood formally. Thirdly, pilots are a central part of the system and are subject to a combination of personal traits, social beliefs and organisational policies. And fourthly, pilot error is given as the primary cause in about 67% of fatal accidents[4].

BBNs have recently been profitably applied to LPHC accidents [28, 24, 3, 25] and this paper applies this technique to civil aviation safety. It should perhaps be pointed out that Bayes' equation for dependent probabilities follows directly from the axioms of probability theory, and is therefore mathematically completely unobjectionable. However, it posits the use of dependencies of known values, and there is some philosophical objection as to whether or not these can be determined in real life situations. However, as indicated above, others have reported success in applying Bayesian probability successfully to real-life problems, and we are following along this path.

Investigations of LPHC accidents, including aviation accidents, have also shown that there is a certain pattern of events and

Paper No. 2922. Manuscript received 18 October 2004, revised version received 26 September 2005 , accepted 30 September 2005.

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

conditions that precedes the occurrence of such accidents[17]. It was further shown that the source of this pattern of conditions and events is rooted in what is called the organisational 'safety culture,' or the way the organisation and society perceive safety[33]. As it turns out, safety culture affects the performance of managerial personnel and operators, and thus also the accident rate.

An important outcome of the above discussion is that any quantitative assessment model of safety has to include at least parts of this pattern. The authors of this paper made a first attempt to model pilots as individuals subjected to a variety of 'modifying factors' that ranged from training efficiency, skills, through psychological factors and their medical situation to organisational and sociological safety culture[13]. The model was based on two major assumptions: (a) the pilot error rate is a product of a basic value and the 'modifying factors,' and (b) the crewmembers are two independent subsystems, each performing a combination of flying and monitoring tasks.

Greenberg *et al*[13] noted, however, that the above two key assumptions (along with others) are problematic. It is believed that a model based on a BBN is a better approach that will overcome, at least, the difficulties that the major assumptions generate. For example, instead of being limited by a product of 'modifying factors', the exact influence can be introduced in terms of probability values. And a BBN can also better model the interaction between the pilots who are, in fact, cooperating in achieving their mission.

The model can be used in various areas and by many bodies. Airlines can use this model to predict how changes in their policy of activities may affect their safety profile. Examples of such policies include training, preferred crew composition, division of activities between crewmembers at the various stages of flight, rest and active periods of time, etc. Regulatory bodies can use the model in predicting the influence of certain instructions on the accident probability. Aircraft manufacturers can analyse alternative designs and choose the most appropriate one for specific uses.

In this paper we intend to show that BBNs are a good vehicle for the above purposes and we intend to support our claim by showing that certain quantitative findings can be replicated by adjusting the attributes of the model. However this claim does not mean that the present form of the model is finalised. On the contrary, changes will be made to its present form depending on the particular use.

In the following we discuss the nature of the problems that qualitative models in civil aviation have to solve. In section 3 we describe the basic ideas of BBN and show two short examples of their possible use. In Section 4 we describe the cockpit model in which the interaction of the two pilots is stressed. A BBN model of a flight that can cover a range from a normal safe ending to a remote accident is found in Section 5. Section 6 presents some results and shows that the model replicates safety findings. In Section 7, we show how the model can practically be used for predictions of safety behaviour. We summarise the paper by providing conclusions and lines for future work in Section 8.

2.0 THE NATURE OF THE PROBLEM

Many studies have already shown that air transport accidents are indeed very rare events, and it is unlikely that an air transport pilot will be involved in an accident[1,2] in his or her working life. For example, if a pilot were to have a 20,000 hour flying career, and the probability of an accident is in the range of $0{\cdot}7{\cdot}10^{-6}$ per hour[1,18], then the probability of a pilot (in a two-person crew) being involved in an accident in their working life is $2{\cdot}0.7{\cdot}20{,}000/1{,}000{,}000 \approx 3\%$.

Such a low accident rate makes quantitative data acquisition for modelling purposes extremely difficult, resulting in models with such high variance that their validity and utility may be marginal. Many phenomena appearing in flights are very difficult to represent and replicate quantitatively. It seems that global data and large-scale experiments are the only reliable information sources.

Moreover, the literature reveals some fascinating information that requires explanation. Examples include the big difference of 1-2 orders of magnitude between the 'more' and 'less' safe airlines[30]. Also, Line operations safety audits (LOSA)[15, 16] have found that pilots commit many errors per flight with no overall level of safety-criticality which relates to the low accident rate discussed above. LOSA also shows a difference between the two crewmembers in their ability to monitor and catch errors[38]. In this audit, captains demonstrated better ability in correcting their own errors as well as catching errors committed by the First Officer.

Reason's 'Swiss-cheese' (or rotating disks) model[32], as well as other models[35,36,9,37] that deal with socio-technical systems attempt to explain qualitatively that the source of these phenomena is in the organisational and sociological safety cultures. According to these models, defence lines and barriers are defeated stochastically by the formation of latent failures that remain in a dormant state and become active at the accident by certain triggering events. This logical explanation that nicely represents accident occurrences, however, falls short in quantifying phenomena such as those discussed in the previous paragraph. Also, this explanation can hardly be used practically for predictions, engineering, policy determination, etc.

3.0 INTRODUCTION TO THE BAYESIAN BELIEF NETWORK

The earlier mathematical model of aircraft safety that was introduced by Howard[18] and expanded later by Greenberg *et al*[13], used probability to calculate equipment and systems performance reliability. Howard[18] accepted that the two pilots are not completely independent and introduced a cross-monitoring limiting term. The more formal way, however, to deal with dependent probabilities is based on Bayesian logic. The following section provides a brief introduction to the Bayes theorem.

Calculations, based on Bayes theorem may become intractable and complex for large numbers of dependencies. The approach to deal with this complexity is to create a number of smaller sub-models that are manageable and to connect them together to form a Bayesian belief network (BBN). BBNs belong to the class of probabilistic graphical model, i.e. a combination of the two fields: probability theory and graph theory. It is this combination, which makes them a powerful methodology within statistics and in our case in probabilistic modelling of complicated, socio-technical systems.

3.1 Understanding Bayes' theorem

The starting point in understanding BBN is Bayes' theorem. This theorem is the general form of the product rule that is used for calculating the probability of two (or more) independent events, A and B, namely $P(AB) = P(A){\cdot}P(B\backslash A) = P(B){\cdot}P(A\backslash B)$, where is the probability of A and B happening. When A and B are dependent the product rule changes into;

$$P(AB) = P(A)\cdot P(B\backslash A) = P(B){\cdot}P(A\backslash B) \quad \ldots (1)$$

Where, $P(B\backslash A)$ is the probability of event B happening given event A already happened, or in short terminology, B given A. From Equation (1) it is easy to see that;

$$P(A\backslash B) = (P(A)\cdot P(B\backslash A)/P(B)\cdot \quad \ldots (2)$$

which is the simplest version of Bayes theorem. Notice that there is no causal relation between the two events, that is, each event, say A, may occur if the other happened or not. In mathematical notation this is expressed as:

$$P(A) = P(A\,|\,B)\ \cdot P(B) + P(A\,|\,\overline{B})\cdot P(\overline{B}) \quad \ldots (3)$$

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

Where $\overline{B}$ means that event B has not happened. Using similar arguments as used in the application of the product rule in Equations (1) and (2), one can chain the simple Bayesian theorem with more events $E_1,E_2,...,E_n$ to form hypothesis H in context I. The general rule that is obtained is:

$$P(H \mid E_1E_2 \cdots E_nI) = \frac{P(H \mid I) \cdot P(E_1E_2 \cdots E_n \mid HI)}{P(E_1E_2 \cdots E_n \mid I)} \quad \ldots (4)$$

Equation (4) gives the probability of a hypothesis H in context I given events $E_1,E_2,...,E_n$. Note that these events can be regarded as evidence that can either support or counter the hypothesis.

As an example of the use of Equation (4), let's assume that Mr Stevenson was found in the rich neighbourhood where many home break-ins have occurred recently. Mr Jefferson, who suffered one of these burglaries, takes Mr Stevenson to the police station and accuses him in breaking into his house and stealing valuables. At a later stage one of these valuables is found in Mr Stevenson's possession. In Bayesian terminology, this scenario can be described in the following way:

- Hypothesis, H: Mr Stevenson is the burglar that has stolen Mr Jefferson's valuables.
- Context, I: Mr Stevenson was found in the Rich neighbourhood.
- Evidence, E: one of Mr Jefferson's valuables is found in Mr Stevenson's possession.

This scenario can further be evaluated in the following equations:

$$P(H \mid EI) = \frac{P(H \mid I) \cdot P(E \mid HI)}{P(E \mid I)} = \frac{P(H \mid I) \cdot P(E \mid HI)}{P(E \mid HI) \cdot P(H \mid I) + P(E \mid \overline{H}I) \cdot P(\overline{H} \mid I)}$$

where:

- $P(H\backslash EI)$: The Probability that Mr Stevenson has stolen Mr Jefferson's valuables, given that he was found in the Rich neighbourhood and he possessed one of those valuables.
- $P(H\backslash I)$: The probability that Mr Jefferson's accusation is true before any evidence was found (i.e. the prior probability of Mr Jefferson hypothesis).
- $P(E\backslash HI)$: The probability that Mr Stevenson has really stolen Mr Jefferson's valuables. Note that the fact that Mr Stevenson possesses one of Mr Jefferson's valuables does not necessarily imply that Mr Stevenson has stolen them. For example, he could have bought it from the real burglar.
- $P(\overline{H} \mid I) = 1 - P(H \mid I)$ The probability that Mr Jefferson's accusation at the police station before any evidence was found (i.e. the prior probability of Mr Jefferson hypothesis) is false.
- $P(E\backslash \overline{H}I)$: The probability that despite the fact that one of Mr Jefferson's valuables was found in Mr Stevenson's possession, he actually hasn't stolen it.

3.2 Bayesian belief networks

We now turn to Bayesian belief networks. The best way to understand them is by a simple example as follows. Suppose that Abraham uses a car to travel to work while Jacob uses his bike. Abraham may be late when a traffic jam occurs while Jacob is subject to weather conditions, i.e. Jacob may be late in rainy days. Rains, however, may slow traffic because of two reasons. Firstly, visibility decreases and secondly, many people prefer to use their cars instead of public transport increasing the probability of traffic jams. Thus a rain may also cause Abraham to be late, either directly by low visibility, or indirectly by causing traffic jams. Having this knowledge we want to infer reasonable conclusions about the probability of Abraham and Jacob being late. We also would like to infer other conclusions when certain other knowledge (or evidence) is available. For instance if we know that there is rain what is the probability of Abraham and Jacob being late. Similarly, suppose we know that Abraham is late, we expect that there was a traffic jam but what about the probability of Jacob also being late.

Figure 1 is the representation of the example in BBN form. It is easy to see that the road situation may affect Abraham, while the rain influences both as well as road situation. The node 'rain' is a 'root node', as nothing (within the graph) influences it. These two nodes 'rain' and 'traffic jam' are 'parent nodes' to the 'child nodes' 'Abraham late' and 'Jacob late'. Notice that the node 'traffic jam' is both a parent and a child. Similarly ancestry and descendant relations can also be defined.

In BBN each node is associated with a Node probability table (NPT) that represents the exact influences of its parents. In our example these NPTs are shown in Tables 1, 2, 3 and 4. For root nodes the NPT represents basic knowledge or beliefs (e.g. statistical data on rainy days collected by meteorological institutions). The NPTs of child nodes represent knowledge or beliefs on the influence of the parent nodes.

These tables may be understood in the following way. In the left side of the bolded line are the 'given' conditions and in the right side the probability values according to our belief or knowledge. For example, the first row in Table 3 means that the probability of Abraham being late given that there is rain and traffic jam is 0·7, or in the mathematical notation P(Abraham late | rain and a traffic Jam) = 0·7.

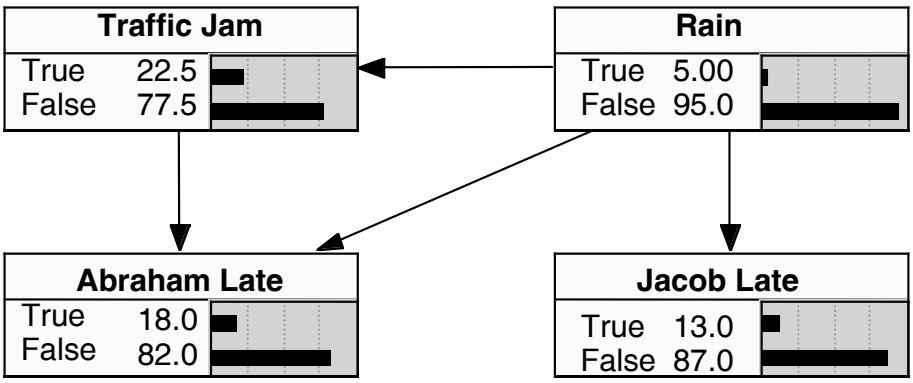


Figure 1. An example of a BBN.

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

Table 1
NPT of 'Rain'

Rain	
FALSE	TRUE
0·95	0·05

Table 2
NPT of 'Traffic Jam'

	Traffic Jam	
Rain	FALSE	FALSE
TRUE	0·3	0·7
FALSE	0·8	0·2

Table 3
NPT of Abraham Late

		Abraham Late	
Rain	Traffic Jam	FALSE	TRUE
TRUE	TRUE	0·3	0·7
FALSE	TRUE	0·4	0·6
TRUE	FALSE	0·8	0·2
FALSE	FALSE	0·95	0·05

Table 4
NPT of Jacob Late

	Jacob Late	
Rain	FALSE	TRUE
TRUE	0·3	0·7
FALSE	0·9	0·1

The values that are shown in Fig. 1 are the a-priori probabilities, that is, they reflect the knowledge in the absence of any specific evidence. Let us calculate one of them, the probability of Jacob being late. Using Equation (3) we get:

P(Jacob Late) = P(Jacob Late|Rain)·P(Rain)+P(Jacob Late|no Rain))·P(no Rain)= 0·7·0·05+0·1·0·95 = 0·13.

This value is equal to the one found in the node 'Jacob Late' of Fig. 1, where the results are given as percentages. In similar ways, other information may be extracted. The calculated values can be changed and revised as more information becomes available. It is easy to see that carrying out the calculations can become very tedious, especially when the BBN grows†. For example, the calculation of Abraham being late depends on whether there was rain and whether there was a traffic jam, but also the traffic jam has to be calculated as it depends on whether there was rain. To continue the example we use a commercial computer program, NETICA[26] that was developed for this purpose.

Suppose we know that Jacob is late. What assumptions can we make on the probability of having rain, of having a traffic jam or even the probability of Abraham being late? Intuitively, we can assume that the probability of having rain will increase and so will the probabilities of having a traffic jam and of Abraham being late. But, we would like to have a quantitative answer and not an intuitive one. Making these calculations we see that indeed that is the case, the probability of having rain increased from 0·05 to 0·269, the probability of there being a traffic jam increased from 0·225 to 0·335 and the probability of Abraham being late increased from 0·18 to 0·265. Now suppose that we know that Abraham is on time (not late). What can we say about the probability of Jacob being late, there being rain, or a traffic jam? As expected, in this case all three probabilities decreased to 0·115, 0·274 and 0·105, respectively. In a similar way other conclusions can be drawn.

Although obvious, it is important to note that the propagation of accumulated information or evidence is in both directions. That is, when inserting new evidence not only are descendant nodes affected but also ancestor ones. This property of BBNs is very important and will be used later in this paper.

3.3 Building a Bayesian belief network

There are two basic stages in the formation of a BBN, building the structure or topology and the determination of the NPTs. From our experience, the first is almost straightforward; the second stage, however, is more difficult. The topology of a network is determined by the causal interrelationships among the factors that compose the network. In a socio-technical system, as the one discussed later in this paper, these factors include physiological, psychological and sociological elements that may be connected in many variations. As a result, both the topology and the NPTs may become very complex. It is, therefore, recommended that only the most important factors will be taken into account. We will return to this problem later in this section.

As mentioned in the previous section, each child node is associated with an NPT that contains the influence of its parent nodes. NPTs associated with root nodes contain basic beliefs or knowledge. In both cases it is very important to understand that for a network to be fully specified (and populated), there must be a relation stored at each node's NPT, which expresses the value of that node in terms of its parents (or as a constant for root nodes). This relation is a function, which provides the probability of each state of the child, for each possible configuration of parent values. The main difficulty is in the formation of this function. The easiest way is when events frequencies are known or can be measured (e.g. climatological statistics or probability of having defective parts in production line). When this knowledge cannot be obtained directly (as is the most frequent case), more sophisticated ways, such as accumulation of indirect data out of which these values can be derived, or when even this cannot be achieved, experts' elicitation, might be required.

Another difficulty is usually encountered when introducing this knowledge into the NPTs. This difficulty is a result of the tables' complexity (actually, the table dimensions in terms of the number of rows and columns). In general, the number of rows is the product of the number of all the states of each parent node and the number of columns is the number of states of the child node (the node that is associated with the particular table). In the example of Subsection 3.2 the biggest table (associated with the node 'Abraham Late') is composed of four rows with one column (the sum of each row has to be one so determining either a FALSE or TRUE value gives the other one). It is easy to see that NPTs tend to grow very fast and dealing with such tables can become extremely difficult. From our experience the maximum manageable number of rows is in the range of 12-16. There are many ways to cope with large NPTs, each with its advantages and drawbacks. In this paper we used two:

- Neglecting parent nodes that seemed to have minor influence on the child node. The problem with this approach stems from the fact that the influence evaluation may turn out to be wrong at a later stage, necessitating major changes.

† Although Bayes' theorem is very well known and the idea of using it in conjunction with networks is mathematically well established, the calculation complexity prevented its commercial use until the beginning of the 1990s. From a theoretical point of view, the complexity of BBN calculations can be shown to be NP-complete and even the introduction of computers did not solve the problem. It was not until the 1980s that mathematical theorems that lessened the number of dependencies and dealt with loops, were found[24, 29]. These theorems enabled very efficient algorithms to be written that cut the run time significantly. Today, many commercial programs can be found. NETICA that is mentioned in Ref 26 is the one used in this paper.

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

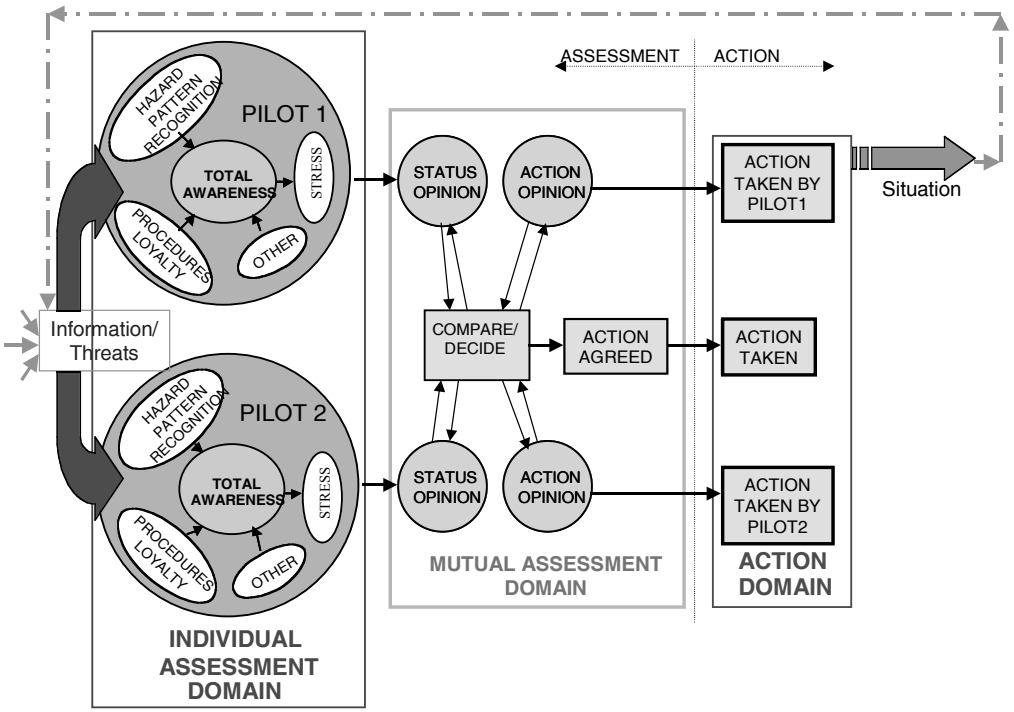


Figure 2. The cockpit model.

- Introducing intermediate nodes that group parts of the parent nodes. Since it is a product that determines the number of rows, their total number after the introduction of intermediate nodes is much smaller than it would have been before. This, however, decouples the parent and child nodes, thereby hiding (and sometimes distorting) the connection between these two node types.

4.0 THE COCKPIT MODEL

An illustration of the activities that take place in the cockpit is shown in Fig. 2, derived from Howard's domain concepts and extracted from Equation (6). These activities can be partitioned into three major domains. In the first one, the individual assessment domain, each pilot perceives the situation of the aircraft according to the information that is perceived by his senses. In the second domain, the mutual assessment domain, the pilots assess the aircraft status as to the course of actions that should be taken in the next domain, the action domain. This assessment is done through an interaction that includes both body language and verbal communication. In this domain, experience, mutual understandings, basic aviation knowledge, etc, as well as the relationship between the two pilots within the 'work culture' of the airline, are articulated. The course of action that will be chosen can either be mutually agreed or be unilateral. The captain predominantly carries out the latter, but not necessarily always so.

The fact that both pilots are exposed to the same incoming information and are able to carry out all required activities is illustrated by the incoming thick arrows and the 'mirror' image of the three domains. The incoming thick arrows represent both the feedback loop that is formed when the aircraft responds to the pilots' actions and input from external threats that emerge outside the cockpit, such as bad weather conditions, area topography, controllers' instructions, and others.

It is worthwhile to stress that the mutual assessment domain, in which, error correction processes take place, is the essence of the cockpit model. As illustrated, both pilots are performing two tasks concurrently; they both fly and monitor. In our view, the monitoring task is as important as the flying one and without it the present figure of civil aviation safety would not have been achieved[18]. Moreover, in our opinion, unless the monitoring ability is enhanced by adding technical devices and appropriate procedures, future improvements in flight safety are doomed to be, at best, marginal. This view is based on the fact that pilots, like any human being, do make errors and it is the responsibility of the designer to implement technical devices and appropriate procedures for error detection and recovery.

It should be noted that this model is similar to the cognitive 'procedural event analysis tool' (PEAT)[12,31] that is used by Boeing to investigate incidents and accidents. The difference between these models is a result of the different purposes of the two. While PEAT is used for understanding the individual pilot performance as governed by his or her long-term and short-term memory interactions, in our model the interactions of the two crewmembers is stressed. Thus in PEAT the individual assessment domain is described in detail, while in our model the mutual assessment domain is described in detail. It is our belief that the mutual assessment and the error correction processes are key factors in understanding the crew behaviour within the cockpit for the purposes of reducing fatal error rate.

5.0 DESCRIPTION OF THE NETWORK

The BBN that we have constructed is a result of a safety analysis of a population of flights. This class of model looks at the relationships between the various factors that influence the accident rate and it can be used for explaining various phenomena, as shown in

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

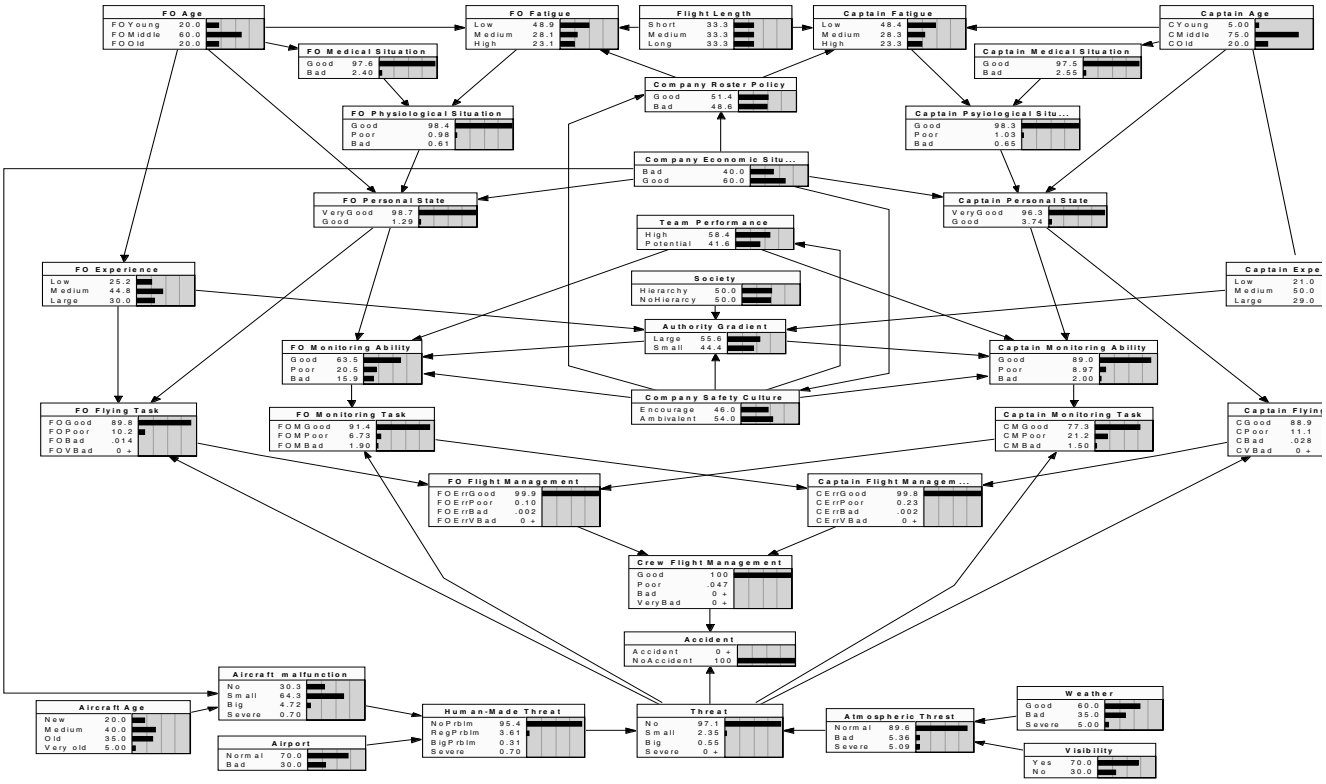


Figure 3. The flight model in BBN form.

the next chapter. As a population model of accidents, entire flights that abstract details of stages and special events are considered. Such flights are going to end safely or (extremely rarely) with a fatal accident. The node 'Accident' represents this last fact. The model was initially populated using global accident statistics and as such it becomes a benchmark for comparison purposes. It is noted again that at this stage the purpose of this paper is to show the feasibility of the model for assessing flight safety and not for practical purposes.

The description of the network is divided into two parts. The first describes the topology and the second discusses the assumptions that were made in the assignment of the NPT values.

5.1 The structure of the network

Figure 3 is a BBN representation of a flight. The network shows causal relationships among the factors that influence accident rate. As was discussed in the introduction, these factors include the personal condition of the pilots, airline policy and safety culture as shaped by sociological beliefs, and external threats. All these factors were identified as contributors to the development of accidents in general[32,20,21,22] and in aviation in particular[35,36]. Thus, excluding the accident node, the network can be divided into four basic parts:

- The Captain and First Officer: The functionality of the captain and the first-officer are represented in the upper right and left parts of the figure, respectively. These parts are actually a mirror image of each other and include all nodes starting with 'Captain' and 'FO' (First Officer).
- The mutual or context part: The middle upper part of the figure contains nodes that are either influencing or influenced by both pilots.
- The threats: The bottom of Fig. 3 represents the threats that might contribute to the occurrence of, or directly cause accidents. Four threat types are represented: malfunction of the aircraft, airport condition, weather, and visibility.

Notice that both the captain and the first officer are engaged in monitoring each other and no assumption of independence is made. However, as discussed in the introduction, the amount of the dependence is introduced through the appropriate NPTs.

5.2 The model in detail

In this section we describe the logic, the topology, and the main assumptions upon which the network is based. Space does not permit us to provide a detailed description of each NPT and its assigned values. Therefore, only main ideas will be discussed. Also, we will use pilot (in italic font) for referring to one of the pilots.

5.2.1 The Captain and First Officer

The main two parts of the network are devoted to the pilots' personal situation and activities. This partition is consistent with the observation that pilots are involved in all significant events that take place in a usual flight and, in particular, in the occurrence of most accidents[4]. In the network, each pilot's personal situation and capabilities are first established and only then the actual activities are modelled.

We start with the personal state of each pilot (nodes 'Captain personal state' and 'FO personal state'). These two nodes 'summarise' the ability or fitness of each pilot to carry out his/her

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

duties during the flight. As is seen, the nodes contain two levels of 'fitness', 'very good' and 'good'. This terminology implies that pilots are professional people and would not fly if they were really in an unfit state. The term 'good' (as opposed to 'very good') just means that the pilot is in a state that makes him more vulnerable to risks he might encounter in the flight and thus the accident probability is increased.

As is shown in the figure, the personal state of each pilot is directly influenced by his/her age, fatigue, physiological situation and the airline's economic situation. It is assumed that a bad economic situation might put the pilots (and the captain in particular) under some stress and affect their decisions along the flight . The physiological situation is a combination of the pilot's fatigue and medical situation, where the latter factor is dominant. Medical situation and the fatigue are also influenced by the pilot's age. Obviously, fatigue is also a function of the flight length and airline roster policy.

There are certainly other factors that may influence each pilot's personal state such as his basic character, temporary moods, etc. However, at this stage of the model, we assume that these factors are averaged across the large population of the world's pilots that create the accident rate 'background.'

In this model we assume a causal relationship between the pilot's age and experience, i.e. we assume that, from a generic point of view, experience is a function of age. This assumption is, however, arguable and in certain cases is not needed. For example, in a particular airline the distribution of pilots' age and experience are known and can be introduced directly into the appropriate nodes. In such cases the causal relationship between the two nodes can be omitted so that the 'pilot experience' nodes are changed into root nodes. It can also be argued that 'pilot experience' influences 'pilot personal state'. We, however, assume that this influence is small and can be neglected. We believe that experience is much more significant in influencing the flying management and the creation of the authority gradient[40].

As we go further down the figure we can see how the personal state of each pilot, together with other factors, is affecting the actual flight managementand monitoring capabilities. We assume that the 'team performance' and 'authority gradient' of the crew and the 'company safety culture' combined with the personal state of each pilot determine his ability to monitor the activities carried out by the other pilot. We further assume that the personal state has a similar and strong influence on the monitoring ability of both pilots. However, the influence of the authority gradient on each pilot is different; while authority gradient strongly affects the first-officer's monitoring ability, its influence on the captain is almost negligible. This difference can be seen in Tables 5 and 6 that are the NPTs associated with captain and first-officer monitoring abilities, respectively. In Table 5 which is associated with the 'captain monitoring ability', there is a total symmetry between eight upper rows (representing a 'large' 'authority gradient') and the eight lower ones (representing a 'small' 'authority gradient'). In Table 6 which is associated with the 'FO Monitoring Ability', there is no symmetry and a 'large' 'authority gradient' limits the maximum value of 'good' to 0·5.

To better understand these zx, the following is a short explanation. The probability values at the right side of the bolded lines in the table are those related to the configurations of the parent nodes given at the left side of that line. For example, in Table 5 the first row has to be read:

P(Captain Monitoring Ability = Good | Authority Gradient = Large and Team Performance = Yes and Company Safety Culture = Encourage and CaptainPersonal Situation = Very Good) = 1,

and,

P(Captain Monitoring Ability=Poor | Authority Gradient = Large and Team Performance = Yes and Company Safety Culture = Encourage and Captain Personal Situation = Very Good) = 0.

Table 5
Captain monitoring ability

Authority gradient	Team performance	Company safety situation	Captain personal situation	Bad	Poor	Good
Large	Yes	Encourage	Very Good	0	0	1
Large	Yes	Ambivalent	Very Good	0	0·05	0·95
Large	No	Encourage	Very Good	0	0·05	0·95
Large	No	Ambivalent	Very Good	0	0·1	0·9
Large	Yes	Encourage	Good	0	0·5	0·5
Large	Yes	Ambivalent	Good	0·01	0·54	0·45
Large	No	Encourage	Good	0·01	0·54	0·45
Large	No	Ambivalent	Good	0·01	0·59	0·4
Small	Yes	Encourage	Very Good	0	0	1
Small	Yes	Ambivalent	Very Good	0	0·05	0·95
Small	No	Encourage	Very Good	0	0·05	0·95
Small	No	Ambivalent	Very Good	0	0·1	0·9
Small	Yes	Encourage	Good	0	0·5	0·5
Small	Yes	Ambivalent	Good	0·01	0·54	0·45
Small	No	Encourage	Good	0·01	0·54	0·45
Small	No	Ambivalent	Good	0·01	0·59	0·4

Table 6
First-officer monitoring ability

Authority gradient	Team performance	Company	Captain safety situation	Bad personal situation	Poor	Good
Large	Yes	Encourage	Very Good	0·1	0·4	0·5
Large	Yes	Ambivalent	Very Good	0·3	0·3	0·4
Large	No	Encourage	Very Good	0·2	0·4	0·4
Large	No	Ambivalent	Very Good	0·4	0·4	0·2
Large	Yes	Encourage	Good	0·2	0·4	0·4
Large	Yes	Ambivalent	Good	0·3	0·4	0·3
Large	No	Encourage	Good	0·2	0·5	0·3
Large	No	Ambivalent	Good	0·5	0·4	0·1
Small	Yes	Encourage	Very Good	0	0	1
Small	Yes	Ambivalent	Very Good	0	0·1	0·9
Small	No	Encourage	Very Good	0·1	0·2	0·7
Small	No	Ambivalent	Very Good	0·05	0·15	0·8
Small	Yes	Encourage	Good	0·05	0·15	0·8
Small	Yes	Ambivalent	Good	0·1	0·2	0·7
Small	No	Encourage	Good	0·2	0·3	0·5
Small	No	Ambivalent	Good	0·01	0·3	0·6

It should be noticed that the two nodes 'pilot monitoring ability' represent their ability to monitor, not the monitoring task as performed during the flight. It is the two nodes 'pilot Monitoring Task' that represent the pilots' actual performances.

Four nodes, two for each pilot, represent their flying activities. The nodes 'pilot flying task' represent the activities before correction by the other pilot and the nodes 'pilot Flight Management' represent these activities after being corrected. The node 'Flight Management' represents the crew performance along the flight.

There are two main differences between the flying and monitoring tasks as expressed in the NPTs. While flying is an active task, monitoring is a vigilance task that is characterised by inactivity that is changed into activity when errors are caught.

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

Humans are much better at active tasks than inactive monitoring ones. The other difference stems from the fact that flying is a unitary task while monitoring requires interaction, at least when an error is being corrected. The representation of both tasks accounts for these differences. While monitoring is affected by the pilots' ability to form a good teamwork environment, flying is mainly affected by individual factors. Furthermore, while flying can directly cause or trigger accidents, monitoring is involved indirectly, therefore only flying nodes include 'very bad' states ('CVBad' for the captain and 'FOVBad' for the first-officer).

It is assumed that both the flying and the monitoring activities of each pilot are influenced by the threat. The exact influence of the threat on the two tasks is different, however. It is assumed that, as the threat increases (especially for 'big' and 'severe' threats) each pilot is forced to concentrate in his own (flying) tasks, thereby significantly decreasing his ability to monitor the other pilot. Flying, on the other hand, is generally assumed to improve as the threat increases, as long as the pilot's personal state is 'very good'. The flying improvement is reduced for pilots that are in state 'good'. A last comment on flying as represented in the figure: it is assumed that when a pilot realises that he has committed an error, he will correct it or compensate for it promptly and usually before the other pilot will. In terms of the network, this last comment means that the flying task nodes include two 'minor' steps flying and correcting errors or compensating for them.

The last node that sums the crew 'flight management', directly affects the accident probability. We use the term 'very bad' to describe the worst level of the flying management of each pilot as well as of the crew. This term doesn't imply that an accident will necessarily occur; it merely means that the performance is so bad that the accident probability is increased significantly. Similar arguments are also used for explaining that the probability of being involved in an accident given a 'Good' flight management is still not zero.

5.2.2 The contextual factors that impinge on both pilots

The influence of some context nodes on the pilots was already discussed. Here is a short discussion on the exact relationships among these contextual factors. We start with the safety culture of the airline or its assembly of characteristics and attitudes towards safety. In our case, this loose definition is translated to the way an airline (especially its managerial ranks) treats safety related issues such as cockpit procedures, policies, training, maintenance, capital investment and so on. Safety culture affects the selection of a roster policy and thus indirectly affects pilot fatigue, team performance or the ability of the two pilots to cooperate in their cockpit tasks and the authority gradient. The latter two factors influence the monitoring tasks of the two pilots.

Authority gradient has been recognised as a strongly influencing factor in the ability to monitor, especially as performed by the first-officer. As seen, in addition to the safety culture the authority gradient is influenced by the attitudinal values within the pilots' society towards authority figures and the difference between the pilots' experiences. Between these three factors we assume that pilots' sociological values are dominant. This dominance has already been shown even in non-hierarchical societies as reported by Wiley in *Business & Commercial Aviation*[40].

Although team performance is less influencing than the authority gradient, it has already been identified as a factor in almost every task that is carried out by more than one person. Katzenbach and Smith[19] describe team performance in a plot with five degrees out of which we consider the 'potential' and 'high performance' ones. While the first term (i.e. 'high') is clear, the second one represents a team that its ability to cooperate is relatively low and their 'theoretic' potential to carry out the mutual tasks has not been materialised. In an airline with an 'Ambivalent' safety culture the team is created randomly, taking no account on the ability of the crew members to cooperate. In an encouraging safety culture this factor is considered[5]. And a last comment on team performance; although in this paper we concentrate on the two pilots, the actual team includes also the stewards and even the passengers[11].

Strict regulations do limit the influence of the airline's economic situation on maintenance policies, cockpit procedures and other practical issues. However, economics have certain important unconscious or irrational influences, which are difficult, if not impossible to ignore. A bad economic situation was suspected as a source of pressure on maintenance personnel and pilots. This was demonstrated in the crash of flight 1420 in Little Rock, Arkansas. Contrary to regulations, the pilots, the captain in particular, attempted to land late at night in severe thunderstorms and strong wind shear alerts[27]. A bad economic situation may also bias the airline policies into modifying procedures to sacrifice some safety in favour of saving costs. This explains the relationship between the economic situation of an airline and other factors as seen in Fig. 3. An example of this phenomenon is the fear of lay-offs by pilots and maintenance personnel may lead to 'round-edges' in their work.

5.2.3 The threats

The description of the threat is straightforward and only three points deserve some elaboration. Firstly, the nodes 'airport & aircraft' and 'visibility & weather' are intermediate nodes that facilitate the formation of the 'threat' NPT. Secondly, the 'Aircraft Age' and the 'company economic situation' influence the probability of an 'aircraft malfunction'[8]. And thirdly, both nodes 'aircraft age' and 'aircraft malfunction' include four states with the last one indicating extreme cases that are indeed rare in regular airlines but can be found in bad ones. 'very old' aircraft that strongly affect safety are indeed rare, but can't be totally ignored especially in airlines that are in a really bad economic situation. A 'severe' malfunction (or worse) is the level of say a major failure in the aircraft structure such that an accident is almost unavoidable. This situation is certainly very rare but has occurred[8]. It should also be mentioned that this last 'severe' state is different from the one that belongs to the node 'weather'. This difference is a result of certain regulations that limit flights to 'manageable' weather conditions, while no rule can avoid major breakdowns.

5.2.4 Accidents and last comments

The last node 'Accident' is the 'climax' of the network. It gives the probability of ending the flight safely or in a fatal accident. Consistent with Boeing findings[4] that pilots are involved in about 70% of accidents, it is assumed that crew flight management is the dominant factor. However, as was mentioned, even very bad flight management does not necessarily lead to an accident. In this first model only the combination of very bad flight management and a severe threat will eventually lead to an accident.

As mentioned the model is based on some arguable assumptions. These assumptions include causal relationships that were neglected, relationships that should have been changed and similar others. There are probably also certain aspects such as training policies, automation types in the piloting system and similar others that were not included in the model. It is not that we consider these factors and extra relationships unimportant. We do believe that at this stage of the model it is better to stick to the 'principle of parsimony'[10] and concentrate on a limited number of factors. Filar further explains that attempts to refine immature models, as this one, may result in increased complexity and inaccuracies.

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

6 0 THE MODEL AND THE PRESENT KNOWLEDGE

In this section we intend to support our claim that BBNs are appropriate for safety assessment by replicating certain global air transport accident data. As mentioned, global data and large-scale experiments are the only reliable information sources.

6.1 Accident rate and differences between airlines

The first and most important source is the accident rate in terms of accidents per departure. The model predicts an accident rate of $2{\cdot}015{\cdot}10^{-6}$, which fits the world accident rate of $\sim 2{\cdot}10^{-6}$[1].

The model replicates another qualitative observation that deals with the captain and first-officer as contributors to accidents. Accident investigations indicate that the captain is more dominant than the first-officer in both dictating the flight process and in contributing to the occurrence of fatal accidents. According to the model, the accident probabilities given very bad 'pilot Fflight management' by the captain ('CErrVBad') or the first-officer ('FOErrVBad') are 0·329 and 0·085, respectively.

Accident statistics further show that there is a significant difference between the various airlines. Accident probability of 'good' airlines is in the range of one to two orders of magnitude better than accident probability of the 'bad' airlines[30]. Using differences in contextual and environmental factors that influence the performance of pilots, airliners and airports, as detailed in Table 7, we succeed in replicating this finding. According to the model, the accident probability of 'bad' airlines is $1{\cdot}416{\cdot}10^{-5}$ and $8{\cdot}23{\cdot}10^{-7}$ for 'good' ones: a ratio of 17·2:1.

Table 7
Assumptions made for 'good' and 'bad' airlines

Node	'Bad' airlines	'Good' airlines
Captain personal state	Good	Very Good
First Officer personal state	Good	Very Good
Company economic situation	Bad	Good
Society	Hierarchy	No Hierarchy
Company safety culture	Ambivalent	Encourage
Aircraft age	New (0·5) Medium (0·25) Old (0·3) Very old (0v4)	New (0·25) Medium (0·4) Old (0·35) Very old (0)
Airport	Bad (0·6) Normal (0·4)	Bad(0·1) Normal (0·9)

6.2 General and fatal errors

A series of Line Operation Safety Audits (LOSA) made by Helmreich and his colleagues[15,16] have shown that the number of pre-defined errors made by the pilots is in the range of 4-13 per flight. These audits further propose a descriptive model that shows a process that 'evolves' these errors into accidents. Their model, however, fails to explain the large gap of 7-8 orders of magnitude between the error rate and accident rate. Nor does any other model of which the authors of this paper are aware succeed in bridging this gap[†]. In this section we can show that our model succeeds in replicating this large gap.

It is understood that pilot error can be divided into two main categories; the first includes inconsequential errors that cannot lead to accidents, and the second is composed of more dangerous ones that may cause accidents, given that certain conditions exist. The model accounts for this difference by using four levels of errors (states that belong to the nodes 'Captain flight management', 'FO flight management', and 'threat'), out of which only the fourth level (states 'CErrVBad', 'FOErrVBad' and 'Severe') has a significant probability of causing accidents. As all LOSA flights ended safely, it must be that errors and other failures that belong to the fourth level did not happen. In the model, we 'express' this assumption by nullifying the values of states 'CErrVBad', 'FOErrVBad' and 'severe', and get an accident probability of $1{\cdot}67{\cdot}10^{-7}$ accidents/departure[3]. Using Helmreich's findings of error rate of 4-13 errors/departure, we get an accident rate in the range of $6{\cdot}68 \times 10^{-7}$–$2{\cdot}17 \times 10^{-6}$ accidents/departure. This result is still very low and means that the probability of these errors leading to an accident is indeed very low and still in the range of the general accident rate.

6.3 Pilot and First Officer monitoring evaluation

We now turn to another finding of LOSA made by Thomas[38]. According to this audit, the captain detects 44·4% of the errors made by the First-Officer, while the First-Officer detects only 36·2% of the errors made by the captain. This finding means that the captain is (44·4/36·2) = 1·2 times better than the First-Officer in catching errors.

This ratio can be evaluated by assessing the improvement of each pilot when monitored by the other one. In the model, the ratio between the two pairs of nodes 'pilot flying task' and 'Pilot Flying Management' expresses this improvement. Specifically, for each pilot, the improvement is the ratio of his error handling while monitored 'badly' (i.e. monitoring task in 'bad' level) and 'regularly' (i.e. monitoring task with no prior knowledge) by the other pilot.

Strictly speaking, it is not possible to draw an inference given the size of the sample represented by LOSA data with respect to the accident rate. Nevertheless, it is interesting to pursue the topic, taking as prior knowledge the fact that no accident has occurred in any LOSA flight to date. This prior knowledge can be introduced into the comparison in two ways:

- Nullifying the accident probability.
- Making this comparison on the states that represent performances of 'averaged' pilots. In our scale of pilots' performance, this is represented by the 'Poor' states ('bad' or 'very bad' performances cannot represent the 'safe' flights as demonstrated in the LOSA. And 'Good' performance means that the number of errors is significantly lower than the average).

The results of this discussion are presented in the Table 8, for which explanation follows. As can be seen, Table 8 is partitioned into three parts that are separated by blank rows. The upper part presents the improvement of the captain in handling errors while the First-Officer monitoring 'regularly' and 'badly'. The middle part presents the same information about the First-Officer. Thus, for example, the first two rows for the column 'Poor' in the captain section has to be read as:

$$P(\text{Captain Flight Management=Poor} \mid \text{First-Officer Monitoring Task='Regular'}) = 2{\cdot}345{\cdot}10^{-3},$$

and

$$P(\text{Captain Flight Management=Poor} \mid \text{First-Officer Monitoring Task=Bad}) = 4{\cdot}378{\cdot}10^{-3}.$$

The third row in each section represents the improvement of each pilot that results from the other pilot monitoring, or the ratio of the above two results. The last section of the table is the ratio of the improvements of the two pilots. It can be seen that this ratio for the 'poor' column is in close agreement with Thomas's findings.

[†] One author of this paper (arris) [14] has attempted to bridge this large gap by taking into account only errors that lead to an "Undesired Aircraft State" (a LOSA-defined event), which is a necessary, but certainly not sufficient, precursor to an accident, but this still leaves unexplained a gap of about 2-3 orders of magnitude.

In mathematical notation this probability is given in the expression:

$$P(\text{Accident} \mid \text{CerrVBad} = 0 \text{ and FOErrVBad} = 0 \text{ and Severe} = 0) = 8{\cdot}81{\cdot}10^{-8}$$

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

Table 8.
Monitoring task comparison between the captain and First Officer

Captain improvement

First Officer monitoring	Very Bad	Bad	Poor	Good
Regularly	5·205E–09	2·363E–05	9·976E–01	9·945E–01
Badly	2·614E–07	1·102E–03	4·378E–03	2·345E–03
Improvement	50·22	46·63	1·867	0·997

First Officer Improvement

First Officer monitoring	Very Bad	Bad	Poor	Good
Regularly	2·852E–10	1·676E–05	1·042E–03	9·989E–01
Badly	1·732E–08	1·099E–03	2·255E–03	9·967E–01
Improvement	60·72	65·58	2·165	0·998

Ratio Captain/First Officer

Ratio	1·21	1·41	1·16	1·00

7.0 USING THE MODEL

In the introduction to this paper we wrote that airlines, regulatory bodies and aircraft manufacturers could use a quantitative model for aviation safety. In this section we show such hypothetical uses. Again we would like to stress the point that this first version of the model is incomplete and is not suitable for practical purposes and that the results obtained in this section are illustrative only. As mentioned, the present form of the model is not final and changes will need to be made depending on the particular use. We stress again that the aim of this paper is to show the feasibility of the model for practical uses by various bodies and for many purposes.

7.1 Causes of accident

The first issue that we want to explore is the main causes of accidents. Boeing's findings have already been mentioned and according to them the crew is involved in about 70% of fatal accidents. Here we intend to explore this finding more thoroughly and search for other major causes. We will conduct this exploration by using the top-down approach, i.e. assume that an accident has happened and then backtrack. In the terminology of the model, this is equivalent to setting the 'Accident' node to state 'Accident' (that is, in the node 'Accident'), p(Accident) = 1 and p(NoAccident) = 0 and analysing the changes in state values of the ancestor nodes. As before, we are interested in those states that might lead to accidents, i.e. the worst states of each node. Table 9 demonstrates this analysis where (a) the column 'Regular Values' means the values of the 'worst states' when no assumption is made, (b) the values in the 'Accident Occurred' column are those values assuming an accident has occurred, and (c) the 'Ratio' column gives the changes of the above values.

Table 9
Contribution of threats to accident probability

Node	Worse state	Regular value	Accident occurred	Ratio
Captain Flight Management	CErrVBad	$7{\cdot}76{\cdot}10^{-9}$	$1{\cdot}27{\cdot}10^{-3}$	$1{\cdot}63{\cdot}10^{5}$
FO Flight Management	FOErrVBad	$3{\cdot}12{\cdot}10^{-10}$	$1{\cdot}32{\cdot}10^{-5}$	$4{\cdot}22{\cdot}10^{4}$
Aircraft Malfunction	Severe	$7{\cdot}00{\cdot}10^{-3}$	0·532	75·99
Airport	Bad	0·3	0·31	1·03
Weather	Severe	0·05	0·96	19·8
Visibility	No	0·3	0·32	1·04
Aircraft Age	Very Old	0·05	0·56	11·2

It is easy to see that the largest three changes belong to the nodes 'Captain flight management', 'FO flight management' and 'aircraft malfunction'. The first two are not a surprise as they represent the pilots. The third largest change is also not a surprise since, according to the discussion near the end of Section 5.2.3, this represents a major breakdown of the piloting system. The largest change after these is the weather.

The previous discussion shows, therefore, that in addition to the pilots, major malfunctions and severe weather conditions are also significant contributors. Both regulatory bodies and airlines can use this analysis for issuing better maintenance and weather related procedures. Furthermore, airlines may also consider allocating resources and more investment to improve maintenance and reduce malfunction probabilities. The introduction of better weather-related procedures and instrumentation for ground controllers is another outcome of this analysis. Aircraft producers might also use this information for introducing better maintenance related systems and detection systems for bad weather. The former systems may enable the reduction of maintenance costs of airlines, thereby allowing more investment or profit.

Another interesting outcome of the table can be deduced from the last row. It infers that the major contributor to aircraft malfunction is its age. This is a very important outcome since it indicates either that the use of very old aircraft should be restricted or some more strict maintenance rules should be applied.

7.2 Economic situation

To what extent does the economic situation of airlines influence accidents? This question is very important especially in the last decade when many airlines report a decline in their profits. This influence can be quantified by comparing accident probability of airlines with 'good' and 'bad' economic situations. Setting the above states of the node 'Economic Situation' results in accident probabilities that range from $1{\cdot}7{\cdot}10^{-6}$ for a 'good' economic situation to $2{\cdot}5{\cdot}10^{-6}$ for a 'Bad' economic situation (50% increase). This information is very valuable and can be used for evaluating depth and frequency of periodical and random inspections by regulatory bodies.

7.3 Monitoring or flying improvement

The model can also be used in many areas of decision-making. An example of such decisions at an airline or producer levels could be whether to introduce better facilities for monitoring or flying. Setting nodes 'pilot flying task' to state 'CGood' and 'FOGood' we get an accident probability of $1{\cdot}85{\cdot}10^{-6}$. Similarly setting nodes 'pilot Monitoring Ability' and 'pilot Monitoring Task' to their best states we get an accident probability of $4{\cdot}60{\cdot}10^{-8}$. Obviously, creating a better monitoring environment and other monitoring facilities is far more beneficial.

8.0 CONCLUSIONS

In this paper we have demonstrated in principle that BBNs can be used for evaluating accident probability of large passenger aircraft. We have also shown how such a model can be used for practical applications. Account is given to environmental factors that influence both the flight-crew and external threats. Temporary and permanent properties of each pilot are also modelled as to their

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

influence on the crew to form a working and cooperating team during the flight. All these factors are used in a structured way that starts with the individual pilot and progresses through their mutual monitoring capabilities to their flight management.

The main purpose of this paper is to demonstrate the feasibility of obtaining quantitative results in the general area of aviation accidents. This is a possibility and we do believe that similar models, based on BBN, could be developed for other safety critical areas.

The model succeeds in replicating statistical factors that were obtained from experiments and 'real life' data. This replication builds a quantitative tier on top of the qualitative explanations of accidents, and as more findings are replicated more confidence is built. As demonstrated, we do believe that such models can become a reliable tool for predicting influence of airline policy changes and even in orienting investments at this level. Furthermore, regulatory bodies can use this model in their long-term plans, decision-making procedures and goals setting.

In this model the crew members are assumed to operate as a team. This assumption is an evolution from previous quantative approaches[18,13] in which pilots were modelled as independent subsystems. The amount of cooperation, the pilots' ability to correct errors and to improve safety, depends on organisational and sociological factors as well as personal ones. The model indicates the value of safety-minded companies creating a safety culture that enhances cooperation and monitoring procedures, thereby reducing accident rates.

Another deviation from previous probabilistic models is the distinction between the captain and the first officer. Equality has been used for ease of analysis in simpler models but this does not fit the knowledge gained from aviation accident records[13]. In accordance with this knowledge, the present model gives more weight to the captain than to the first officer. This weight is expressed in both contributing to and preventing accidents.

This work can be expanded and continued in two directions. The first direction has already been discussed and deals with the development of similar models in other safety critical areas. We do believe that even if a general quantitative model cannot be composed, area-specific models are achievable and are potentially very valuable.

The other direction deals with a deeper study of this area in order to qualify this model for practical uses. This direction will probably necessitate both restructuring the net in order to account for influences that were neglected at this stage and introducing other influences on the pilots' behaviour, such as training policies, cockpit facilities, etc. These other aspects were not used in this early model in order to avoid more complexity and inaccuracy[10]. An improved net will also necessitate more accurate data for the construction of NPTs. Another aspect of this expansion direction includes refining the model to be applicable to different flight stages as done with earlier models[18]. This refinement may reveal many other aspects of pilots' behaviour for differing workloads.

REFERENCES

1. AirDisaster, viewed 15 August, 2004, http://www.airdisaster.com/statistics/yearly.shtml.
2. GORE, A, Vice President (Chairman) Final Report to President Clinton. White House Commission on Aviation Safety and Security, 12 February 1987. Viewed 28 May 2004. http://avsp.larc.nasa.gov/WHC_Aviation_Safety.pdf, 1987.
3. BOUISSON, M., MARTIN, F. and OURGHANLIAN, A. Assessment of a safety critical system including software: A Bayesian belief network for evidence sources, 1999 Proceedings of the Annual Reliability and Maintainability Symposium, 1999, pp 142-150.
4. Boeing, Statistical Summary of Commercial Jet Airplane Accidents, World Operations 1959 – 2002, 2003.
5. SHI-LIE, C. (GARY) and LI, L. Modelling team member characteristics for the formation of a multifunctional team in concurrent engineering, IEEE Transactions on Engineering Management, 2004, **51**, (2), pp 111-124.
6. COOK, S.C., HARRIS, D.D. and AL-ZUBAIDY, S. A systems approach to modelling the piloting system in multi-crew aircraft. In proceedings of the 10th Australian International Aerospace Congress and 14th National Space Engineering Symposium, 29/7/03-1/8/03, 2003, Brisbane, Australia.
7. DEGANI, A. and WEINER, E.L. Procedures in complex systems: The airline cockpit, IEEE Transactions on Systems, Man Cybernetics – Part A, Systems Humans, **27,** (3), pp #02-312, 1997.
8. FAITH, N. *Black Box*: *The Further Investigations*, Boxtree, Great Britain 1998.
9. FIDEL, R. Collaborative Information Retrieval (CIR), viewed January 15, 2004, http://www.ischool.washington.edu/cir/ISICtextWeb.doc, 2003.
10. FILAR, J.A. Mathematical models, In: knowledge for sustainable development: An Insight into the Encyclopaedia of Life Support Systems, UNESCO/EOLSS, released at the World Summit on Sustainable Development, September 2002, Johannesburg, South Africa.
11. GINNETT, R.C. Crews as groups: their formation and their leadership. In WIENER, E.L., KANKI, B.G. and HELMREICH, R.L. *Cockpit Resource Management*, 1995, Academic Press, pp 71-98.
12. GRAEBER, R.C. and MOODI, M.M. Understanding flight crew adherence to procedures: The procedural event analysis tool (PEAT). Report of Boeing Commercial Airplane Group Seattle, Washington, USA Safety seminar Capetown, South Africa November 17-19, 1998.
13. GREENBERG R., COOK, S.C. and HARRIS, D.D. A probability model of the piloting system in large passenger transport aircraft. In proceedings of the 10th Australian International Aerospace Congress and 14th National Space Engineering Symposium, 29/7/03-1/8/03, 2003, Brisbane, Australia.
14. HARRIS D.D. Quantifying accident precipitating events, To be published, 2005.
15. HELMREICH, R.L., KLINECT, J.R., and WILHELM, J.A. Models of threat, error, and CRM in flight operations. Proceedings of the 10th International Symposium on Aviation Psychology Columbus, Ohio 3-6 May, 1999.
16. HELMREICH, R.L., KLINECT, J.R. and WILHELM, J.A, Threat and error management: Data from line operations safety audits. Proceedings of the 10th International Symposium on Aviation Psychology Columbus, Ohio 3-6 May 1999.
17. HOPKINS, H. *Managing Major Hazards: The Lessons of the Moura Mine Disaster*, Allen & Unwin, 1999.
18. HOWARD, R.W. Planning for super safety: The fail-safe dimension, *Aeronaut J*, 2000, **104**, (1041), pp 517-555.
19. KATZENBACH, J.R. and SMITH, D.K. *The Wisdom of teams*, Boston, MA. 1993. The actual discussion is brought in SNOOK, S.A. *Friendly Fire*, Prinston University Pres, 2000.
20. LEVESON, N.G. *Safeware: system safety and computers*, Addison-Wesley, 1995.
21. LEVESON, N.G. An accident model for engineering safer systems, *Safety Science,* 2000, **42**, pp 237-270.
22. MANION, M., and EVAN, W.M. Technological catastrophes: their causes and prevention, Technology in Society, 2002, **24**, pp 207-224.
23. MOORE, W.H. The grounding of Exxon Waldez: An examine of the human and organizational factors, *Marine Technology*, 1994, **31**, (1), pp 41-51.
24. NEAPOLITAN, R.E. *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*, John Wiley & Sons, Inc 1990.
25. NEIL, M., FENTON, N., FOREY, S. and HARRIS, R. Using Bayesian belief networks to predict the reliability of military vehicles, *Computing & Control*, *Engineering J*, 11-20 February 2001.
26. NETICA, Details on NETICA can be found in www.norsys.com.
27. NTSB Aircraft Accident Report, Runway Overrun During Landing, American Airlines Flight 1420, McDonnell Douglas MD-82, N215AA, Little Rock, Arkansas, 1 June, 1999. Viewed, 3 September, 2004, http://www.ntsb.gov/ Publictn/2001/AAR0102.pdf. 1999.
28. PARK, J., KIM, J. and JUNG, W. Comparing the complexity of procedural steps with the operators performance observed under stressful conditions, *Reliability Engineering and System Safety*, **83**, pp 79-91, 2004.
29. PEARL, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, 1991.
30. PlaneCrashInfo, This figure along with other important statistical information can be found in http://www.planecrashinfo.com/cause.htm.
31. PROCTOR, P. Boeing safety tool provides insights into human factors errors, *Aviation Week Space and Technology*, 21 June 1998.
32. REASON, J. *Human Error*, Cambridge University Press, 1990.
33. REASON, J. *Managing the Risks of Organizational Accidents*, Ashgate Publishing Limited, 1997.

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press

34. ROTH, E.M., MUMAW, R.J. and LEWIS, P.M. An empirical investigation of operator performance in cognitively demanding simulated emergencies, NUREG/CR-6208, Washington DC, US Nuclear Regulatory Commission (1992).
35. SHAPPELL, S.A. and WIEGMANN, D.A. The Human factors analysis and classification system-HFACS, Report DOT/FAA/AM-00/7 of the Office of Aviation Medicine, FAA, 2000.
36. WIEGMANN, D.A. and SHAPPELL, S.A., A human error analysis of commercial aviation accidents using the human factors analysis and classification system (HFACS), Report DOT/FAA/AM-01/3 of the Office of Aviation Medicine, FAA, 2001.
37. SVENSON, O. The accident evolution and barrier function (AEB) Model Applied to Incident Analysis in the Processing Industries, *Risk Analysis*, 1991, **11**, (3), 499-507.
38. THOMAS, M.J.W. Improving organisational safety through the integrated evaluation of operational and training performance: An adaptation of the Line Operations Safety Audit (LOSA) methodology. *Human Factors and Aerospace Safety*, 2003, **3**, (1), pp 25-45.
39. TOLSTYKH, V. Brief Summary of incident reporting system events with procedural deficiencies, Proceedings of a Specialists Meeting on Operating Procedures of Nuclear Power Plants and Their Presentations, International Atomic Energy Agency, 1992.
40. WILEY, J. Right talk from the right seat, *Business and Commercial Aviation*, July 2001, pp 48-54, 2001.

https://doi.org/10.1017/S0001924000000981 Published online by Cambridge University Press