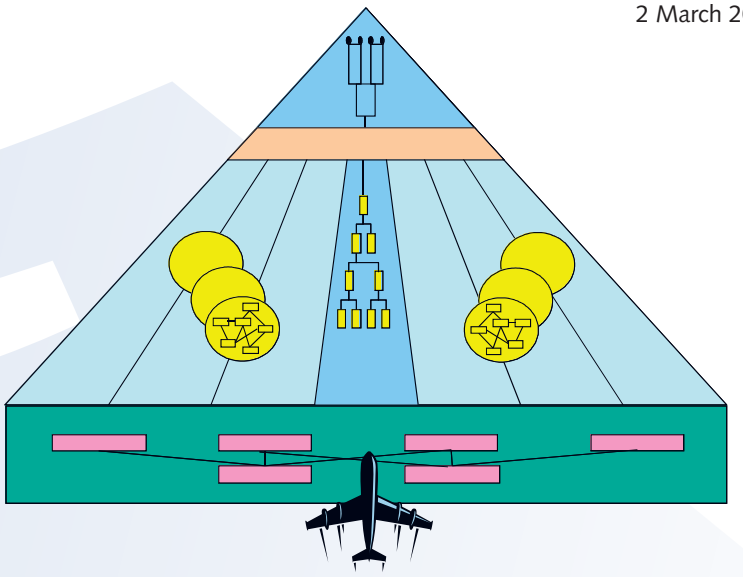
Causal Model for Air Transport Safety

Final report

Causal Model for Air Transport Safety

Final report

2 March 2009



"CATS is the second best representation of reality, reality itself being the best"

Prof dr Patrick Hudson

Preface

The Netherlands Ministry of Transport, Public Works and Water Management (VenW) continuously pursues the improvement of transport safety. To this aim, the Directorate General of Civil Aviation and Maritime Affairs (DGLM) also invests in innovative safety research. It gives me great pleasure to present the results of such a research project, the Causal Model for Air Transport Safety, or CATS in short. Over a period of three years, a multi-disciplinary consortium of experts and organisations under the supervision of Delft University have further developed the CATS scientific framework and database into a working application. The enclosed report describes the model's structure and potential application areas.

Historical Background

Air traffic at Amsterdam Schiphol Airport grew rapidly during the nineties. The Airport planned a territorial expansion in combination with a new runway. In October 1992, a Boeing 747 freighter crashed into an apartment complex in Amsterdam. This led to great public concern and debate on the issue of Third Party Risk of residents, better known as 'external safety' in the Netherlands. As a result a formal external safety policy was implemented in Dutch law, based on the calculation of the risk of potential airplane crashes near airports. A static statistical model for third party risk was developed, using generic accident probabilities, to calculate individual risk contours and societal risk curves. The risk contours are used to identify areas to which residential and commercial land use planning restrictions apply. The societal risk identifies the probability of numerous casualties on the ground in airplane accidents. As the statistical model only uses generic accident scenarios, the underlying causes of accidents are not considered. This makes it almost impossible to identify specific measures, for instance for Amsterdam Schiphol Airport, to improve safety at source, i.e. the aviation system. In contrast, a causal model includes the underlying causes of accidents.
To control the societal risk of residents on the ground, while at the same time providing possibilities for airport expansion, the Dutch Parliament in 2001 amended the Aviation Act requiring the Ministry of Transport to develop a causal model for aviation safety. The initial results however made it clear that a causal model for aviation alone does not suffice to adequately control societal risk, as societal risk is also determined by the growth of the population. However, it also became clear that causal modelling does provide practical insights for improving aviation, or internal safety and understanding the related aviation risks. Therefore in 2004, the Aviation Act was further amended to reflect the intention of developing a causal model for air transport safety with a strong focus on internal aviation safety, rather than external safety. This report presents the results of this effort.

Integrated Tool for Safety Management Systems

ICAO introduced safety management as a means of controlling and further improving aviation safety. Risk management was introduced as part of this concept. Aviation processes now have reached such a degree of complexity that traditional analytical methods are no longer able to deal with the entire aviation system spectrum. The CATS model provides insight into cause-effect relationships in the event sequences leading up to potential incidents and accidents. These event sequences cover all potential failure modes of the operations during the different gate to gate flight phases. CATS enables quantitative risk assessments of existing and new operations to be carried out, while providing

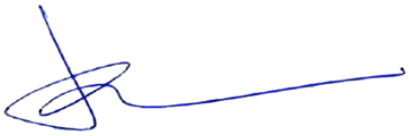
insight into the effectiveness and efficiency of risk-reducing measures. CATS is a meaningful addition to the conventional safety management tools that are based on organisation structures. Furthermore it also meets the new ICAO Safety Management Systems' requirements.

Application

Since it describes the gate to gate risks inherent in the complete aviation system, CATS can in principle be used by different aviation organisations to tackle different problems. The model provides a strategic perspective of aviation risk that can be used as input into the policy formulation process. This insight can be used to achieve a proper balance between alternative solutions to safety issues and policy can be focused on the most promising measures. For policy purposes, understanding minor risks can be important as well, since incidents may focus attention on minor risks making them appear more significant than they really are. Pursuant to the adopted causal modelling regulation, the Ministry intends to use the model as a support tool for implementing a regulatory risk-based safety oversight system. Such a system will help streamline the resources required to efficiently conduct safety oversight without compromising the effectiveness of the safety improvement process.
An essential element of the new safety management approach is for aviation sector organisations to take responsibility for safety. The model can support the discharge of this responsibility and be used as a tool to understand the risk of air operations and possible improvements by implementing risk mitigating measures. We hope that this effort will also inspire foreign aviation authorities to implement a similar system.

Finally, I would like to express my appreciation to the researchers who have achieved this result. Also I would like to thank the group of national and international participants that supported the project. I specially would like to show my appreciation to the FAA and Eurocontrol for their research contribution to this project.

Sincerely,



M.E.P. Dierikx
Director General of Civil Aviation and Maritime Affairs

Table of contents

Preamble

The purpose of this project was to create new ways of finding the complex causes of air transport accidents and to help in their prevention. The development of a Causal Model for Air Transport Safety, CATS, has been a journey of discovery for the members of the development team who contributed to various parts of the project, some from start to finish. At the beginning in 2005 only the classical tools of quantified risk analysis were available. Now at the end in 2008 a whole new level of modelling has been achieved. The use of a fully integrated "directed acyclic graph" made it possible to model much more realistically the complex interactions between the technical and organisational systems that make an aircraft operate and sometimes fail. In order to achieve this, close cooperation was needed between mathematicians, computer programmers, data analysts and aeronautics experts.

The Ministry of Transport and Water Management of the Netherlands embarked on a difficult scientific project. The result is a new generation of modelling and an immediately useable application. The result can also form the basis for further development using the full power of this new technology.

There were stimulating discussions with the *Core Group in the Ministry of Transport* (V&W): Andre Muyselaar, Hok Goei, Frederik Demeyere, Michael Portier (NIVR/To70) the *CATS Advisory Group*: Joy Oh (SZW), Jan Busstra (V&W) Michael Portier (NIVR/To70), Rob van der Boom (V&W), Henk van Leeuwen (NIVR), Job Brüggen (LVNL), Bart de Vries (KLM), Jos Wilbrink (IVW), Erik Lagerweij (AAS), Eric Perrin (Eurocontrol), John Lapointe (FAA), Graham Greene (CAA), Werner Kleine-Beek and Illmar Bilas (BMVBW), Stephane Deharvengt (Aviation Civile), John Vincent and Michel Masson (EASA) and the *CATS Group of Experts:* Rob van der Boom (V&W), Rudi den Hertog, (Stork-Fokker services), Ali Mosleh (University of Maryland), Patrick Hudson (University of Leiden), Michel van Tooren (TU Delft), Pietro Carlo Cacciabue (JRC), Thomas Bos (VNV), Jurgen van Avermaete (LVNL), Bert Kraan, Arthur Dijkstra (KLM), Adrian Young (AAS) Hans Offerman (LVNL).

We especially thank the client, Rob van der Boom of the Ministry of Transport and Water Management, for his support and patience.

Delft, Amsterdam, London
31 July 2008

Authors:	Programmers:	Researchers
Ben Ale, TU Delft (Project Manager)	Dan Ababei, TU Delft	Yan Chan, DNV, UK
Linda J. Bellamy, White Queen BV	Daniel Lewandowski, TU Delft	Tina Florentina, TU Delft
Roger Cooke, TU Delft	John Cooper, JPSC Ltd, UK	Coen van Gulijk, TU Delft
Martina Duyvis, NIFV		Hans de Jong, NLR
Dorota Kurowicka, TU Delft		Kasia Krugla, TU Delft
Pei Hui Lin, TU Delft		Gavin Osborn, DNV, UK
Oswaldo Morales, TU Delft		Job Smelting, NLR
Alfred Roelen, NLR		
John Spouge , DNV, UK		

Executive summary

The Ministry of Transport and Water Management (V&W) commissioned a causal model of air transport safety. The model was developed by a consortium including Delft University of Technology (TUD), National Aerospace Laboratory (NLR), White Queen Safety Strategies (WQ) and the National Institute for Physical Safety (NIFV) in The Netherlands, and Det Norske Veritas (DNV) and JPSC consulting in the UK. The project is known as Causal Model of Air Transport Safety (CATS).

The motivation for the project is the need for a thorough understanding of the causal factors underlying the risks of air transport and their relation to the different possible consequences so that efforts to improve safety can be made as effective as possible.

The project as a whole is defined in the program of work from the consortium manager, TUD Risk Centre (23 August 2004). The project started on 14 July 2005. The project ended on 31 October 2008. The specified goals of the project were completed while full validation and further development of management influences remain subjects for further work as described later in this report.

The Ministry expressed the following potential uses of the model:

- To enable comparative judgements
  - With other means of transport
  - Over time
  - Between airports
  - Between links in the chain of safety
- To diagnose risk situations within and outside one's own organization to enhance the safety of air transport proactively
- To prioritise potential safety measures (e.g. on the basis of the expected effectiveness)
- To support supervision of the quality of safety delivered by the air transport system and its parts
- To inform the public on risk policy and risk reduction efforts.

The Ministry wanted an interactive and collective approach leading to an authoritative model, that advances the frontier of science and which would allow a comprehensive analysis of safety problems in air transport. This model would provide increased knowledge and add an management tool to the current instruments of the Ministry. In 2003 the strict requirement for proof of a non-increasing disaster potential was relaxed and is no longer a legal requirement. This took the development of the causal model out of the political arena and out of the sometimes heated debate between the different parties involved.

The objective of the CATS project was to develop a fully operational causal model that represents the causes of commercial air transport accidents and the safeguards that are in place to prevent them, building on the experience gained in the demonstration causal models developed by DNV and a consortium led by NLR during 2001-2002. (DNV, 2001; Roelen et al 2002)

According to the brief by the Ministry the CATS model should be useable for.

- Identifying areas for improvement to the technical and managerial safeguards against accidents.
- Quantifying the risk implications of alternative technical and management changes, allowing evaluation of their cost-effectiveness.

This requires a model of the causes of accidents, based on a realistic description of the air transport industry and its safety functions, including the relationship between technical and management systems.

General approach

Aviation accidents tend to result from a combination of many different causal factors (human errors, technical failures, environmental and management influences) in certain characteristic accident categories (loss of control, collision, fire etc), whose causes and consequences differ according to the phase of flight in which they occur (taxi, take-off, en-route etc). The CATS project approaches this complexity by developing separate causal models for each accident category in each flight phase. These are represented as Event Sequence Diagrams (ESDs) and Fault Trees. The back-bone of the model consists of a chain of these ESDs. All these separate elements are then converted into a single Bayesian Belief Net (BBN). This allows the model to take into account dependencies and also to model the softer influences such as management in a homogeneous manner. These modelling components are described further in the report.

The resulting model is a true "causal risk model", since it covers consequences of accidents as well as their causes. As the consequences have been dealt with in "statistical-causal" models that are used extensively for quantification of third party risk (risks to people living around airports), there is no need for further analysis of these aspects in this project. The majority of the modelling effort focuses instead on the causes The consequence model conventionally forms the right-hand side of a bow-tie model. This ensures that the outputs of the causal models (i.e. the frequencies of the different accident categories and their causal breakdowns) combine in a consistent way to give the risk results, allowing valid comparison of options affecting different accident categories.

The safety management model represents the elements of the safety management systems of the different actors in the air transport system (aircraft operator, air traffic services, airports, maintenance etc), which may influence many of the elements of the causal and consequence models. In simple terms, safety management controls the safeguards (what are now commonly referred to as barriers) intended to prevent hazards leading to accidents. In CATS these are represented by management influences on the behaviour of people involved in the air transport system.

Mathematical tools

The development of the CATS model required some significant developments of the mathematical tools available. The final outcome of a calculation with CATS is the probability of an accident. Using Bayesian Belief Nets as the modelling vehicle for calculating accident probability proved to have several advantages over using Fault Trees and event trees. First of all in BBNs the events do not have to be linked deterministically as in the trees. In trees the state of an event can only be a binary quantity: yes or no, true or false. BBNs support both functional and probabilistic nodes. Roughly, this means that they can capture all functional relations and also dependences between probabilities of occurrence of base events.

The BBN structure also allows analysis of the correlation of accidents with the underlying causes. In a system which is highly reliable such as the air transport system there are not many accidents for which a single defined cause can be established. Correlation analysis may give a lead to combinations of more extreme values of parameters in the system that could cause an accident. A system was developed which displays the distributions of parameters associated with a certain selection of values of other parameters or variables.

Using a BBN the interdependencies between different sections of the model, such as the relationship between engine failure, fuel starvation and go-around manoeuvres can rigorously be modelled. Here the real power of using a BBN

over the event and Fault Trees manifests itself. The effects of interdependencies on the final result can be modelled directly.
No less useful is the fact that the states of the nodes can be distributed over many values and that this distribution can be continuous rather than discrete and that the edges of the BBN are – conditional – correlations.

Data

A model such as CATS has large data requirements, the major problem being the exposure data. It is not sufficient to know how many failures of a certain piece of equipment are recorded in an accident database. It is necessary to also know how many failures of that same instrument occurred without an accident and in how many flights the equipment did not fail at all.
Data are gathered from ICAOs ADREP database, from data made available by airlines and by airports. In addition work data is used from the Line Operation Safety Audit (LOSA) database to establish the performance of pilots with and without accidents. If the performance was – in part – influenced by the equipment or by circumstances these underlying causes were taken into the model whenever possible.
For the development of CATS in all a few thousand numbers needed to be extracted or estimated. The origin and a characterisation of the quality of the data are held in a separate database. This not only helps future users of CATS in interpreting the results of an analysis, but also forms a basis for recording data in the future. By targeted recording, weaknesses and holes in the data structure can gradually be remedied.
UNINET is the software to drive the BBN and is open source The software to build the model is developed by TU-Delft especially for the project. Full documentation can be found on *http://dutoisc.twi.tudelft.nl/~risk*
In many cases experts use aggregate notions such as the complexity of an airport, the complexity of airspace, good or adverse runway conditions and aircraft generation. These notions translate into changes in probabilities of many of the model constituents. Therefore a translation or mapping has to be made of the variables or notions common in the industry onto the base events of the BBN.
In CATS the estimates from experts and the estimates from data are brought together in one system. Calculations are performed to establish a consistent picture between all the "known" quantities in the BBN by adjusting the "unknown" quantities.

Uncertainties

In the course of the development and testing several occasions have been identified where the total of the information is inconsistent. At this stage of development this issue has to remain unresolved. The next stage CATS will be used to explore discrepancies between expectations, judgments and reported facts. Even when the model is kept relatively simple there are many layers in the model where safety management systems can be taken into account. Differences of a factor of 1.5 build up quickly to orders of magnitude. This may be seen as an argument against quantitative modelling as the accuracy of these models then cannot be better than orders of magnitude. It should be borne in mind though that the estimates of experts are equally loaded with uncertainties. The currently dominant way of making decisions on the cost effectiveness of investments in safety, safety measures and safety management is mainly based on expert opinion. The deception after some time that measures did not bring what was expected is the unavoidable result, if these opinions consistently overestimate effects of change.

Care

Validation of the CATS model has only been possible to the extent that past changes in safety performance resulting from design decisions are calculated correctly. The available data are barely enough to populate the model with the required initial set. Independent quantitative validation is impossible. Therefore other approaches will be used to maximise the validity of the model, such as comparison with other existing models, expert and peer review on the equations, probabilities and distributions used. Once this validation has been done, the model will be used first as an additional input to safety decisions in the Dutch air transport industries. It took about 20 years between the conception of a causal model for chemical plants (Ale and Whitehouse, 1984) and the introduction into the legal system in the Netherlands (NNm, 2004). A similar cautious introduction of these sorts of techniques in the air transport industry should be expected.

Care must always be taken in generating measurable performance shaping factors of human beings. Some influences are too complicated to represent at this stage or we are unable to quantify them in numerical units. Therefore the nodes have been limited in their definition and modelled in a way that can be quantified by the BBN. However, this does not necessarily tell the whole story because important influences may have been lost.

Care must also be taken in interpreting the results of the expert judgements. The judgement is crucially determined by the original list of possible influences and their phrasing. Ideally more experts should have been involved to make sure no relevant factors were left out. More work should be also done to fine-tune the method for the cut-offs of the distribution.

The work that started three years ago resulted in a single Bayesian Belief Net structure to describe the probability of an air transport accident. The first applications indicate that the model functions correctly and produces results that are in accordance with observations and expert insight.

However CATS or similar quantitative methods, which bring together reports, observations, facts, opinions, judgments and expectations, can help to improve our insight into what can make air transport safer. It also suggests a pathway to a further development of methodology in other strands of quantified risk analysis.

Expected and unexpected outcomes will need to be carefully evaluated in the next period to gain confidence in this new way of building a causal model.

By virtue of the use of a single BBN, interdependencies could be rigorously modelled and the human performance models could be integrated. For this reason alone the results and performance of the model already exceed the initial expectations.

Further work

Further work needs to be done on validation and on human response and management modelling.

The current workings of the model were carefully checked and rechecked to avoid errors. The inputs were reproduced and a few preliminary case studies showed good behaviour. Nevertheless full validation against an independent dataset was impossible and therefore trust in the model can only be gained by applying the model to a series of test and real cases if possible comparing the results with results from other modelling efforts elsewhere in the world, even if the latter are much less comprehensive.

There is an obvious need for further data. It would be of great help if company specific data on incidents could be used to get an even better estimate of the

probabilities of events earlier in the causal chain.
The human response modelling although much improved when compared to models in other fields still needs much improvement. The most important of these is to get a better understanding of the relationship between qualitative generally understood notions and the translation of these in real observable and thus quantifiable influence on risk and risk reduction.
Maintenance is an underdeveloped area in CATS. Although the maintenance technician is modelled, he has a much more indirect influence on the system than crew and ATCs, whose decisions and actions are directly in the causal chain. FAA have a model developed aimed at the probability that an airplane takes to the sky without being formally airworthy, which is no the same as being in danger of crashing. It would be advisable to investigate whether a part of the maintenance model developed for the FAA could be attached to CATS.

Finally

The current model is – as one member of the international expert committee put it – the second best representation of the reality, reality itself being the best. It provides a much safer testing ground for extreme and unexpected circumstances and new developments than reality. But is remains a model. Therefore caution with the results is always a good strategy.

1 Introduction

The Ministry of Transport and Water Management (V&W) commissioned a causal model of air transport safety. The model was developed by a consortium including Delft University of Technology (TUD), National Aerospace Laboratory (NLR), White Queen Safety Strategies (WQ) and the National Institute for Physical Safety (NIFV) in The Netherlands, and Det Norske Veritas (DNV) and JPSC consulting (JPSC) in the UK. The project is known as Causal model of Air Transport Safety (CATS).

The motivation for the project is the need for a thorough understanding of the causal factors underlying the risks of air transport so that efforts to improve safety can be made as effective as possible.

The project as a whole is defined in the program of work from the consortium manager, TUD Risk Centre (23 August 2004)[1]. The work program is appended to this report. Most of the work program has been carried out. The project kick-off date was 14 July 2005. The project ended on 31 October 2008. Full validation and further development of management influences remain a subject for further work as will be described later in this report.

Third party risks of air transport have been a subject of political debate in the Netherlands ever since the expansion of the airport was proposed in 1989. This has lead to a continuous effort in developing and improving the understanding of these accidents. Originally efforts were aimed at developing models to describe the probability and consequences of crashes based on statistical evaluation of similar accidents in the past. Such modelling is very limited in its ability to investigate and evaluate actions to reduce the probability of these accidents.

The desire of the Ministry to have such a model stems from extensive discussions in parliament about the further development of Schiphol Airport. The debate about the safety of these developments was intensified by the crash of a Boeing 747 into an apartment building in one of the densely populated suburbs of Amsterdam in 1992 Ale and Piers, 2000). This led parliament in 2001 to adopt an amendment to the Air Navigation Act (NN, 1992) in which it was demanded that a causal model would be developed and that it would be used to show that the increase in traffic would not lead to an increase in disaster potential.

The Ministry expressed the following potential uses of the model:

- To enable comparative judgements
  - With other means of transport
  - Over time
  - Between airports
  - Between links in the chain of safety
- To diagnose risk situations within the own organization and outside the boundaries thereof to enhance the safety of air transport pro-actively
- To prioritise the potential safety measures (e.g. on the basis of the expected effectiveness).
- To support supervision of the quality of safety delivered by the air transport system and its parts.
- To inform the public on risk policy and risk reduction efforts.

The ministry wanted an interactive and collective approach leading to an authoritative model, which would allow an integral analysis of safety problems in air

[1] Appendix: work program

transport. This model would provide increase knowledge and add an integral tool of management to the current instruments of the ministry. In 2003 the strict demand of proof of a not increased disaster potential was relaxed and is no longer a demand of law. This took the development of the causal model out of the political arena and the sometimes heated debate between the different parties involved.

The objective of the CATS project thus is to develop a fully operational causal model, that represents the causes of commercial air transport accidents, and the safeguards that are in place to prevent them, building on the experience gained in the demonstration causal models developed by DNV and a consortium led by NLR during 2001-2002.

According to the brief by the ministry the resulting model should be useable for.

- Identifying areas for improvement to the technical and managerial safeguards against accidents.
- Quantifying the risk implications of alternative technical and management changes, allowing evaluation of their cost-effectiveness.

This requires a model of the causes of accidents, based on a realistic description of the air transport industry and its safety functions, including the relationship between technical and management systems.

Aviation accidents tend to result from a combination of many different causal factors (human errors, technical failures, environmental and management influences) in certain characteristic accident categories (loss of control, collision, fire etc), whose causes and consequences differ according to the phase of flight in which they occur (taxi, take-off, en-route etc). The CATS project approaches this complexity by developing separate causal models for each accident category in each flight phase. These in turn are represented as Event Sequence Diagrams (ESDs) and Fault trees. The back-bone of the model consists of a chain of these ESDs. All these separate elements are then converted into a single Bayesian Belief Net (BBN). This allows the model to take into account dependencies and also to model the softer influences such as management in a homogeneous manner. These modelling components are described further in the report.

The resulting model is a true "causal risk model", since it covers consequences of accidents as well as their causes. As the consequences have been dealt with in "statistical-causal" models that are used extensively for quantification of third party risk (risks to people living around airports), there is no need for further analysis of these aspects in this project. The majority of the modelling effort in this project instead focuses on the causes and on the coupling of the consequences to the decision-making.

As mentioned above, separate models were initially developed for each accident category. The main reason for this was to allow a staged development, integrating work by different organisations. Common elements in the causal models of different accident categories can be modelled separately by different organisations. An example is flight crew fatigue which may influence the causes of all accident categories.

The consequence model conventionally forms the right-hand side of a bow-tie model. This ensures that the outputs of the causal models (i.e. the frequencies of the different accident categories and their causal breakdowns) combine in a consistent way to give the risk results, allowing valid comparison of options affecting different accident categories. The input requirements for the consequence models form the output specifications for the causal models. If followed, this will ensure that they are integrated as required.

The safety management model represents the elements of the safety management systems of the different actors in the air transport system (aircraft

operator, air traffic services, airports, maintenance etc), which may influence many of the elements of the causal and consequence models. In simple terms, safety management controls the safeguards (what are now commonly referred to as barriers) intended to prevent hazards leading to accidents. In CATS these are represented by management influences on the behaviour of people involved in the air transport system.

The development of the CATS model required some significant developments of the mathematical tools available. A true report on such a development therefore cannot go without some description of the mathematics involved. This is also necessary because the readers of this report will not only be civil servants in the Ministry; the report will also be the starting point for scientists and technicians who want to further build on this project. Nevertheless the mathematics in this report has been kept as simple as possible, with references to more complete descriptions elsewhere. The remainder can be read without these mathematical parts. The descriptions and examples used to clarify the approach are kept as easily understandable as was possible. This sometimes implies that the full depth of the scientific background and more intricate technical details had to be omitted. Readers with scientific interests are referred to the appendices and the literature cited for further background.

The remaining chapters of the report are as follows:

In chapter 2 the model is described. It is set out how it was conceived and constructed, how the data were analysed and used to support the quantification. The usability is discussed and the limitation of the current development.

In chapter 3 the underlying concepts of risk and safety modelling and the mathematics involved in building the model are expanded upon. Examples of human response modelling are used to illustrate these mathematics and to describe the modelling of human error in more detail.

In chapter 4 an overview is given of the quantification methods used in building CATS.

In chapter 5 the principles of human error modelling are further explained and in chapter 6 it is described how management influences have been incorporated into CATS.

In chapter 7 the modelling and quantification of the consequences of accidents are explained.

In chapter 8 a discussion is given on the validation work that was undertaken

In chapter 9 some conclusions are drawn and recommendations are given for further work on CATS and on air transport safety.

2 CATS

In this chapter the final product CATS is described. The description is based on the normal use of the model through the provided interfaces. The support programs that come with CATS are also described. At the end of this chapter the potential uses of the current CATS model, its limitations and potential future developments are explained. In later chapters in the report further descriptions are given of the underlying methodology and the mathematics involved. From these descriptions the scope of the use of CATS, the limitations and the need for further development can be more fully understood.
The use of the program can be understood from this chapter alone. Some concepts used in CATS are explained in this chapter in general terms. This explanation is necessary in order to understand the data structure in CATS and how it can be manipulated. In particular the role of Event Sequence Diagrams, Fault trees and Bayesian Belief Nets are briefly described. After this description in general terms the particular application to CATS and the air transport safety problem are discussed from section 2.8.

2.1 The purpose of CATS

CATS is a causal model for air transport safety. Its purpose is to establish in quantitative terms the risks of air transport.
The current development of CATS makes it suitable for supporting strategic and tactical decisions with risk information. It can be used to evaluate the effect on risk of proposed measures and it can support the analysis of what measures could be the optimal way of reducing risk in general or certain types of risk in particular.
In principle risk calculations could be done for a particular airport and the software is made to support such analyses. However it should be noted that many of the parameters which influence the risk picture and for which general values are known, are less known for a particular airport and care should be taken each time the calculations are made for a subset of the population of airfields and aircraft. In a subset all input should be conditionalised tailored to the subset chosen, and the correctness of these settings determines the correctness of the final result.
Not all possible risks of air transport are modelled by CATS. CATS is primarily aimed at calculating the probability and the damage of accidents.
ICAO Annex 13 provides definitions of accidents and incidents that may be summarised as follows (ICAO, 2005):
An accident is an occurrence during the operation of an aircraft that entails:

1. A fatality or serious injury;
2. Substantial damage to the aircraft involving structural failure or requiring major repair of the aircraft; or
3. The aircraft is missing.

By calculating the risk of accidents for different circumstances it can be investigated whether certain conditions make risks worse or better and how safety could be improved by taking certain measures.
Risk analysis used to be relatively simple, when the causes of accidents could be readily identified. However, the current level of safety in aviation is such that causes of future accidents are much more difficult to find, even though after the fact, i.e. after an accident has occurred, the cause can always be found.
In many cases a cause is not a singular event. Often it is a combination of many

factors. These factors may be the combination of extreme values of parameters within the defined or designed range of allowable values.
CATS is especially designed to make it possible to find and study these combinations and the interactions between the many systems and parameters in an aircraft, the people who operate and maintain the aircraft and the system, and the people who make the airplane fly, such as pilots.

2.2 The CATS system

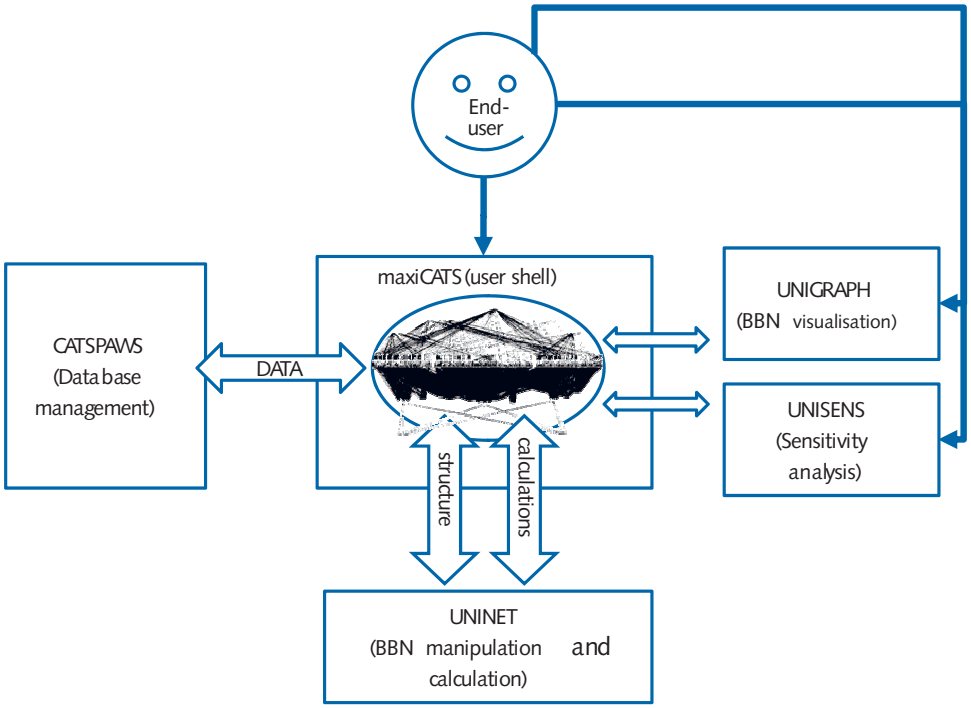


Figure 1: The CATS system.

CATS consists of a number of programs to allow the analyst to perform these analyses. These programs are: maxiCATS, CATSPAWS, UNINET, UNISENS and UNIGRAPH. These programs are used in concert to manipulate the integrated data structure that forms the heart of the CATS system. This data structure is referred to as "The BBN" throughout this report. BBN is the acronym for Bayesian Belief Network. In a BBN events, also referred to as NODES, are connected by causal relationships, also referred to as ARCs. In later chapters of this report this is explained in full. For this section it is sufficient to just know that the BBN represents the events in air transport that are relevant for safety and the causation of accidents and the relations and interactions between them.
The CATS system is depicted in Figure 1. CATS is the user shell around the BBN. It allows the user to manipulate the input of the calculations and view the output. CATSPAWS (CATS PArameters With Sources) is used to maintain the database which holds the names and units of the BBN nodes, the underlying data and also information of where and how these data were obtained. However, all the BBN parameter values in maxiCATS are read directly from the BBN. UNINET performs all the mathematical operations in the BNN. It also is

used to modify and expand the structure of the BBN when needed. UNISENS is used to perform statistical analyses on the results of a calculation such as determining importance measures and dependent correlations. UNIGRAPH is used to display the BBN and also graphs of results. Later in this report these programs will be described in more detail.

2.3 The design of CATS

The design of CATS builds on work done in preparatory projects on air transport risk estimation (DNV 2002, Roelen et al 2000) and work done in the area of occupational safety, linking technological risks to management influences (Ale et al 1998, Bellamy et al, 1999). The design described in these reports called for the combination of three modelling techniques in a single model: Event Sequence Diagrams (ESD), Fault Trees (FT) and Bayesian Belief Nets (BBN).

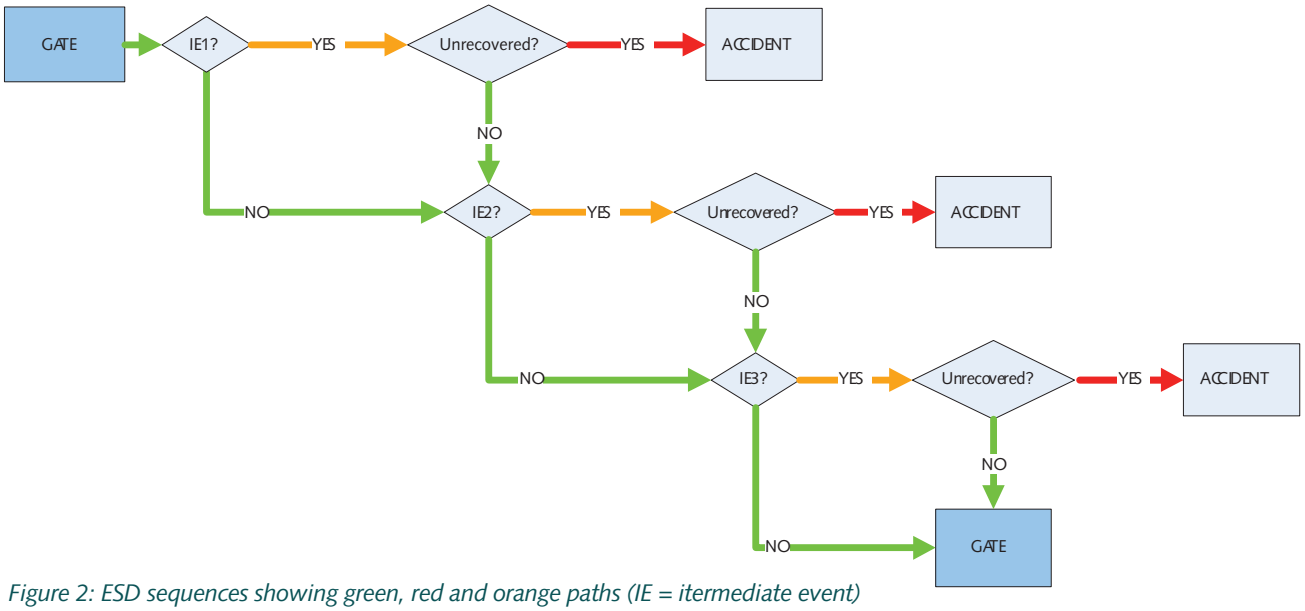


Figure 2: ESD sequences showing green, red and orange paths (IE = itermediate event)

2.4 Event Sequences

The potential accidents that could take place in a journey are divided into accident categories, which connect similar types of accidents with similar groups of causal factors for analysis in one part of the model. The choice of these categories is more a matter of convenience than of principle. Nevertheless they should cover the whole range of potential accidents. Existing taxonomies, such as the one used in the International Civil Aviation Organisation (ICAO) Accident/Incident Data Report (ADREP) database are readily available. While this has several advantages (amongst others the fact that the existing ADREP database stored in the European Aviation Safety Agency software ECCAIRS can be used in a straightforward manner for quantification) there are disadvantages as well. The main disadvantage is that the taxonomy contains many ambiguous elements. While this is less of a problem for the original use of the taxonomy (coding of accidents and incidents), for the causal model it is an undesirable characteristic.

Event Sequences are depicted in Event Sequence Diagrams (ESD). In terms of the logic of CATS the ESDs can be seen as representing the dangers or hazards that each flight has to overcome in order to safely complete the journey (Figure 2). Whether a particular flight encounters one of the hazards depends on whether the initiator occurs. Whether the flight survives depends on how the crew/equipment system copes with the hazard.

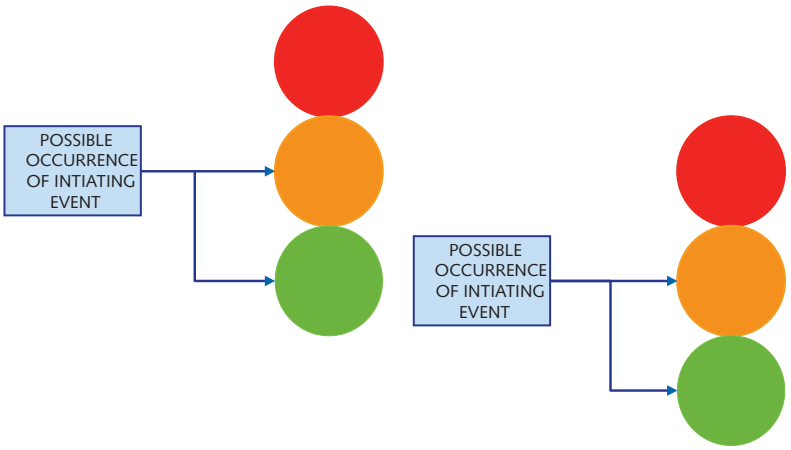


Figure 3: The traffic light model

When an initiator occurs there is an abnormal state (orange). If there is a failure to recover there is an accident (red). If there is recovery then the state becomes normal (green). In this way a traffic light model results(Figure 3). It is possible – and in practice quite likely – to go gate to gate without encountering any initiators.

It was also found convenient to divide these ESDs according to flight phase: Taxi (TA), Take-off (TO), Climb (CL), en route (ER) and Approach and Landing (AL). The original 37 ESDs of the NLR work (Roelen and Wever, 2005) resulted in 33 ESDs incorporated in CATS (Table 1)

Each event is defined in an ESD such that it can go in only two directions. The probability of going in either direction is determined by the outcome of a fault-tree.

In a project for the FAA NLR quantified the ESDs directly from data (Roelen et al, 2006). These quantifications formed the starting point for further analysis of data and further quantification by DNV (Spouge, 2008). The latter was based on quantification of the underlying Fault Trees.

2.5 Fault trees

For each of the pivotal events in the ESDs Fault Trees were developed. Although these could have been written in a compact way, in CATS they are depicted somewhat more elaborately. In this way each so-called "end gate" consists of a challenge and of a failure of a barrier to prevent the challenge to propagate as described earlier in this report. This gives the Fault Trees the ladder type appearance as depicted as the example given in Figure 4. This way of constructing the Fault Trees supports dialogue with technical experts from the industry.

At this stage of building the model, states could only be failed or not failed. However later, when the Fault Trees were converted to elements of the BBN, multiple (degraded) states were allowed and the Boolean logic was replaced by the probabilistic relationships which are used in the BBN.

The Fault trees were constructed from the analysis of the accident descriptions which are associated with the accidents that were the basis for quantification. This analysis was performed by dissecting these accident histories one by one to find potential causes of events already in the Fault Tree or new events in the causal chain towards a pivotal event in the ESD until no new events could be identified for several accident histories or the event frequency could be established from data – which means the failure of an identifiable technical system – or the event was a human action.

In developing the Fault Tree, a top-down approach is followed, which reverses these calculations. The top events of the Fault Trees are known from the initiating and pivotal events from the ESD. These are split into events corresponding to unsuccessful performance of each barrier. At each AND gate, additional probability data or assumptions are needed to quantify the input events. These unsuccessful barrier events are then further split into the causes of barrier failure. At each OR gate, causal distributions are needed as described below.

Development of the Fault Tree Model has followed the same approach as used by the EUROCONTROL Integrated Risk Picture (IRP) (Eurocontrol, 2006). By agreement with EUROCONTROL, the IRP models for collisions have been adopted from this source. Some changes have been necessary because the explicit modelling of common-cause events in the IRP is not required in CATS, since this aspect is represented by BBNs.

Quantification of the Fault Tree Model uses distributions of causes obtained from accident and incident experience. The quality of information in ADREP about accident causes is not sufficient to support the present analysis. Therefore original accident investigation reports have been used where available. In other cases, the summary information from Airclaims, Aviation Safety Network, Flight Safety Foundation and others explains the causes in sufficient detail to relate to the barrier model. Incident reports have also been used where available. The term "event" is used below to refer to both accident and incidents.

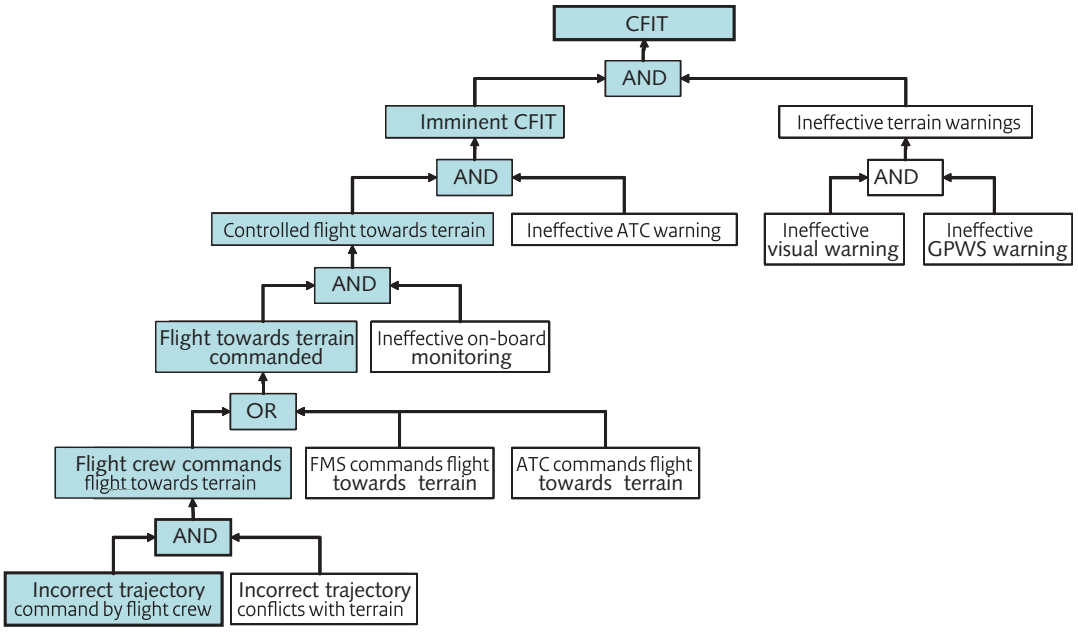


Figure 4: Fault tree for CFIT

To quantify the Fault Tree, it is not necessary to know the causes of every event that has occurred. Since the ESDs have been quantified using probability data, consisting of comprehensive counts of the numbers of events among known flight exposure, the causal breakdown in the Fault Trees can be quantified from a representative sample of events. It is therefore assumed that the events whose causes are known, and which are used to quantify the causal breakdowns in the Fault Trees, are representative of the causes of the full set of accidents.
For ESDs with little or no accident experience, the Fault Trees are quantified using experience from precursor incidents. These are incidents that were prevented from developing into the relevant accident by the success of one or more barriers. It is assumed that the causes of these incidents indicate the likely causes of initiating events in future accidents. The causes of the necessary further barrier failures can be obtained from other ESDs in which the same barriers are relevant, or as a last resort from expert judgement about their relative likelihood. In general, the Fault Trees have been developed only to a level that can be quantified mainly from available accident or incident data, and pure judgements about event probabilities have been minimised.

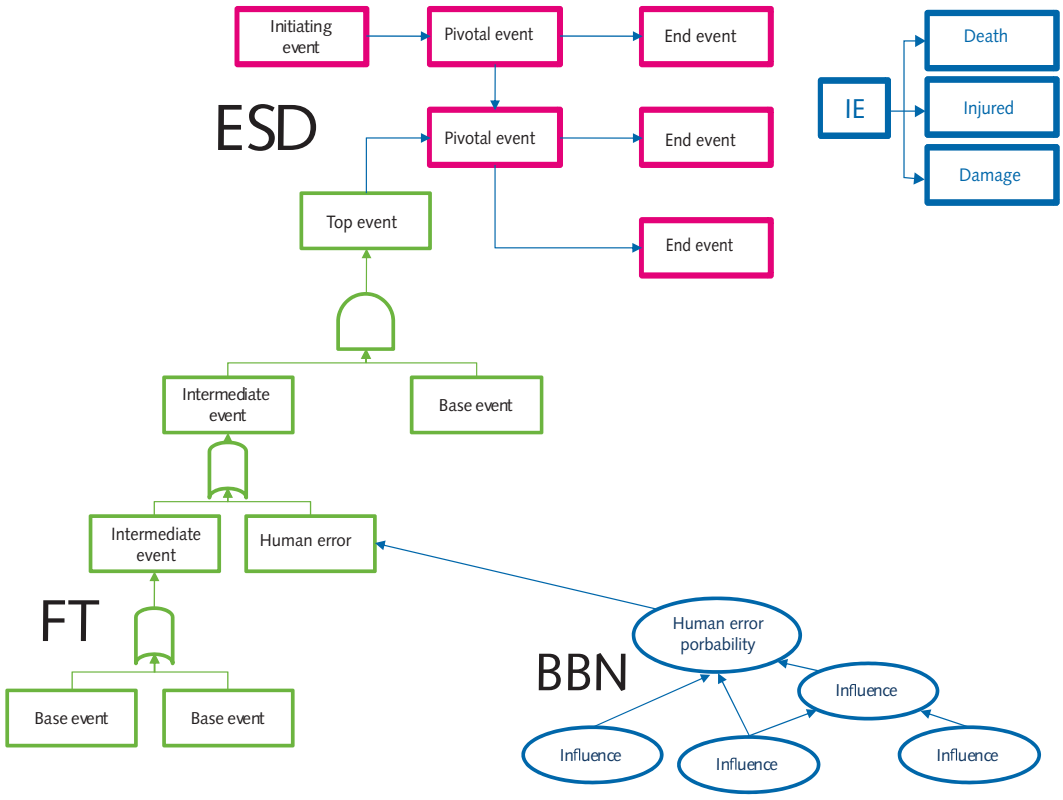


Figure 5: The basic constituents of CATS.

2.6 Human Error Probability

In over 100 places in the final model human intervention is needed to avoid an accident. The probability that these actions do not result in the desired effect are described in human performance models (HPM) , which are models of influence on human error probability (HEP). Of these there are three: the crew (Roelen et al 2007), the ATC controller (Roelen et al 2008a) and the maintenance technician (Roelen et al, 2008). These models hav been built on the basis using

the guidance on human performance modelling by the US Nuclear Regulatory Commission (NUREG 2005). The performance shaping factors were analysed for applicability for the performance of operators in the air transport industry. In order to make these performance shaping factors amenable for quantification, these performance shaping factors were translated into proxy quantities. The rationale behind the choices made can be found in annex NLR12.
In chapter 6 the quantification of human error will be described in more detail.

2.7 A single BBN

One of the main developments in making CATS a much more advanced tool than earlier systems for Quantified Risk Analysis is that the ESDs and the FTs are converted into BBNs and from that the CATS model is constructed as one integrated BBN. This allows the use of distributions of values rather than point estimates wherever appropriate. It also allows a convenient and consistent handling of dependencies and interdependencies throughout the model. It finally takes away the need for artificial transfer points in the model between ESDs, FTs and BBNs.
This however did not take away the need to first develop the ESDs, FTs and BBNs separately as these and their quantification form the basic material on which the integrated CATS BBN is built. The Causal model for Air Transport Safety (CATS) therefore integrates models for technical failures such as event sequence diagrams, Fault Trees, models for human behaviour in a single BBN. Many of the model elements are repeated. For instance, although the pilots remain the same during the flight, they may be tired at the end of the journey. The weather could be different for the two ends of the flight. Separate instances of the pilot model, of the weather influence and parameters associated with

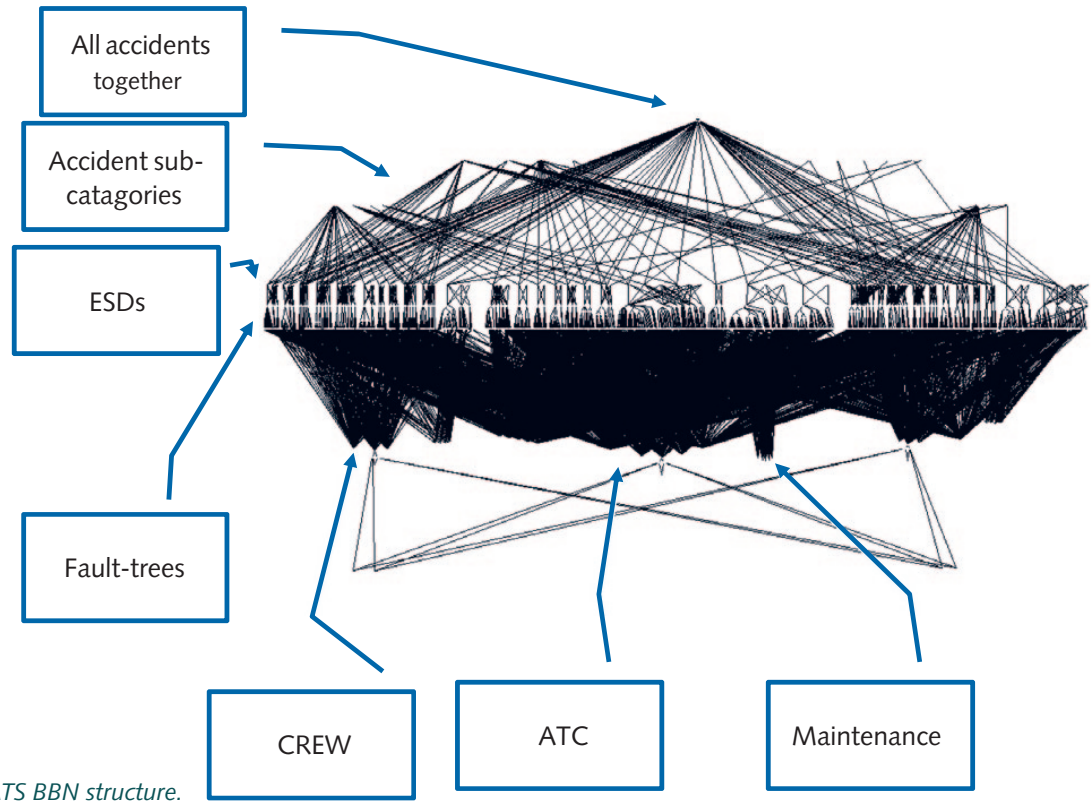


Figure 6: The CATS BBN structure.

airports are used when required.
In Figure 5 the principle of the building of the CATS BBN is depicted. The final outcome is the probability of an accident. In this BBN the interdependencies between different sections of the model, such as the relationship between engine failure, fuel starvation and go-around manoeuvres are already introduced. Here the real power of using a BBN over the ESDs and Fault Trees starts to manifest itself. The effects of interdependencies on the final result can be modelled directly.
No less useful is the fact that the states of the nodes can be distributed over many values and that this distribution can be continuous rather than discrete and that the edges of the BBN are – conditional – correlations.
The final BBN consists of approximately 1400 nodes and 5000 arcs. An impression of the structure is given in Figure 6. For an exploration of the model the CATS software needs to be used, as described in later in this chapter.

Table 1: Classes of accidents in CATS

ESD	Initiating event	Flight phases	Flight phases in CATS
1	Aircraft system failure	TO	TO
2	ATC event	TO	TO
3	Aircraft handling by flight crew inappropriate	TO	TO
4	Aircraft directional control related systems failure	TO	TO
5	Operation of aircraft systems by flight crew inappropriate	TO	TO
6	Aircraft takes off with contaminated wing	TO	TO
7	Aircraft weight and balance outside limits	TO	TO
8	Aircraft encounters windshear after rotation	TO	TO
9	Single engine failure	TO	TO
10	Pitch control problem	TO	TO
11	Fire on board aircraft	CL, ER, AL	ER
12	Flight crew member spatially disorientated	CL, ER, AL	ER
13	Flight control system failure	CL, ER, AL	ER
14	Flight crew incapacitation	TO, CL, ER, AL	ER
15	Anti-ice system not operating	CL, ER, AL	ER
16	Flight instrument failure	CL, ER, AL	ER
17	Aircraft encounters adverse weather	CL, ER, AL	ER
18	Single engine failure	CL, ER, AL	ER
19	Unstable approach	AL	AL
20	Deleted (incorporated in ESD 19)	-	-
21	Aircraft weight and balance outside limits	AL	AL
22	Deleted	-	-
23	Aircraft encounters windshear during approach/landing	AL	AL
24	Deleted (incorporated in ESD 19)	-	-
25	Aircraft handling by flight crew during flare inappropriate	AL	AL
26	Aircraft handling by flight crew during roll inappropriate	AL	AL
27	Aircraft direction control related systems failure	AL	AL
28	Single engine failure	AL	AL
29	Thrust reverser failure	AL	AL
30	Aircraft encounters unexpected wind	AL	AL
31	Aircraft are positioned on collision course	CL, ER, AL	ER
32	Incorrect presence of aircraft/vehicle on runway in use	TA, TO, AL	TO, AL
33	Cracks in aircraft pressure cabin	CL, ER, AL	ER
34	Deleted (incorporated in ESD 17)	-	-
35	Flight crew decision error/operation of equipment error	CL, ER, AL	AL
36	Ground collision imminent	TA	TO, AL
37	Wake vortex encounter	CL, ER, AL	ER

2.8 Accident scenarios

In the previous sections the technical concepts used in CATS were discussed in general terms. From this section onwards the application to the air safety problem is described. In CATS a number of accident scenarios are distinguished that can take place in one of three phases in a journey: take-off (TO), which includes taxi and climb, en route (ER), and approach and landing (AL), which includes the final taxiing.
There are 33 different accident scenarios as given in Table 1. Note that the numbering in this table runs from 1 to 37, but that there are 4 numbers designated as deleted. These classes of accidents were deleted in the course of the development of CATS. In order to maintain the correspondence of the numbering in CATS with the numbering in the technical reports that were produced during development, the original numbering is maintained. The actual number of classes therefore is 33.
In CATS there are some 1200 initial and intermediate events associated with these accidents. The probability of initial events can be changed directly by the analyst or indirectly through a mechanism called mapping, which will be explained later. The probability of all the intermediate events can be seen by the analyst in the output, as also will be explained later in this chapter.

2.9 ESD

These classes of accidents can also be seen as challenges that may present themselves in the course of a journey gate to gate. These challenges have to be overcome to complete the journey successfully. Each challenge is initiated by an initiating event. Following this initiating event actions and reactions of systems in the aircraft or of the pilots may lead to success, to partial success or to failure. Success means that the journey continues normally and no residual effects remain.

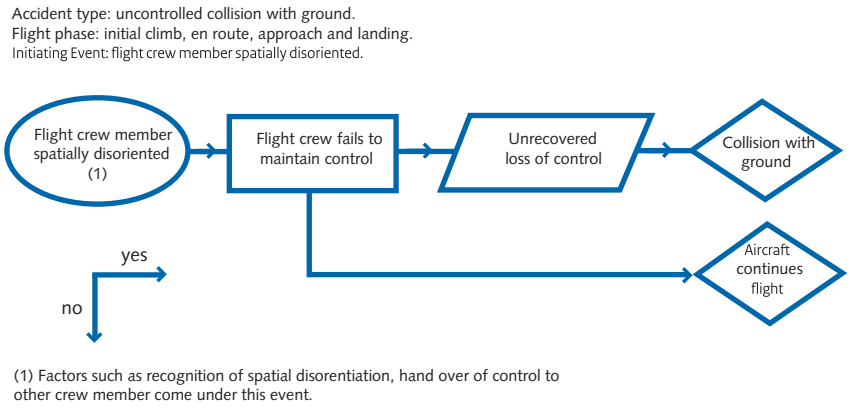


Figure 7: Schematic of an ESD (nr 12 used as an example)

Partial success means that the journey continues normally but that some effects of this challenge remain that may have consequences later in the journey. Such an effect could be additional use of fuel, which may lead to fuel starvation of the engines later.

Failure means that the journey ends in an accident. As described earlier, this usually means the premature end of the journey.
In the later more technical chapters of this report and in the technical reports in the annexes, the ESDs are depicted in three ways.
The first is the schematic, showing the initial event, the so called pivotal events and the possible outcomes. The pivotal events are those events in the model where the chain of events – also called the scenario – may turn good or bad, depending on an action. In Figure 7 the schematic is given for ESD 12, uncontrolled collision with the ground. In this example there are only two outcomes: failure (or BAD), which means collision with the ground, and success (or GOOD), which means continuation of the flight.
The second is the quantification. Here the probabilities and frequencies derived from data are shown in conjunction with a diagram of the ESD. In diagrams such as Figure 8, the user and the interested reader not only can find the numbers, but also the referencesto reports containing the source data.
The third way of depicting an ESD is given in Figure 9. This is a so called Storybuilder diagram (Bellamy et al, 2008).

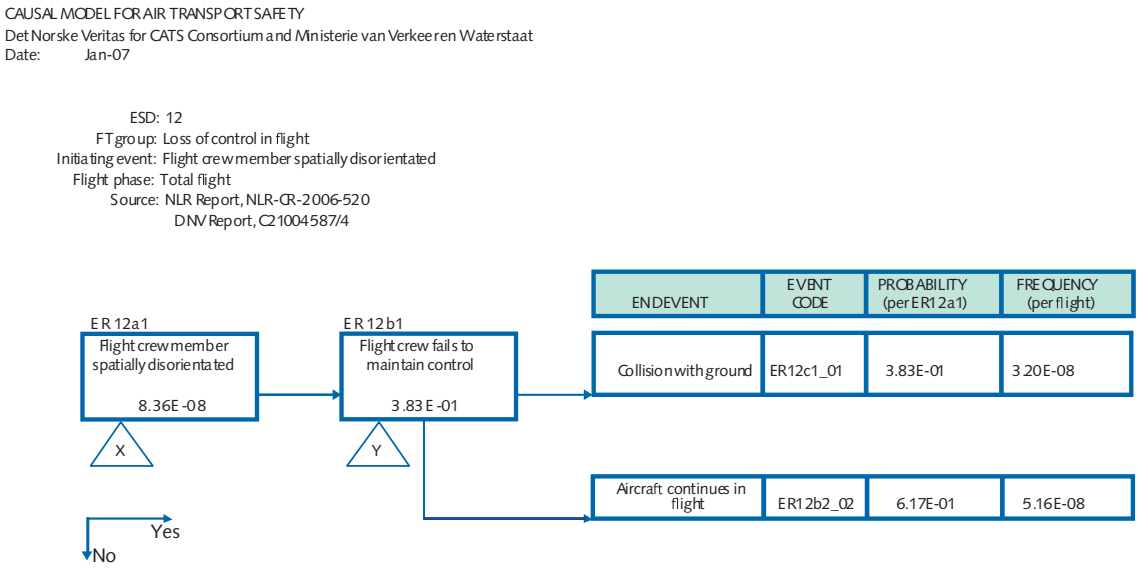


Figure 8: Quantification diagram of ESD12

These diagrams are used to support the analysis of accident reports and derive relative frequencies of chains of events in the historical evidence. In this representation the ESDs can be strung together to depict a continuous story of what happened during a certain journey and, in the case of an accident, what happened in the causal chain leading up to an accident. The complete structure of all the ESDs strung together is called the BACKBONE of the model. Storybuilder software was used to create this backbone as described in (Bellamy et Roelen, 2006). Storybuilder is an instrument for data analysis developed in the framework of another project Bellamy, 2006). The backbone structure is used in CATS and in the BBN as the reference structure of the model. All nodes and gates are identified with codings that refer to this backbone. This guarantees that all nodes are unique and the structure is consistent.
In the backbone there are more than 33 ESDs, as some ESDs can occur in more than one phase of the flight. The probabilities in these phases may differ. Thus several instances of the same ESD structure would be needed to implement these differences.

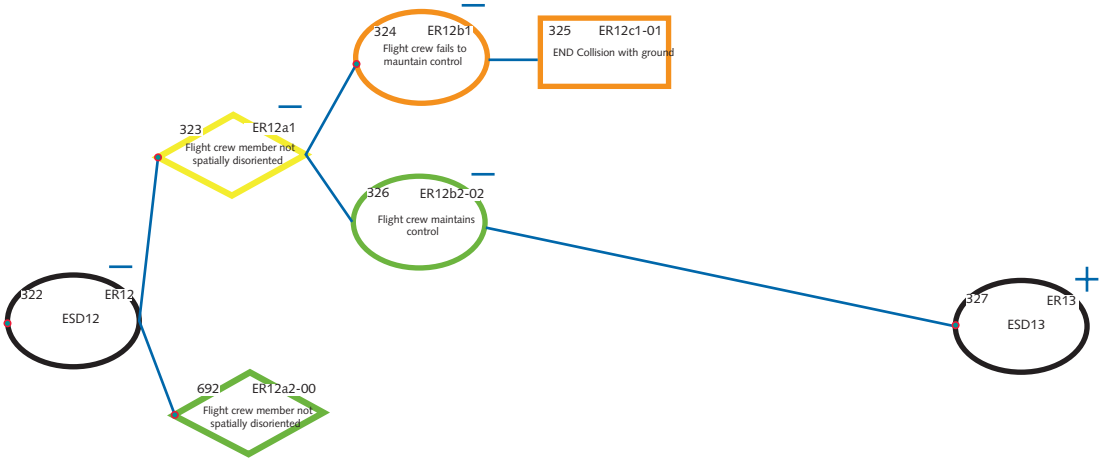


Figure 9: Storybuilder diagram of ESD12

For the current version of CATS insufficient data were available to actually implement multiple instances of the same ESD. Wherever the ESD could figure in different flight phases itv has been assigned to ER. In a future further development of CATS multiple instances of ESDs can be added when sufficient data become available.

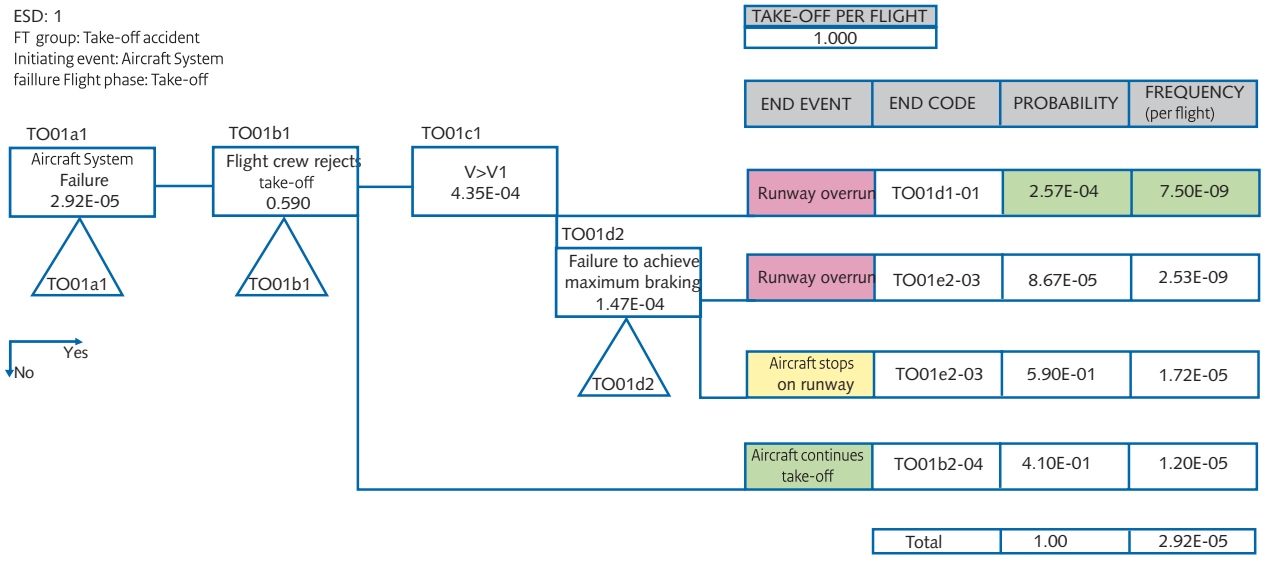


Figure 10: ESD1, showing GOOD, SORT OF GOOD and BAD end states and quantification

A more complicated ESD is ESD1, which is initiated by a system failure on the aircraft. Figure 10 depicts the ESD when this happens on the runway during take-off. Now also a partial success end state is possible (SORT OF GOOD), in this case when the aircraft stops on the runway. In principle it could now turn back and start again – preferably after repair of the failure – and continue its journey. In Figure 10 under the column END EVENT, the events are coloured green, yellow and red corresponding to the GOOD, SORT OF GOOD and BAD end states. This figure also illustrates the pedigree system used in CATS.

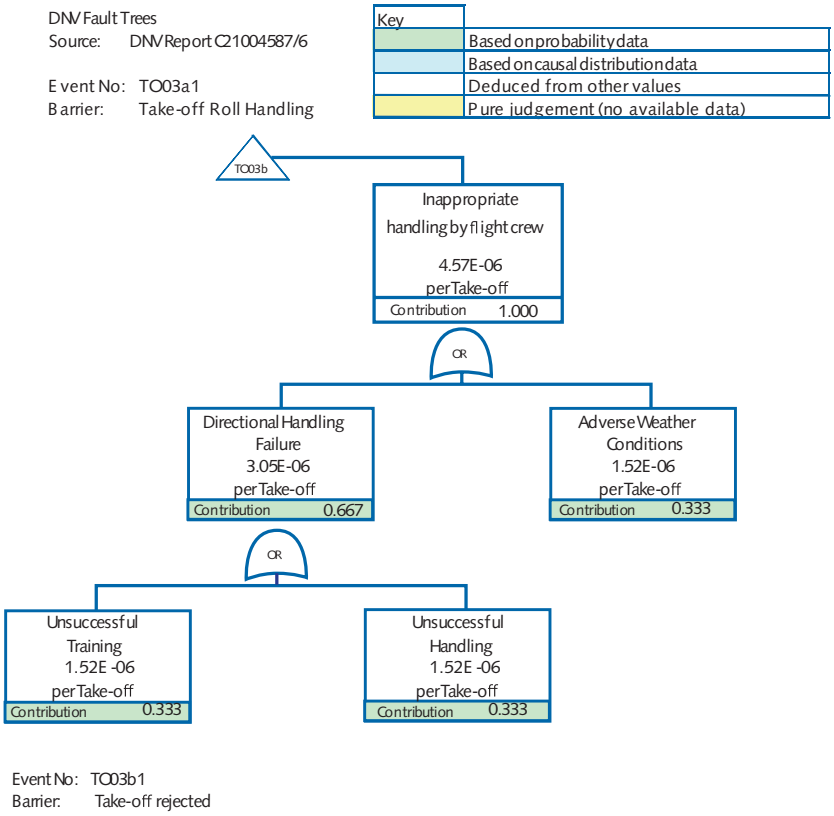


Figure 11: Example event tree from the Causal Model for Air Transport Safety

The colouring in the probability and frequency column refer to the pedigree of the data. Data are given a pedigree which for the convenience of the user is reflected in the colour used to display these data. These are

- Directly from event and exposure data (Green)
- Based on the distribution of causes in the event data (Blue)
- Deduced from other parameters (White)
- Based on expert judgement (Yellow).

2.10 Backbone

As mentioned earlier, all the ESDs were strung together to form the backbone of the model which is given in detail in (Bellamy and Roelen, 2006).
In Figure 12 a reduced size picture of this model is given.

Figure 12: the backbone of CATS

2.11 Fault trees

It can be seen in the examples of the ESDs that at each pivotal event the course of events can go in two ways. The probability of going in either direction is determined by a Fault tree (FT). Fault trees are designed to capture the causal chains that may lead to a certain outcome and to depict and describe the logical relationships between the events leading to this outcome.

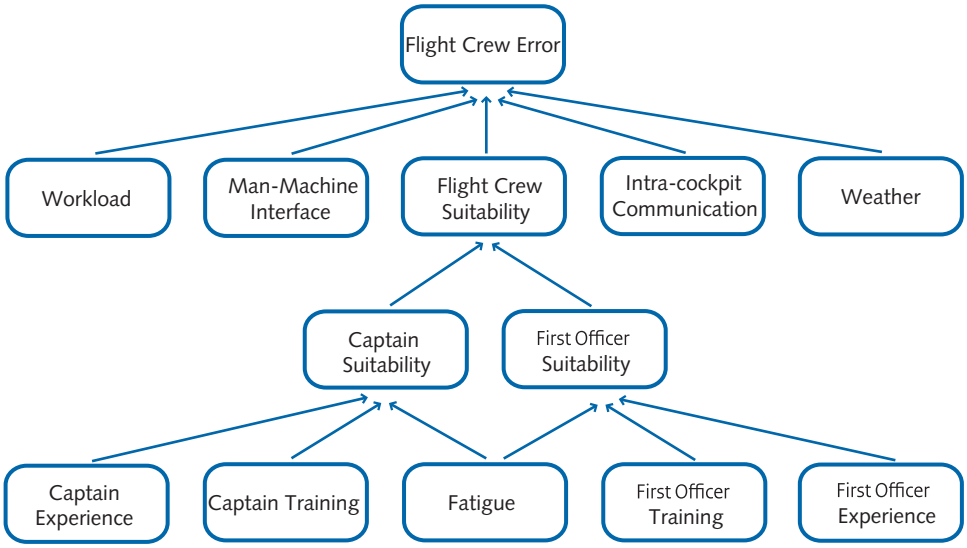


Figure 13: Flight crew performance model

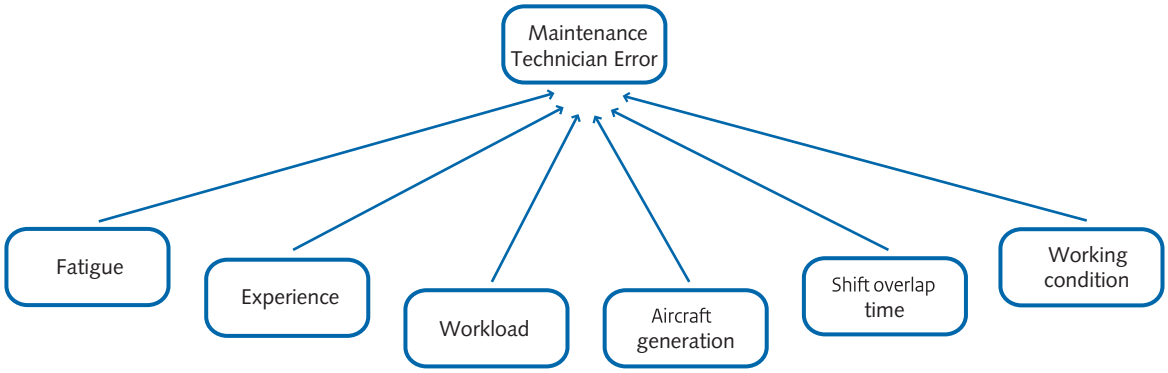


Figure 14: Maintenance Technician Performance Model

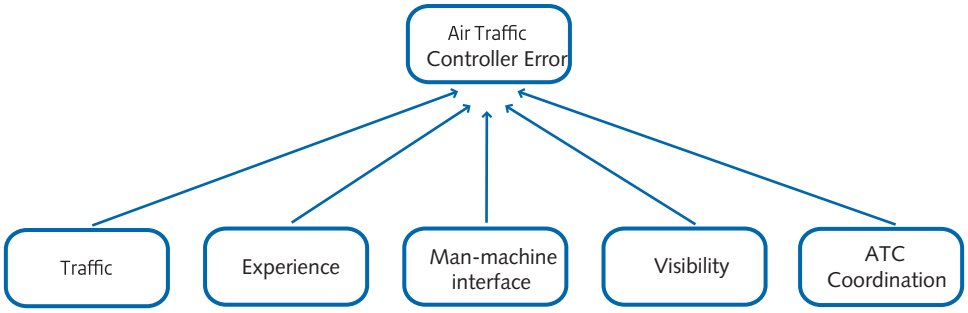


Figure 15: Air Traffic Controller Performance Model

This outcome is also called the TOP EVENT of the tree. The probability of the top event is logically and quantitatively dependent on the probabilities of base events. The top event and the base events are connected through intermediate events and gates. Gates determine how the events feeding into the gates combine. The probability of the top event can be calculated when the probabilities of all the base events are known. In Figure 11 an example of a Fault Tree in CATS is given. In this figure the use of the pedigree of the numbers involved can also be seen.

2.12 Human action

In over 100 places in the final model human intervention is needed to avoid an accident. The probability that these actions do not result in the desired effect are described in human performance models (HPM) , which are models of influence on human error probability (HEP). Of these there are three: the crew (Roelen et al 2007) (Figure 13), the ATC controller (Roelen et al 2008a) (Figure 15) and the maintenance technician (Roelen et al, 2008) (Figure 14). Since they are models of influences rather than dirct causes they are directly modelled as BBNs. These models have been built on the basis of the guidance on human performance modelling by the US Nuclear Regulatory Commission (NUREG 2005). The performance shaping factors identified in general in these guidelines were analysed for applicability for the performance of operators in the air transport industry. In order to make these performance shaping factors amenable for quantification, these performance shaping factors were translated into proxy quantities. The rationale behind the choices made can be found in annex NLR12.

2.13 Data Flow

It may be obvious that an elaborate model such as CATS needs a vast amount of data. The flow of data is depicted in Figure 16. The various elements depicted in this figure will be explained in the sections which follow.

2.14 HELP

The model and the computer program that embodies it have been provided with help facilities, which the analyst can use to guide him through the process of performing an analysis. These help facilities suppose that the analyst has a minimal understanding of the principles of performing a quantitative risk analysis (QRA). Some of the principles of QRA are explained later in this report, but only in as far it is necessary to underpin the choices and modelling techniques in CATS. The user of CATS is supposed to be able to perform a risk analysis.
In the next sections the general workings of CATS are explained. For details the user is referred to running the program and looking at the software help files and supporting technical report documentation where the program is not self explanatory.

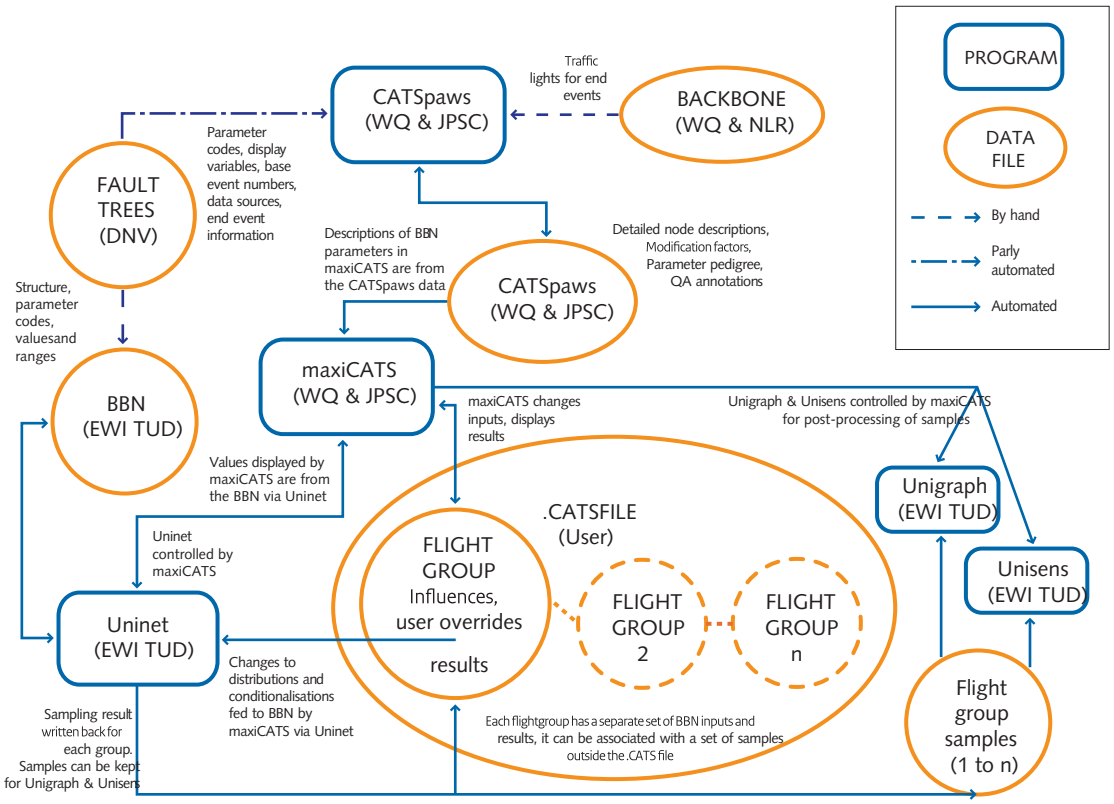


Figure 16: Data flow in CATS

2.15 Setting up a case

The base setting of CATS, also called the baseline case, corresponds to average values for Europa and North-America and distributions for all of the parameters in CATS and the result is the corresponding average probability of having an accident.

The user can set up their own case or cases by setting the values of parameters to be different from the world average. The user can set up different cases or FLIGHT GROUPS. By specifying the number of flights in each group an operation consisting of a number of groups of flights, each with different

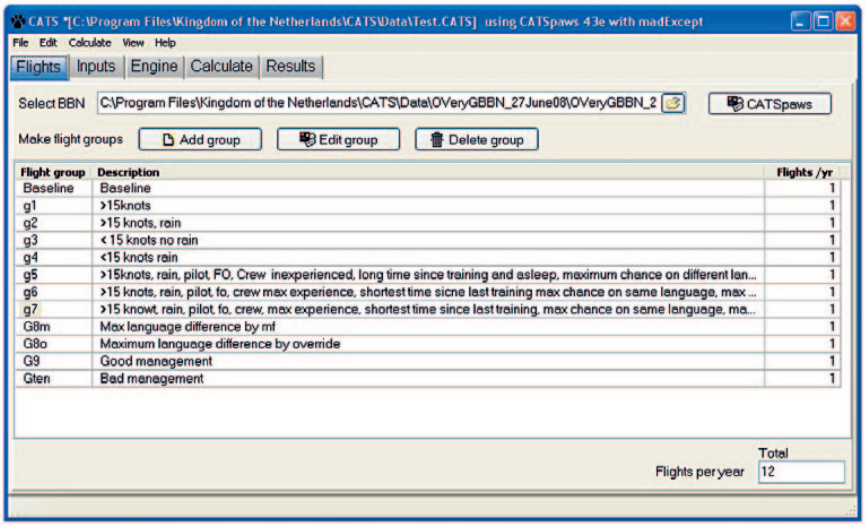


Figure 17: Setting up flightgroups

properties, can be built. An example could be the operations of a complete airline, consisting of specific number of flights with various generations of aircraft and under various weather conditions. The user could also leave the number of flights in each group equal to 1 (the default). Such a setting is useful for comparing the risk of various circumstances. Such a use is illustrated in Figure 17. Several specific weather cases are specified, a number of cases with various – extreme – properties of the crew and a number of cases that are associated with the management of the company.
The setting of the parameters can be done by "mappings" or by direct setting of the probabilities of the base events.

2.15.1 Mappings

The user can select a number of preset conditions, such as good or bad weather. If such a setting is chosen, the CATS system automatically sets all parameters associated with this condition to the corresponding values. In the case of bad weather, this would correspond to high wind and rain. The probability distributions used in the BBN will then automatically be adapted to these conditions by multiplication factors that are contained in the database.

2.15.2 Direct setting

The user can also set parameters of the BBN directly. In that case the user is responsible for the internal consistency. However, in CATS, integrity is guarded to the extent that when parameters or distributions are set and some of these are closer to the accident (i.e.less nodes and arcs away from the top event of the BBN) than others, and there is an inconsistency between the farther away values and the closer values, the values of the parameters closer to the accident are given priority. This will be reported to the user in the output of CATS.

2.16 Engine

The settings selected by the user are presented in the ENGINE settings. An example is given in Figure 18. The screen shows all the variables, the units in which the value is expressed, the kind of variable – to be explained later in this report –, the minimum value the variable can take, its maximum, the mean value in the historical data base, which is also the default value, the standard

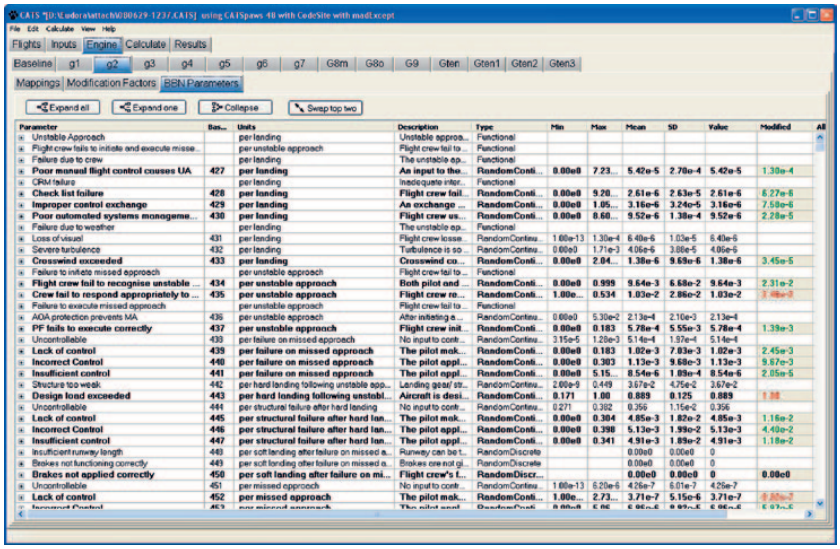


Figure 18: Presentation of the settings

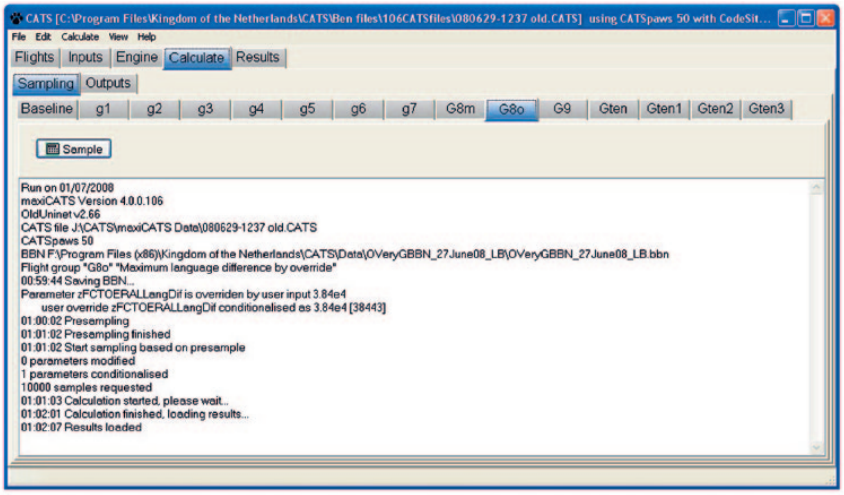


Figure 19:
Report of the calculation

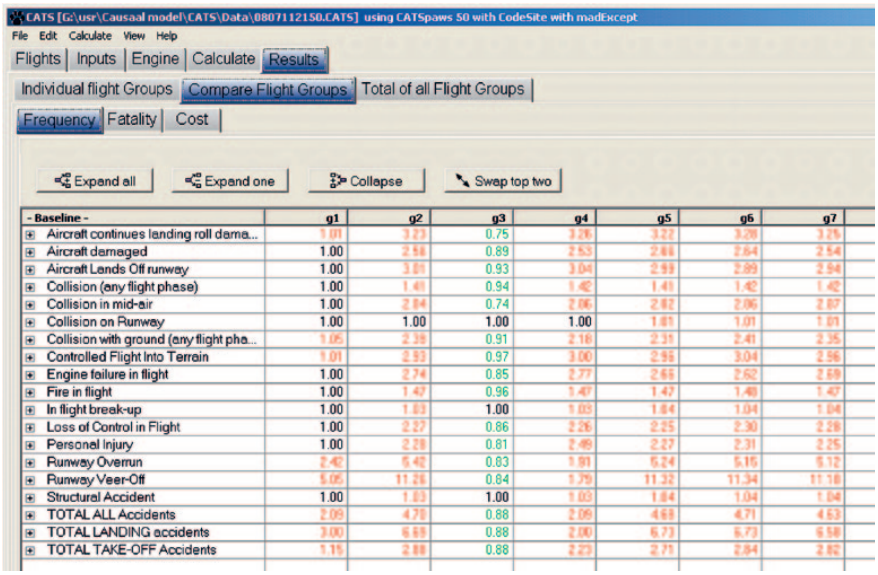
deviation, if it is a distributed variable rather than one with a fixed value, and the current setting. If the setting chosen by the user is out of the allowable range given by the minimum and maximum value, the value will be set to the extreme of the range later in the operation of CATS.

2.17 UNINET

The user determines the number of samples to be taken and from this point UNINET begins its calculation. It takes the user BBN parameter settings and starts calculating then feeds the data back to maxiCATS for the user to see. Once the calculations are done the user gets a report on them, including any adaptations to the parameters that were needed to retain consistency (Figure 19). This information is also written back to the data base for this case file. The final result can be observed in the Results screens. The user can see the calculated probability of any node in the BBN, get aggregated results such as the probability of a veer-off on landing and compare the results of the various flight groups they may have defined. This comparison can be done in terms of absolute probabilities and in terms of relative changes with respect to the base line, as is depicted in Figure 20. The total of all flight groups, taking account of number of flights, is also given.

2.18 CATSPAWS

CATSPAWS is supplied as a separate program by which the displays of BBN node characteristics can be directly manipulated. The connection between the BBN and the CATSPAWS database is between the coded identifiers of the nodes. All information in CATS is held in a single database. Detailed descriptors can be supplied and edited through CATSPAWS. That is also the way to adapt values when new information is obtained by analysis of data. CATSPAWS will identify inconsistencies that may develop in the coding of nodes and of missing data.



- Baseline -	g1	g2	g3	g4	g5	g6	g7
Aircraft continues landing roll dama...	1.01	3.23	0.75	3.26	3.22	3.28	3.25
Aircraft damaged	1.00	2.58	0.89	2.53	2.68	2.64	2.54
Aircraft Lands Off runway	1.00	3.01	0.93	3.04	2.99	2.89	2.94
Collision (any flight phase)	1.00	1.41	0.94	1.42	1.41	1.42	1.42
Collision in mid-air	1.00	2.84	0.74	2.06	2.82	2.06	2.87
Collision on Runway	1.00	1.00	1.00	1.00	1.01	1.01	1.01
Collision with ground (any flight pha...	1.05	2.35	0.91	2.18	2.31	2.41	2.35
Controlled Flight Into Terrain	1.01	2.93	0.97	3.00	2.95	3.04	2.96
Engine failure in flight	1.00	2.74	0.85	2.77	2.65	2.62	2.69
Fire in flight	1.00	1.47	0.96	1.47	1.47	1.48	1.47
In flight break-up	1.00	1.03	1.00	1.03	1.04	1.04	1.04
Loss of Control in Flight	1.00	2.27	0.86	2.26	2.25	2.30	2.28
Personal Injury	1.00	2.28	0.81	2.49	2.27	2.31	2.25
Runway Overrun	2.42	5.42	0.83	1.91	5.24	5.15	5.12
Runway Veer-Off	5.05	11.26	0.84	1.79	11.32	11.34	11.18
Structural Accident	1.00	1.03	1.00	1.03	1.04	1.04	1.04
TOTAL ALL Accidents	2.09	4.70	0.88	2.09	4.69	4.71	4.63
TOTAL LANDING accidents	3.00	6.69	0.88	2.00	6.73	6.73	6.58
TOTAL TAKE-OFF Accidents	1.15	2.88	0.88	2.23	2.71	2.84	2.82

Figure 20: Final results compare flightgroups

2.19 UNISENS

Once the calculations are finished, further processing of the data can be performed using UNISENS. With this program importance measures can be derived. These are used to determine what factors are mainly important for changes in the probabilities of accidents. These can be presented in various graphical forms using UNIGRAPH.

UNISENS can also be used to evaluate correlations. As an example it could be derived what sort of accidents are associated with what sort of airplanes. In the figure it can be seen that generation 4 aircraft are associated with the lowest percentile of accidents. That means that they are associated with accidents that, within the accidents, have the lowest probability of occurrence. Further inspection of this graph under fatigue shows that these accidents also are associated with relatively high values of crew fatigue.

Although these examples are from test calculations and the results should be considered with care, they show how powerful an integrated model is. Rather than looking at the effect of a single instrument or a single change, the effect of certain parameters can be seen in the context of simultaneous changes somewhere else in the system that may or may not be the side effect of the intended change.

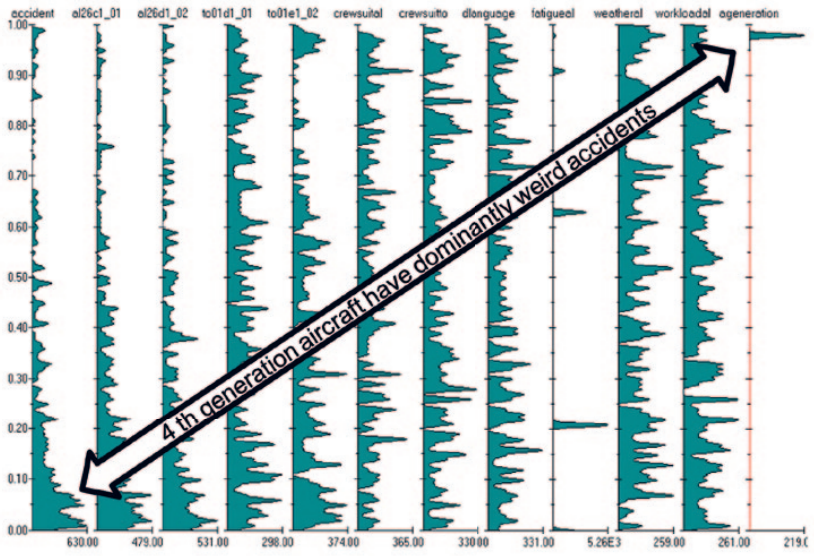


Figure 21:Sensitivity analysis output. Each column represents a distribution of values for the parameter mentioned at the top.

2.20 Data

A number of data sources were used as a basis for quantification.
Airclaims and ADREP were the primary accident data sources. The time period considered was 1990-2003. This period provided a dataset that is large enough for quantification and is considered representative for 'current' air transport. When only Airclaims was used the time period was slightly expanded to 1985 - 2005 to provide a larger data sample. Because of the size of most databases involved, much of the initial analysis was done by running queries, e.g. looking for particular key words. Each incident in the resulting dataset was then individually analysed to verify whether it 'fitted' the particular ESD under consideration. The quality of information in ADREP about accident causes is not sufficient to support the present analysis. Therefore original accident investigation reports have been used where available. In other cases, the summary information from Airclaims, Aviation Safety Network, Flight Safety Foundation and others explains the causes in sufficient detail to relate to the barrier model. Incident reports have also been used where available.
The primary source of data for quantification of the probability of occurrence of the initiating event are the databases of Service Difficulty Reports and Air Safety Reports but sometimes other sources of data were used if these were considered to be more accurate.

2.20.1 Using Storybuilder to analyse ADREP data

The Ministry of V&W supplied the ADREP database containing around 35,000 accidents together with the ECCAIRS software which is used by aviation authorities to store occurrence data. Using this database it was intended to superimpose accident scenarios of interest (fatal accidents since 1990) onto the ESD integrated model structure, combined with the developed Fault Trees of DNV. This could not be done automatically because many accident data records proved to be incomplete. As an example only a few percent of the CFIT accident reports supply any information with regards to the status of the GPWS. Therefore it would have been necessary to go to original accident reports for accuracy and detail beyond what was available in the ADREP/ECCAIRS database because many fields in the database are not filled in. The analysis of accident data by NLR and DNV has only enabled its classification as scenarios in the ESD backbone, and then only a limited number of accidents have been so classified according to reference to the ADREP state file numbers.
Storybuilder (Bellamy et al, 2007) was used to capture data from the ICAO ADREP database and combine these data with some of the modelling within CATS, ultimately performing a bookkeeping function as well as the basic groundwork showing the feasibility of using ADREP data for validation purposes. 8969 occurrences are to be found in this final bookkeeping Storybuild. This provides the foundation for future bookkeeping, checking and validation work, using the import function of Storybuilder to acquire the necessary data. Storybuilder is a graphical interface tool tailor-made for developing scenarios for accident modelling (Bellamy et al, 2000; Bellamy et al, 2006; Bellamy et al, 2008). The tool was developed in the Dutch WORM project (Ale, 2006 and was available for use and further development in the CATS project..
Hundreds if not thousands of scenarios can be captured in a single storybuild diagram. The underlying database that defines the graphical model properties is structured in MS Access. Storybuild files (*.sb) can be created and opened in Storybuilder and in Access. Data contained in an sb file can be exported in different tabulated forms which for convenience is done in MS Excel or

MS Word. For the CATS project the following functionalities were developed:
A small component of the overall CATS model called MiniCATS was developed using the CFIT backbone component (ESD35) and the Fault Trees and BBN attachments to the pivotal events. This was developed as the "Storybuilder Evaluator" and was the first software model of the CATS risk assessment used in the "miniCATS" pilot 23 November 2006. JPSC Ltd developed this miniCATS program in Delphi to calculate this one backbone segment. The model took inputs (equations) from the Fault Trees of DNV and BBNs of TUDelft. Inputs could be modified to different values to show how the outputs change.
A printing function and user-defined placing of boxes was developed to enable the integrated ESD backbone model to be both printed and be more visible in Storybuilder. Instead of having automated box layout the user could hand place the position of boxes allowing a better graphical representation of the events.
Modifications were made to Storybuilder to improve the import function for importing ADREP data from exports in MS Excel from ECCAIRS, as well as for the analysis of audit data from IVW's BReS (Bedrijfs Registatie Systeem) database.
Storybuilder was used subsequently to model the accident sequences for the integration of the Event Sequence Diagrams produced by NLR into an integrated single gate to gate model (ESD backbone). The backbone formed the basis for establishing the feasibility of developing one giant BBN. The alphanumeric coding of backbone events in CATSPAWS.
Based on these, the quantifications supplied by DNV and NLR could be cross-checked on a number of topics:

1 Whether all the accidents relevant for a certain quantification were indeed used
2 Whether all accidents were only used once
3 Whether all the accidents making up the total number of accidents or another aggregate were used in the underlying analysis

2.20.2 Scenarios

In order to perform the analysis the concept of scenarios of storybuilder were used.
A scenario was defined as a path along the yes-no decision tree of the ESDs. Scenarios were numbered from top to bottom of the consequences in the ESD diagrams with the all-Yes scenarios being 1 eg. ESD35-01, ESD35-02 etc. where ESD35 is the 35th ESD diagram in the NLR report. This coding system was later developed further to be included in CATSPAWS. All end events of the ESDs have a unique coding which represents a unique scenario (pathway) through the ESD backbone. There were 723 nodes with 328 generic scenarios passing through these.

2.20.3 ADREP Data

The International Civil Aviation Organisation (ICAO) Accident/Incident Data Report (ADREP) database was supplied for the purpose of assisting in the modelling and quantification. The type of incidents which are of main interest to the International Civil Aviation Organization for accident prevention studies are listed in the ICAO Accident/Incident Reporting Manual (Doc 9156-AN/900 Accident/incident reporting manual ADREP 1987, ICAO). In accordance with Annex 13 - Aircraft Accident Investigation, States report to ICAO information on all aircraft accidents which involve aircraft of a maximum certificated take-off mass of over 2 250 kg. The Organization also gathers information on aircraft incidents considered important for safety and accident prevention. For ease of reference the term "occurrence" is used to signify both accidents and incidents.

In the CATS project ECCAIRS (release 4) is the software package developed by the EU that was used to access the reported occurrence data. The European Aviation Safety Agency uses the ECCAIRS software to store occurrence data. The ICAO ADREP 2000 taxonomy has been implemented in ECCAIRS (European Co-Ordination Centre for Aviation Incident Reporting Systems). The Dutch incident analysis bureau (ABL) also use ECCAIRS but with a much shallower taxonomy. Directive 2003/42/EC of the European Parliament and the Council of 13 June 2003 on occurrence reporting in civil aviation make occurrence reporting compulsory from 2005. ABL now has at least one complete year of reported occurrences for The Netherlands. These data could in the future provide much sought data for model validation and further development when it becomes possible to access to this database.

The total number in the database provided was 34912 occurrences from 2/01/1970 (but also one accident from 1953) until 20/05/2007. In ECCAIRS the query building for simple queries for data sets (e.g. all CFIT accidents) was not too difficult, but more detailed queries and export of data to say an MS Excel file was found to be extremely user-unfriendly. The basic problem is that although one can select a set of data fairly easily using a build query language, to export to Excel or other format the user has to select each attribute parameter value for export individually.

Since the list of attributes is in itself a 100 page long document (R4LDAttributesvaluesbyattributeid.pdf) - available from ECCAIRS website - this is no mean feat. This document does not, for example, include missed approach as an attribute or value. But that does not mean missed approach is not in ADREP. The problem is knowing how to find it, although it is recognised that normally appropriate training in using ECCAIRS is required. These problem issues have been further elaborated in a report which has been issued to a contact at JRC (Bellamy, 2007). For finding variables of interest a pdf was created using all available ADREP taxonomy documents, which was named the Giant Dictionary. This document has 1665 pages and includes the following taxonomy documents

- Aircraft type designators by designator
- Aircraft type designators by manufacturer
- Aircraft make/models
- ANS Services
- ATM Ratings and endorsements
- Aviation operations
- Descriptive factors
- Engines
- Event phases
- Events
- Explanatory factor
- Fuel types
- License types
- Location indicators by indicator
- Location indicators by state
- Modifiers
- Occurrence classes
- Operators by state
- Organisations and persons
- Propellers
- Recommendation types
- Report forms and types
- States

- Attribute values by attribute id
- Entities and attributes
- Entity structure
- ICAO ADREP 2000 taxonomy
- Topics sections and attributes

The most important needs from the accident data are:

- Identify subsets of accident types as mutually exclusive groups
- Quantify influences, cause, event and consequence frequencies and simultaneous occurrence of events
- Avoid double counting, and have completeness and accuracy in counting

Many different queries were made to extract data of use to team modellers, notably for TU Delft for management modelling and DNV for quantification. In terms of providing data to DNV it was found that ultimately the best approach was to provide a complete text file exports of all the data in a query. Although this produced documents several hundred pages long, for the various information requests from DNV it gave a better overview and offered search facilities within MS Word or Adobe Acrobat. However there were cases where specific exported data selections were asked for. For example, for Fault Tree Modelling the following were considered needed:

ADREP Taxonomy Name	Taxonomy code
Location of Occurrence	440
State of Occurrence	454
Occurrence Class	431
Flight Phase	121
Maximum Take-off Mass	175
Occurrence Category	430
Manufacturer Model	21
Operation Type	214
Operator	215
Injury Severity Level	451
Mass Group	319
Propulsion Type	232
Narrative	425
Turbulence Intensity	293
Visibility Restrictions	311
Amount of cloud	266
Light conditions	168
Weather conditions	127
Precipitation Intensity	230
Runway Length	501
Experience all a/c	410
Experience this aircraft	411
Dew point temperature	85
Air temperature	287

And for management modelling the following were considered useful:

ADREP Taxonomy name	Taxonomy code
Occurrence category	430
Occurrence class	431
Event Type	390
Flight phase	121
Phase	391
Descriptive factor subject	385
Descriptive factor modifiers	386
Organization/person	394
Explanatory factor subject	392
Explanatory factor modifier	393
Maintenance docs	174
Duty last 24 hours	403
Rest before duty	408

Various tests of importing data into storybuilder were made using the state file as the occurrence name to examine the general feasibility of importing and analysing ADREP data en masse. The value of such imports is that it prevents double counting (an occurrence can only pass through an event box once) and subsets of data can be easily selected. This then avoids the long wait for exporting of data from ECCAIRS and of the difficulty of then making data counts. The basic query used initially was:

```
ALL from 1990 not russian COMMERCIAL with mass groups and
unrestricted
Find all Occurrences where { Local date. after 31/12/1989 }
and { Aircraft category. equal to Fixed wing } and { Aircraft
manufacturer/model. doesn't have any of ANTONOV,  ILYUSHIN,  LET
AERONAUTICAL WORKS,  TUPOLEV,  YAKOLEV } and { Operation type.
equal to Commercial Air Transport  } and { [ Mass group. equal to
27 001 to 272 000 Kg ] or [ Mass group. equal to > 272 000 Kg ] or
[ Mass group. equal to 5 701 to 27 000 Kg ] } COUNT 8665
```

2.20.4 Bookkeeping

8969 occurrences were imported into Storybuilder using query 79 (Annex I) described earlier with the Excel export of some of the ADREP data.
The following data were included:
ADREP data categories:

- Damage to aircraft
- Injury
- Make
- Mass
- Model
- Occurrence category
- Occurrence class
- Operator
- Year
- CATS data categories
- Aircraft generation
- DNV scenario categories

- Delivery systems
- ESDs of the DNV fatalities
- ESDs of NLR
- NLR ESD end states
- Extra ESD end events of NLR added later
- NLR damage proportion aircraft
- NLR dead proportion
- NLR year of NLR dataset (because include before 1990)

In carrying out the analyses a generally good correspondence was found between the DNV and NLR ESD classifications. Differences have been investigated and where necessary the data have been amended.[2]

2.21 Limitations

With a first edition of a system such as CATS limitations are unavoidable. As will be described in the next chapters, building and calibrating the model has exhausted the available data. It was therefore not possible to perform an independent validation using an independent dataset.
Human performance modelling is still difficult territory when it is aimed at quantifying the probability human error in particular settings. The expert judgement methods used in CATS are what are available but the maintenance model especially needs further development and pertinent data are rare and because of confidentiality difficult to obtain. Given the importance of the role of human operators in the air transport system this is an important weakness, not only in CATS, but also – and this is even more serious – in managing risks in the real world. Apparently the responsible authorities inspect, take measures and manage safety in the absence of pertinent information about the effect – be it good or detrimental or absent – their actions have on the safety of air transport.
The depth of the modelling was determined by opportunity. When there were data the model goes deeper than when there were not. This means that the model needs further extension when it is used for even more detailed analyses.

2.22 Usability

Despite these limitations CATS is already a powerful tool in the analysis of the air transport system. It allows detailed calculations of probabilities of accidents and of intermediate events. It is based on a detailed analysis of world wide data and on the insights of experts. The use of an integrated BBN makes it possible to take the many interdependencies into account in an integrated way, which makes the model much closer to reality that any previous modelling could.
CATS makes it possible to investigate what happens if extreme values of distributed quantities combine and the preliminary analyses show that these extremes indeed correspond to the rare but nevertheless normal accidents.
This use of CATS may prevent accidents in the future. It allows the user to vary values and parameters on a scale that cannot be done in real life, not even in the extensive test and certification processes common in the industry.
The model can be used in industry and by policy makers to support policy development and decision making with risk information. Future changes in operation and oversight arrangements such as changing the structure of air

[2] Exported data set from Storybuilder is shown in Annex WQ 4.
Exported data using just the 325 (327 aircraft) DNV fatalities (since 1990) is given in Annex IV

traffic management can be analysed on their risk implications.
Insight in risk is a necessary precondition for effective risk and safety management and CATS can be used to generate this insight.
As a comparative tool, i.e. to compare different strategies and different measures, CATS is the most powerful. The absolute values of the outcomes are as close to reality as could be made. However, this is the first integrated model of its kind. Never before could the effect of interdependencies be studied in a comprehensive system. Therefore it is wise to exercise caution.

2.23 Further work

In the near future further development is needed on five fronts

1 Further validation by using the model for studies of various problems in the air transport system, such as changing air traffic management or the changing pattern of air traffic caused by the changing behaviour of the air traveller. This could involve further detailing of the model e.g. replacing "airport complexity" by parameters that constitute complexity.
2 Further analysis of data, especially data on human performance. Amongst others these analyses should comprise acquiring and analysis of audit data.
3 Improvement of the human response model, particularly the model for maintenance and the actions of management in conjunction with the modelling of his operational environment as shaped by company management. This involves replacing the current proxy variables by the real variables and modelling the management organisation.
4 Improvement of the management influence model. This includes a further enrichment of the – abstract – deliveries in terms of actual actions
5 Further acquisition of (confidential) data. This could involve co-operation with database owners to facilitate the extraction of information and the transfer of this information into the CATS data set.

Finally a few preliminary results indicate that fatigue should be a major concern and further research into what constitutes workload, the onset of saturation, boredom and sleep deprivation and the relationship could be beneficial for future safety.

3 General Methodology

The methodology used in CATS differs from the technology used in many risk analysis models by its use of a single model structure based on the theory of Bayesian Belief Nets (BBNs). In CATS the traditional techniques such as Fault Trees, event trees and influence diagrams are combined in an integrated network model. The model is constructed such that more familiar representations of accident causation can be recognised in the structure. This facilitates discussion with and among experts in the aviation field about technical issues and measures. It also allows the future addition of links to safety policies, design standards and legal requirements.
In this chapter we first describe the models and metaphors that are used in accident causation modelling. Then the design principles of the model are described.
Section 2 of this chapter has much mathematics. The reader may skip the formulae and still obtain a general picture of what the mathematics are meant to do.

3.1 Accident Causation Metaphors

There are various metaphors used to capture the essence of accident causation and the protection against harm. These are amongst others the Hazard-Barrier-Target model, the Swiss Cheese Model and the Bow-Tie model.
These are described in the following sections.

3.1.1 Hazard Barrier Target Model

The Hazard barrier Target Model sees accidents as the result of a continuous threat, the hazard, on a target (Schupp et al, 2004). This target is shielded from

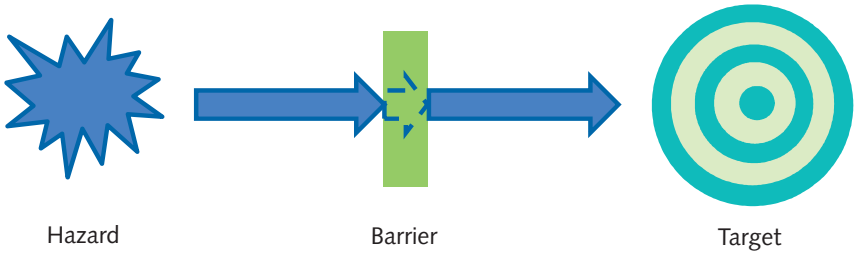


Figure 22: Hazard Barrier Target Model

the hazard by a barrier. There may be one barrier but there could also be several. These barriers need to be maintained. The maintenance of the barrier has to be secured by a – barrier – management system, also called safety management system or SMS. Barriers may be imperfect or absent as result of technical or human failure.

3.1.2 The Swiss Cheese Model

The Swiss Cheese Model (Reason, 1999) takes the Hazard Barrier Target Model one step further. It recognises that no barrier is ever perfect and that several barriers have to be in place to prevent a cause from progressing to an accident. These barriers can be of any type. Typically design, construction, operation and maintenance are included. Defects in the barriers can be latent, as other barriers prevent the progression of the cause to an accident.

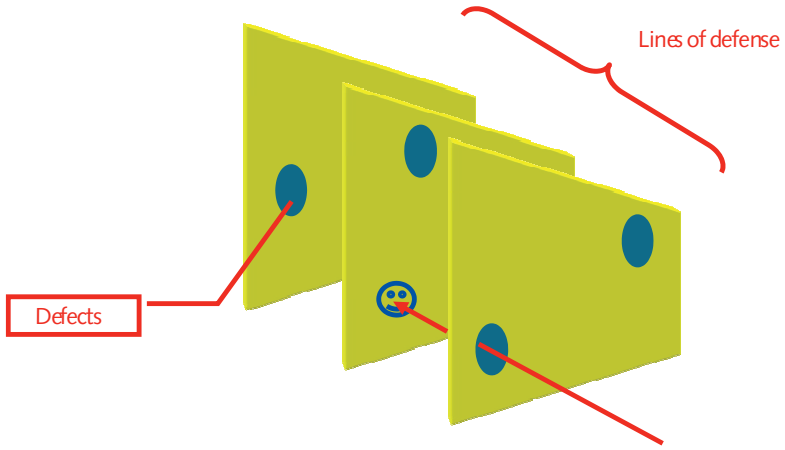


Figure 23: Swiss cheese model

This is also referred to as endemic problems in a system or organisation. There are numerous variations on the same theme, but the original by James Reason remains the most consistent (Reason, 1990).

3.1.3 The Bow-Tie model

In the Bow-Tie model not only the causes of accidents are considered but also the consequences , such a material damage, health effects or death (Visser, 1998). Barriers can be put in the path from cause to accident, but also from

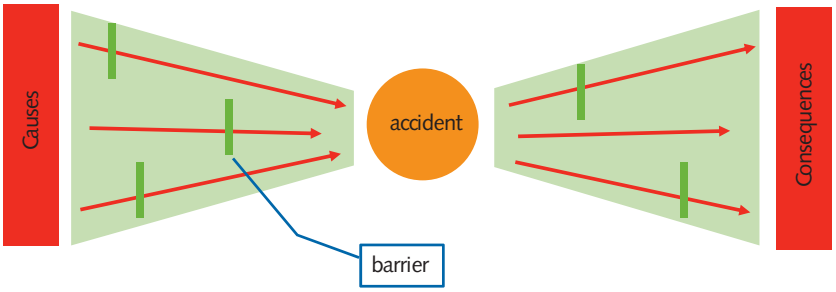


Figure 24: Bow-tie model

accident to consequence. As there can be many causes for a single accident and an accident may have a variety of consequences, diagrams depicting this idea have the form of a Bow-Tie, hence the name (Figure 24). This idea was further developed into the TRIPOD method. Here the latent defects are considered to be precursors and preconditions, which can exist for a long time, before the accident happens and which tend to grow in severity when they are not detected (Groeneweg, 1998). The Bow-tie and Barrier model was embraced by the EU in their attempt to harmonise the thinking about accident causation and prevention in the chemical industry in the ARAMIS project. (Debrey et al, 2004; Duijm et al, 2004; Duijm et al 2004a).

3.1.4 Causality

The concepts and metaphors described above originated from the practical question of where and how to prevent future accidents. For the very rare accidents the underlying causes are sometimes difficult to identify. Furthermore, every underlying cause has another deeper cause. Thus the thinking about accidents evolves into ever deeper constructs and models. Nevertheless any effort to construct a model describing chains of causality of events in a system must be based on the assumption that causality exists and that causality even in systems as complex as the aviation industry can be described. This unavoidably leads to the question: whether the work will ever be finished and whether there is a *last cause* to be found.

Several lines of discussion are continuing. Some of these are triggered by the

perceived incomprehensibility of low probability – high consequence events. Some of these by the notion that analysis of causality seems to have no end and some by the more legalistic discussion on whether a probabilistic progression of a sequence of events should lead to a negation of the certainty of the cause after the fact.
The matter of causality is a highly philosophical question. We describe our position with respect to these questions briefly below, in order to justify the continuation of our efforts to those in the scientific community that have reached the point of seeing no further point in causal analysis and modelling.
The discussion about the infinity of the chain of causality is an old one and goes back to the Greek atomists some 400 years BC (Russel, 1946). The why question in this context can have two meanings: "to what purpose" and "with what cause". Both questions can only be answered within a bounded system, because they imply that there is something causing the system to exist. The cause of the system to exist can be referred to as a "maker", who self is not part of the system, but resides outside the system.
A bounded system can show behaviour that the makers of the system did not anticipate. In most cases the cause of this behaviour can be found as a combination of behaviours of parts of the system that the makers of the system did not consider. Projective analyses take time and effort, and efficiency demands these analyses to be limited. The fact that a behaviour was not anticipated does not imply that anticipation was impossible, merely that it was deemed impractical.
Nevertheless, one could make the proposition that complex systems show unexpected behaviour that is not only surprising, but could not be anticipated in principle. This is referred to as emergent behaviour, meaning behaviour that spontaneously arises, with no cause and was not the result of any of the properties or combination there-of that the makers put into the system. It is argued that such behaviour can only be shown by living organisms (Chalmers, 1996). We share this position. However although human beings are part of the aviation system we take the position that the aviation system is put together by humans and run by humans but is in itself inanimate. (Arshinov, 2003).
As regards causality in the "legal" sense, this is an issue that also plays a role in the discussion about flood defences: what causes a flood: high water or a low dike. This is a question like what is the contribution of the left hand to the noise when clapping hands. We consider the cause of the flood to be the combination of height of water and height of dike where the latter is lower than the former. A cause therefore is a multi-attribute entity. More generally a cause is the occurrence of a particular combination of the values of relevant parameters that give rise to an effect.

3.1.5 Complexity

In many cases the causes of accidents even in complex systems seem to be the straightforward result of a single technical failure or a simple human error. However the world is not that simple. In many cases these causes are rooted in the way people and organisations deal with technology, its advantages and its risks. It is the shape of the human organisation in which the technology is used that often conditions the scene towards the moment that the poor victim steps on the slippery slope sliding towards disaster. In CATS an attempt is made to model the relevant properties of the organisation, called deliveries and delivery systems, to grasp the complex interrelations between organisation and technology, in an attempt to help the user to identify weaknesses in the organisation that may have caused an accident by a particular path on this particular

occasion, but may as well have caused accidents through another path for the same organisational reason on some other occasion (Oh et al, 1998; Papazoglou et al, 2002 and 2003; Bellamy et al 1999)

3.2 Graphical Models: Event Trees, Fault Trees and BBNs

We compare and contrast three common types of graphical models used in risk and reliability theory. Event Trees, Fault Trees and Bayesian Belief nets (BBNs) are all used in the CATS model. We follow customary parlance that "arc" denotes a directed arrow between two nodes. Event Trees and Fault trees both have implicit directionality associated with their links, and we therefore term them arcs, even though they are usually drawn without arrowheads.

3.2.1 Event trees

Event trees consist of two basic elements. Nodes and arcs. Nodes are also called events. Even if there is not really anything that "happens". The arcs connect the events. They are usually represented by arrows, indicating the logical or temporal progression through the tree. In most cases the logic of the tree runs from cause to consequence. The cause is also called the parent, the consequence is called the child. Nodes generally have only two states,. Yes, or True and No, or False. We restrict attention to two-valued events. If we exit the parent node via the 'true' arc, then all subsequent events on this path are conditional upon the Parent being True; similarly for False.

Since the tree has a direction: it belongs to the class of Directed Graphs. Since a child cannot be the parent of it ancestors, a cycle is not possible. So an event tree belongs to the class of Directed Acyclic Graphs or DAGs. A distinctive feature of event trees among DAG's is that they have exactly one source node, that is, one node without a parent. On the other hand they have many sink nodes – nodes without children. These correspond to all possible end states of the system being modelled.

Although we tend to think that time progresses if we go from parent to child to grandchild etc., this is actually not what the tree represents. The tree is strictly LOGICAL. This means that the states set themselves instantaneously.

Therefore in event trees, events can be strung together that do not have a

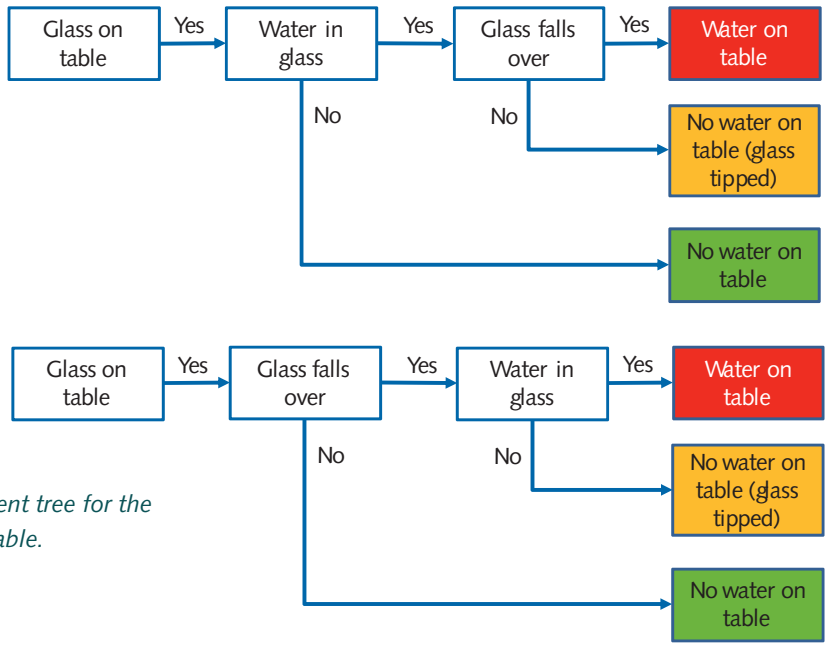


Figure 25: Two versions of the event tree for the same accident: water spilled on table.

causal relationship. For instance to have water spilled from a glass (Figure 25), there has to be water in the glass and the glass has to fall over. The water in the glass is not the cause of it falling over, nor the other way around. In an event tree the two events: water in the glass and glass falls over can occur in any order.
The two states of each node have a probability associated with them. The total of these probabilities must be one. The probabilities are the chances that the state of the node is True or False. It is a conditional probability, conditional on all predecessors. This means that the total probability of the true state of a node is the probability that the node is reached in the tree multiplied by the probability of going into the true exit.
One can think of a node in an event tree as a street crossing with one side street on a road. One can either go straight or exit into the street.
Crossroads often offer three alternatives: left, right and straight. In event tree language they have to be modelled as two nodes: first a node for straight versus turn, and second a node for turn left or turn right.

3.2.2 Fault trees

Fault trees have three basic elements: nodes and arcs, as in event trees, and gates. In event trees the tree branches out at every node. Fault trees combine arcs in gates. The logical flow is from basic events to a "Top Event". Cycles are forbidden, so Fault Trees are also DAGs. In contrast to event trees, there are many source nodes – these are the basic events of the Fault Tree. There is typically one sink node – a node without children, namely the Top Event. Whereas Event Trees model the possible states of a system, Fault Trees model the possible ways in which a given event can occur.
A node can have only two states: true or false. This state is completely determined by the state of the parents and the type of gate in which they combine. If they combine in an AND gate, all the parents have to be True for the child to be True. We note that this is an asymmetrical property. It is not necessary for all the parents to be False for the child to be False. If only one parent is False, the child is False. (Figure 26).

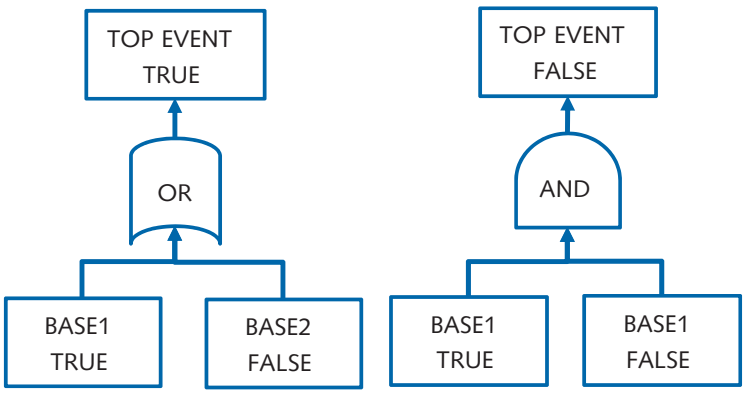


Figure 26: Examples of OR and AND gates

If the parents combine in an OR gate, only one of the parents need to be True for the child to be True. This is again asymmetrical. If only one parent is False the Child is still True if one of the other parents is True. There are more types of gates, but they occur less frequently.
There is one final node. This node is reached when the last remaining parents combine in a single gate. This event is called the top event of the tree.
The states of the base events can have probabilities associated with them. It is sufficient to know the probability of the True state for each base event. The

probability of the false state is 1 minus the probability of the True state. This is also called the one-complement. Once these probabilities are known the probability of the True state of top event follows from a straightforward calculation, for which there are simple rules. These will be discussed later.

3.2.3 Interdependencies in fault and event trees

In many real systems the logic is not as straightforward as described above. In many cases what goes on in one branch of the tree actually influences what goes on in another branch. For example a system might have two redundant pumps for coolant. For the cooling system to fail it is sufficient that both pumps fail (the cooling system can also fail due to pipe break, but loss of both pumps is enough to bring it down). Under normal circumstances the probability of both pumps failing independently would be the product of their individual failure probabilities. However, if they both depend on the electrical system, a power failure would take out both pumps together. In this case the pumps are said to fail due to a common cause. Whenever possible, common cause dependencies should be modelled in the Fault Tree. This is easily said, but often less easily done. In scramming a nuclear reactor, an accident may result if three or more neighbouring control rods fail to insert properly. These rods may fail independently. However, of greater concern is the possibility that they fail due to some common cause. In this case the set of possible common causes is more difficult to identify and more difficult to model in the Fault Tree. Mathematical models must be employed. Detecting the effect of common causes from data is often very challenging.

Whereas hardware common causes can in principle be identified and modelled with engineering knowledge of the system, with other common causes this is more problematic. A few examples illustrate this point:

1 A fire near a cable tray can disable multiple components simultaneously,
2 A flaw in the training of maintenance personnel can cause maintenance errors to occur simultaneously in physically separated sub-systems,
3 An error in the building design may force evacuation of the control room in certain otherwise non-critical situations, thus increasing the chance of operator error.
4 A poor safety culture at top management levels can degrade operator alertness, maintenance performance, training, acquisition and testing of spare parts, quality of emergency procedures, etc, etc.
5 A flaw in the regulatory regime may allow unsafe situations and practices to arise simultaneously in an entire population of systems.

All of these features can increase the probability of multiple simultaneous failures in a system. Hence if we simply gathered failure data from system components, and assume that these systems can only fail independently, we may produce an unrealistically optimistic prediction of system performance.

Dependence modelling in Fault Trees works well when dependences can be associated with failures of support system hardware components. Such support systems might include the electrical system, the sprinkler system, the lubrication system, and the software control system. In such cases, the failure of a support system *causes* the failures, or unavailabilities, of multiple components.

Other dependences, such as numbers (2) – (5) above, do not express themselves directly by causing simultaneous component failures. Rather, they simultaneously influence the *probability* of failure of multiple components. Examples include poor maintenance, poor operator training, poor incident reporting, poor safety culture, etc. Fault tree modelling cannot readily capture dependences that influence the probabilities of failure. Influences acting on the probability of failure, rather than on failure itself, must be captured in the uncertainty analysis

of Fault Trees. Bayesian Belief Nets are a modelling tool specifically designed to capture probabilistic influence. Before we can fully understand uncertainty analysis of Fault Trees, we must understand the difference between Boolean and ordinary arithmetic.

3.2.4 Boolean Arithmetic for Fault Trees

A Fault Tree is just a picture of a Boolean formula. In Boolean arithmetic, variables take only the values 0 or 1. Suppose X and Y are Boolean variables, in Boolean arithmetic, $X +_b Y$, and $X \times_b Y$ are also Boolean variables, and hence take values 0 or 1. This means that $X + Y$ must take only values 0 or 1, which is arranged by defining $X +_b Y = X + Y - XY$; $X \times_b Y = XY$. In Boolean arithmetic, addition and multiplication correspond to the operators AND and OR in propositional logic. Thus, $X +_b Y$ = "either X or Y or both", $X \times_b Y$ = "X and Y". In particular, in Boolean arithmetic, $X^2 = X$; this corresponds to saying that the event X AND X is the same as the event X.

A Fault Tree is a Boolean formula, when we fill in 0's and 1's for the basic events, and apply the AND and OR operators, we always obtain a 0 or 1 for the top event. In most cases, we don't know whether a given basic event occurs, we know only its probability of occurrence. The probability of event X occurring is the expectation of the random variable X; since

$$E(X) = Prob\{X=1\} \times 1 + Prob\{X=0\} \times 0 = Prob\{X=1\}.$$

If the events X and Y are independent, then it is easy to check that

$$E(X \times_b Y) = E(X)E(Y);$$

$$E(X +_b Y) = E(X) + E(Y) - E(XY) =$$
$$Prob\{X=1\} + Prob\{Y=1\} - Prob\{X \text{ AND } Y = 1\} = Prob\{X \text{ OR } Y\}.$$

This might suggest that we can just replace the Boolean variables at the base of a Fault Tree with their probabilities (i.e. their expectations) and compute the probability of the top event with ordinary arithmetic. This is NOT true, in general, and it may depend on how the Fault Tree is displayed. A simple example illustrates this feature. Suppose the Top Event occurs when either X AND Y or X AND Z occurs. We could represent this with two simple, logically equivalent, Fault Trees:

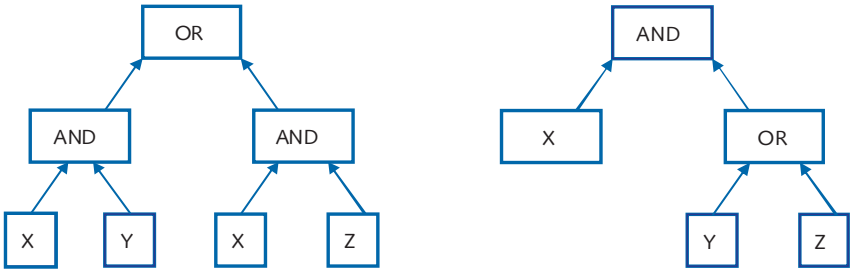


Figure 27: Two Fault Trees

The Boolean formula from the left tree is

$$X \times_b Y +_b X \times_b Z = XY + XZ - XYXZ$$

The Boolean formula from the right tree is

$$X \times_b (Y +_b Z) = X(Y + Y - YZ) = XY + XY - XYZ$$

If we apply Boolean reduction to the first formula: XYXZ = XYZ, we see that these two formulae are equivalent. However, if we replace the variables by their expectations and apply ordinary arithmetic, we will get different answers. From the first tree we would get the incorrect formula:

Prob(Top Event) =
Prob{X=1}Prob{Y=1} + Prob{X=1}Prob{Z=1} Prob{X=1}Prob{Y=1}Prob{X=1} Prob{Z=1}

The correct calculation would be obtained from the right tree:

Prob(Top Event) =
Prob{X=1}Prob{Y=1}+Prob{X=1}Prob{Z=1}- Prob{X=1}Prob{Y=1}Prob{Z=1}

The problem is that when we compute (*X* AND *Y*) AND (*X* AND *Z*) with expectations in the left tree, we would include the term "P{*X*=1}" term twice.
In general, computing probabilities of occurrence from a Fault Tree requires some careful manipulations, before substituting Boolean variables with their expected values. However, if our Fault Trees contain no "repeated events" then we can replace Boolean variables with expectations and replace Boolean arithmetic with ordinary arithmetic. This assumes that we have captured all common cause dependencies in the Fault Tree. This means that once probabilities are assigned to the basic events, the probability of joint occurrence is computed as the product of the probabilities.
When repeated events are present, we can estimate the probability of the top event by various methods. One method is to reduce the Fault Tree to a *minimal cut set* representation: we write the top event in a special form known as the cut set equation. A prose reading of the minimal cut set equation would be something like:

Top event happens if and only if EITHER:
BE1 And BE5 And BE23
Or
BE23 And BE8 And BE31 And BE26
Or
Etc.

We can estimate the probability of occurrence using the inclusion-exclusion principle. However, if the combinations which fail the system have low probability, then we may approximate the probability of the disjunction as simply the sum of the probabilities of the combinations. We do not explain this further here, as this material is available in any standard text, and the Fault Trees in the CATS model do not involve repeated events. In the CATS model, we can simply replace basic events with expectations and compute with ordinary arithmetic.

3.2.5 Bayesian Belief Net

A Bayesian Belief Net is a special kind of directed acyclic graph. In a BBN nodes represent random variables and arcs represent probabilistic or functional influence. Since Boolean gates in a Fault Tree are simply functional dependences of a particular sort, and Event trees are probabilistic dependences of a particular sort, it is evident that BBNs are a more general structure than either event of Fault Trees. Thus it is possible to represent both as BBN's. In the CATS model there are event trees and Fault Trees, which are represented as parts of an over-arching BBN. The point of using a BBN, however, is not simply to replicate

modelling of event trees and Fault Trees; rather, the point resides in the fact that BBNs can capture probabilistic influences between random variables – not just two valued events - which cannot be modelled functionally. This enables us, in principle, to capture factors which influence the *probability* of failure for basic events in a Fault Tree, the factors (2) – (5) in section 3.2.3

A simple example helps. Suppose we are interested in a particular human error event. In a Fault Tree this is a Boolean variable, but since our Fault Tree has no repeated events, we may replace this Boolean variable by its expectation. We wish to model those factors which influence this human error probability. Suppose we identify Training and Fatigue as relevant influences. We construct the simple BBN shown below

Human error probability is operationalised as the relative frequency of error per demand. Training is operationalised as 'hours spent in refresher courses over last 3 years' and Fatigue is operationalised on the Stanford sleepiness scale. From data we can recover the distribution of operators training and the distribution over possible fatigue states. We also have human error data which, with some appropriate (Bayesian) procedure will enable us to form a distribution over the probability of human error. Thus we have the individual distributions for all the above variables. However, data does not tell us how Training and Fatigue influence Human Error Probability.

3.2.6 Functional influence

As good engineers, we might try to capture the relation between training with some mathematical model. We could always start with a Taylor expansion around certain nominal values (not the inclusion of the first interaction term):

$$HEP = HEP_0 + (TR - TR_0)\partial HEP/\partial TR_0 + (FA - FA_0)\partial HEP/\partial FA_0 + (TR - TR_0)(FA - FA_0)\partial^2 HEP/(\partial TA_0\, \partial FA_0), \quad (3)$$

This of course is a possible approach, but it requires an appropriate choice of nominal values, a truncation of higher order terms at an appropriate point, and an estimation of partial derivatives. If none of these choices can be supported by data, it might make more sense to capture the influence as probabilistic influence.

3.2.7 Probabilistic influence

In opting for probabilistic influence, we are acknowledging that our dependence modelling is not supported by data, and that detailed functional modelling is not indicated. Instead, we opt for coarse modelling of high / low values to occur together. This tendency is measured as rank correlation.

We may have some multivariate data of simultaneous observations of all three variables. That is observations of equivalent systems s = 1,…K in which we observe (HEP_s, TR_s, FA_s) for each s = 1…K. We could then capture the influence as probabilistic influence in various ways, we might ask, for example:

In how many systems are HEP_s AND TR above their median values?

(HEP is HumanErrorProbability and TR is (level of) Training)

If that number is larger than one quarter of all observed systems, then high values of HEP and TR tend to occur together, if it is lower than one quarter, then high values of HEP tend to occur with low values of TR, and conversely. From this sort of information we could estimate a rank correlation between HEP and TR. The reader is spared the mathematical details for the time being.

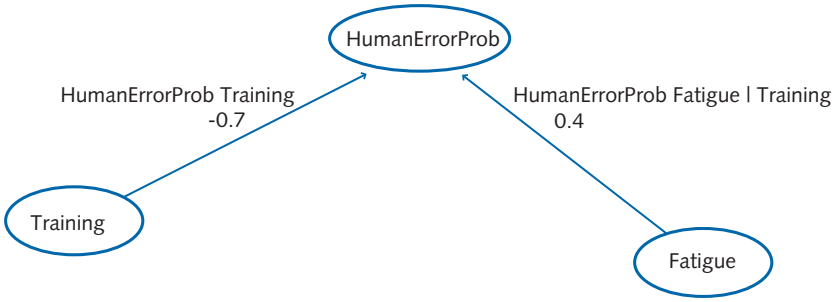


Figure 28: BBN depicting a relationship between the Human Error Probability, Training and Fatigue

To capture the influence of Fatigue in addition to TR, we might ask:

In how many of those systems in which both HEP_s and TR_s were above their medians, was also FA_s above its median?

With this information we could estimate the conditional rank correlation of FA and HEP given TR.
In Figure 28 a possible relationship is depicted between Human Error Probability, Fatigue and Training. The graph in Figure 28 has no influence between FA and TR. That is a modelling assumption which might be checked against data. For example, we might ask whether systems characterised by high values of TR tend to show lower values of FA. For the time being we analyse the above graph.
The reader may accept on faith that the information acquired, together with the independence implied by the graph, together with an assumption on the type of distributions realising the rank correlations, completely determine the joint distribution of HEP, TR and FA. The rank and conditional rank correlations obtained may be shown in the BBN:
The correlation between TR and HEP is negative: high values of training tend to go with low values of HEP.
What if we do not have the simultaneous measurements needed to answer the above questions? In that case we ask the very same questions which we *would* ask to the data, if we had it, to knowledgeable experts. Many people balk at the introduction of expert judgment into quantitative risk studies. Of course, real data is always preferable. On the other hand, it would be very unscientific to assume that influences for which we have no data, therefore do not exist.

3.2.8 Using a BBN

The first thing we can do with a BBN is to compute joint and individual distributions. Switching to the histogram view, we first show the unconditional (marginal) distributions. The mean and standard deviations for each variable are shown beneath the histogram.
Suppose we want to set a training level equal to 50 hrs per year; or 150 hours

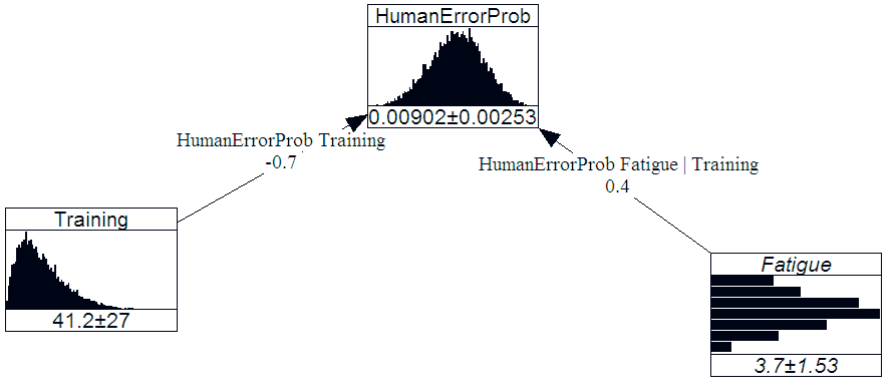


Figure 29: BBN in histogram view

in three years. How might that affect the human error probability distribution? If our training program does not change the system in any other way, our answer to this question is obtained by conditionalising on TR = 150. In other words, we look at the subset of measurements in which TR = 150 and construct distributions for other variables from this subset. Of course in the actual data this subset might be too sparse for this purpose. That is why we build a density function for our data; this allows us to smoothly approximate the distributions which we would find in these subsets.

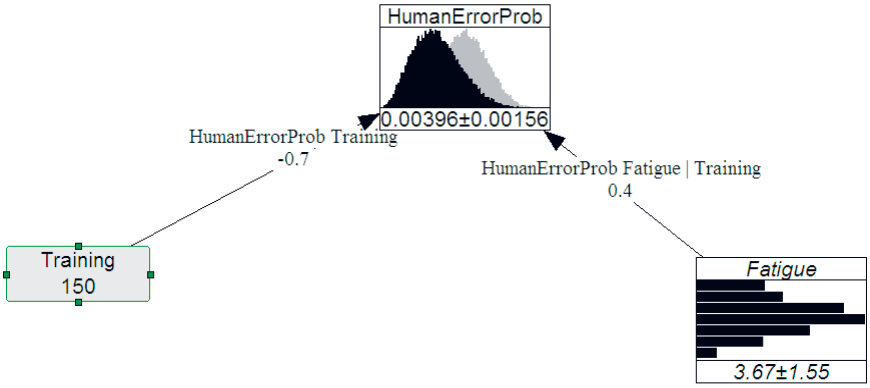


Figure 30: Approximate distribution

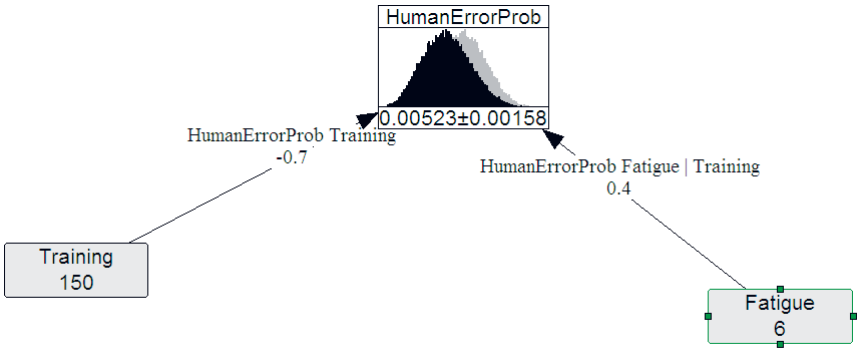


Figure 31: Fatigue increases Human Error Probability

The distribution of HEP is shifted to the left; the mean has dropped from 0.00902 to 0.00396. We might ask whether training enables operators to perform at higher fatigue levels.

We see that at fatigue level 6, the well trained operators still have a lower average human error probability than the overall average.

3.2.9 Fault Trees in BBN

Our purpose of modelling HEP is to capture probabilistic influence on basic events of a Fault Tree. Let us assume that our Top Event, System Failure, can occur in only two ways, either failure scenario 1 occurs, or failure scenario 2 occurs (or both). In both these scenarios, failure occurs if a subsystem fails AND human operator fails to recover. The subsystems are different, and the operators of the two systems may be different. The occurrence of human error in subsystem 1 is not related to human error in subsystem 2. However, the *probability* of error in both cases is influenced by training and fatigue. We assume that training and fatigue are the same for both subsystems. The combined system Fault Tree with HEP BBN may be represented in one BBN, as shown below.

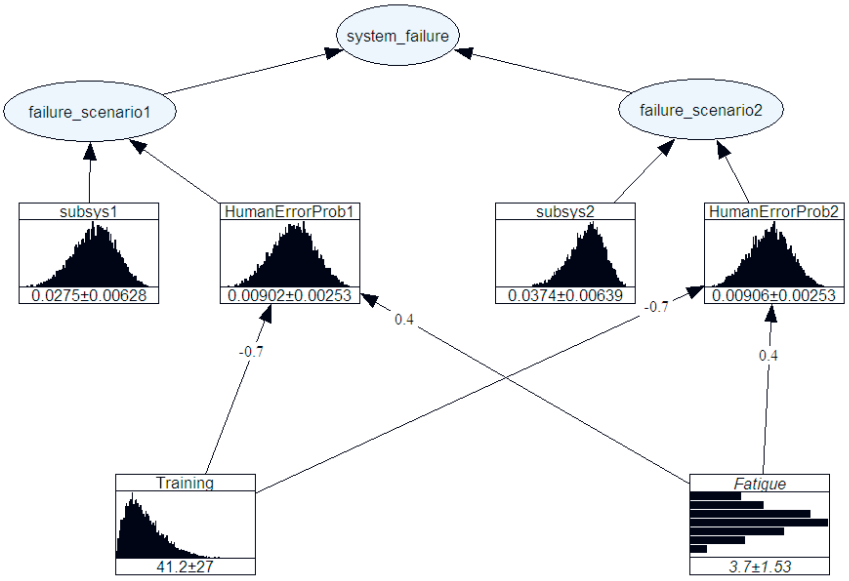


Figure 32: Fault Tree and Human Error Probability model combined in BBN

Failure_scenario1, failure+scenario2 and system failure are represented as functional nodes. Failure_scenario1 is a probability, which is the probability of subsys1 and HEP1, similarly for failure_scenario2. All these probabilities are random variables. The system failure is a function of the two failure scenarios. In a histogram view of this BBN, we can see the distributions of all variables.

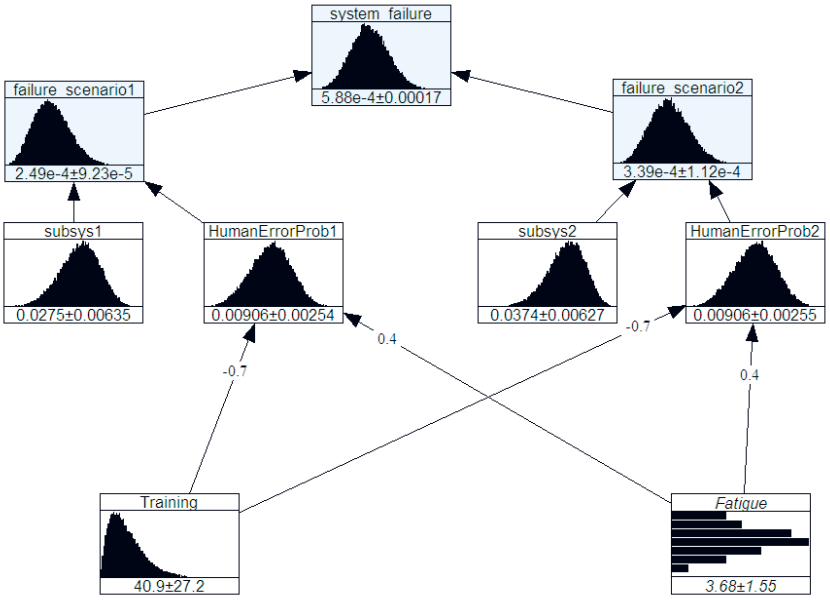


Figure 33: Histogram view of the BBN of Figure 32

A detailed distribution of the probability of system_failure is shown below. The median probability is 5.8148 E-4, with 95%-tile 8.7752E-4.

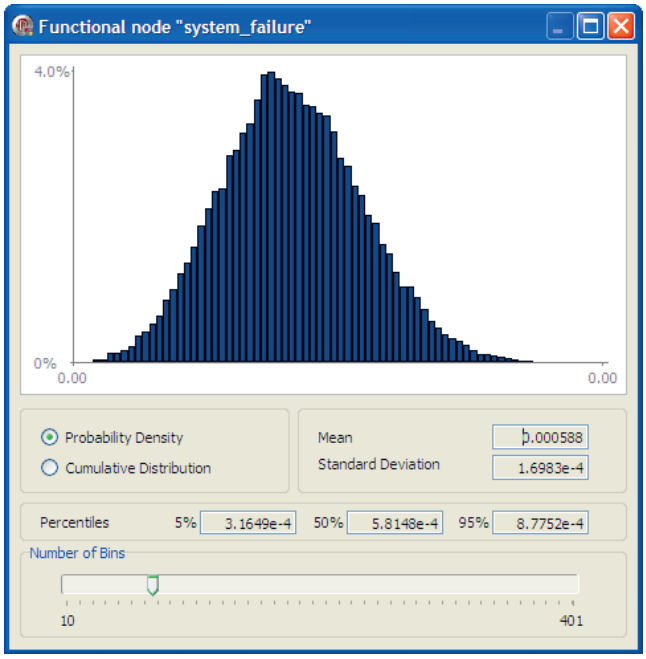


Figure 34: Histogram of functional node

The BBN for HEP has the effect of correlating the variables HEP1 and HEP2. In fact their correlation is 0.58. If we had not introduced the BBN for HEP, or equivalently, if we had made the influence of TR and FA on HEP equal to 0, then the distribution of the top event would change slightly. The result of this change is shown below:
Without dependence on TR and FA, the 95%-tile of system failure shifts from 8.7752E-4 to 8.3506E-4.

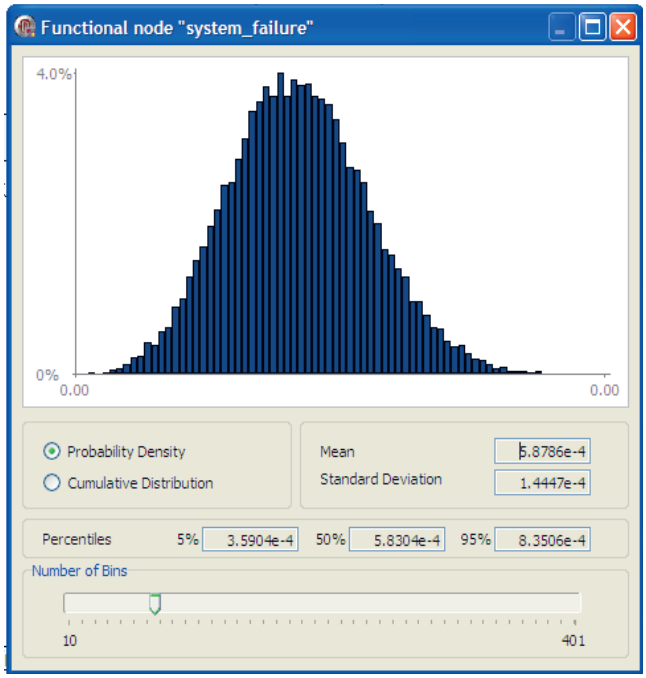


Figure 35: Histogram for HEP = 0

The effect of conditionalising TR at 150 hours can now be propagated through the Fault Tree up to the system failure. The mean probability for system failure shifts from 5.88E-4 to 2.75E-4.

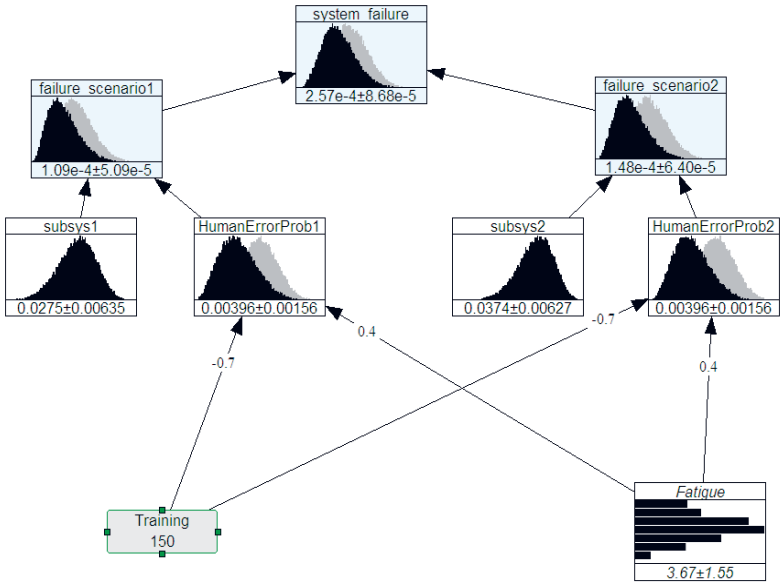


Figure 36: Influence of training propagated through the Fault Tree

3.2.10 Calibration and apportionment

In the model discussed above we had data on the HEP's, on Training, Fatigue and on subsys1 and subsys2. With this information we predict the distribution of System failure. In some cases, such as CATS, we have some data on System failure, typically this will be in the form of expected values or generic probabilities for System failure In such cases the modelling serves to apportion this generic probability over the underlying basic events and BBN nodes. The goal is to be able to predict how changes at a lower level will impact the probabilities at the higher levels.

In performing this apportionment, we may have some expected values at intermediate levels to guide us. The apportionment activity consists of:

- Developing expected values from data
- Assessing the degree of variability at those nodes, consistent with the expected values, at those nodes for which distributional information is lacking.
- Checking that with these variabilities, the expectations of the higher events are still calibrated on the data from step (1).

We saw in the example above that the dependence in basic events introduced by the BBN caused a slight shift in the distribution of System failure. If the distributions are highly skewed, or if the dependences are very strong, the induced shift might be enough to disturb the calibration. In this case the above steps must be iterated.

3.3 Using the BBN for accident analysis

The features of the BBN system described above can be exploited to investigate the effect of influences, to search for potential causes of accidents and to investigate the effects of changes in the system.

3.3.1 Finding potential accidents

Although up to this point the full power of the BBN modelling has not been described, the advantages and potential of this way of modelling can already be

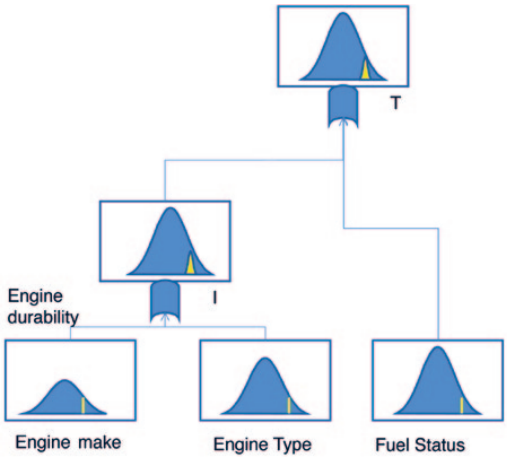


Figure 37: conditionalising

seen. Many studies struggle with the explanation of rare events or accidents that seem to be almost extraneous to the system at hand. The problem of explaining accidents and incidents in high reliability organisations has led to many metaphors such as the functional resonance metaphor by Hollnagel (2006).

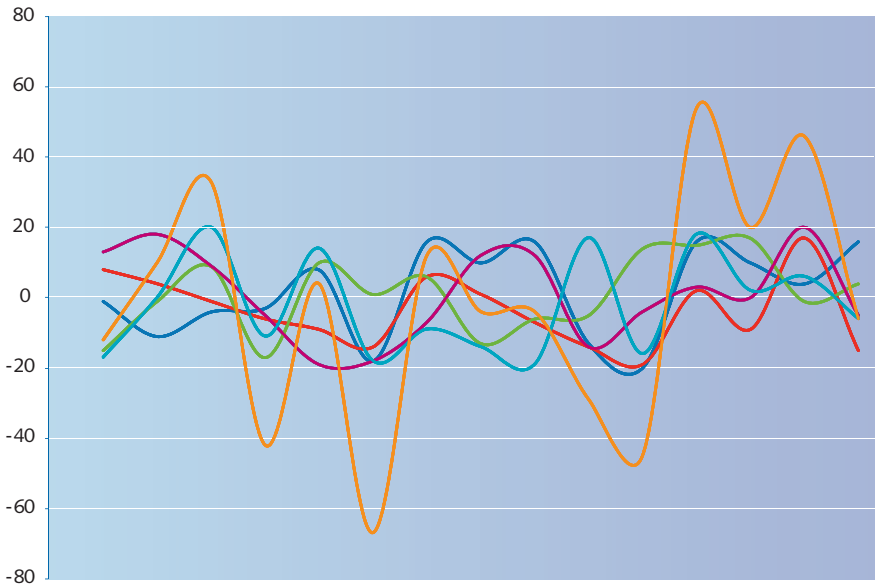


Figure 38: How random numbers combine

Using distributions rather than point estimates shows that accidents may just be the result of the normal variations of properties of the constituents of the system. If in a particular instance more extreme values of a number of parameters combine, the probability of an accident increases to an extreme value in the distribution of the accident probabilities, making an accident much more likely than an analysis based on central, point estimates would otherwise indicate (Figure 37).

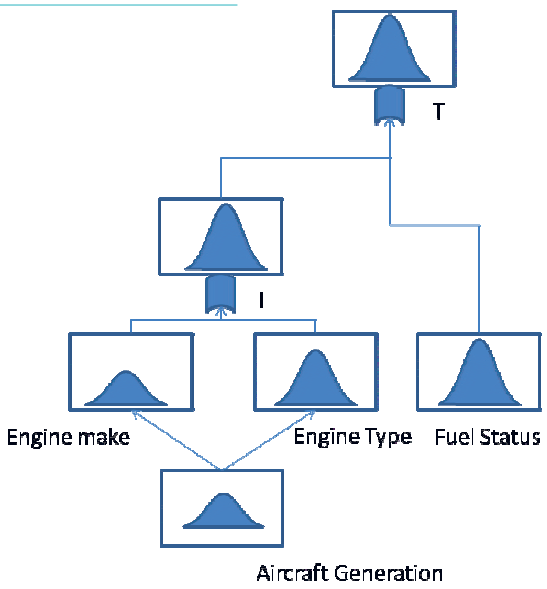


Figure 39: Interrelationship

That extreme values resulting from random combinations of extreme events is not as rare one might think is illustrated in Figure 38. In this figure the random number generator of Excel was used to create five series of values between -20 and 20. The orange line shows the sum. Although the average of each separate series is 0, within 20 "tries" extreme values of -67 and +54 already are reached.

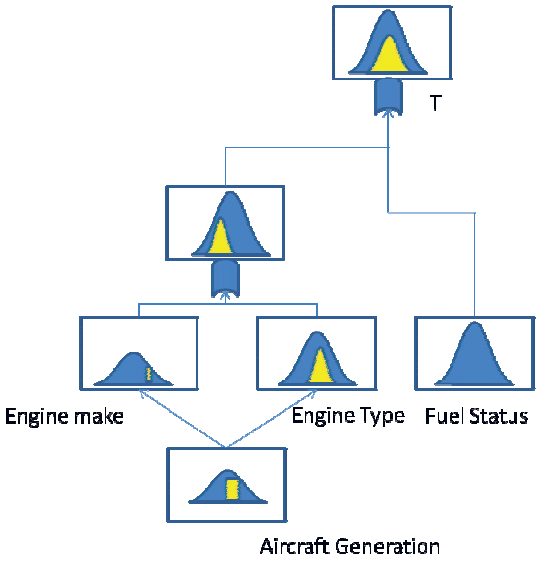


Figure 40: Conditionalising on interrelationship

This shows that such random variations may indeed give rise to exteme values fairly quickly, without these values influencing or resonating between each other. These accidents then prove indeed to be normal accidents (Perrow, 1994).

3.3.2 Interdepencies

The full power of the BBN technique can be exploited by modelling interdependencies rigorously. In (Figure 39) two of the base events Engine type and Engine make now are connected by a common parent node Aircraft generation. Given the structure given in this figure, there are only two ways in which the value of Engine type can be changed from its original distribution. These correspond to two different types of analysis that can be performed using the model.
The first is aimed at analysing why or by what cause a system differs from the average systems for which the model was quantified to begin with. The second is to investigate the result of a deliberate change in the system, which gives it a fundamentally different distribution.

3.3.3 Analysis of members of the population.

As an example, a type one analysis may be aimed at investigating why a certain airline has more incidents of a certain type than another, and whether that should be considered as serious. Suppose the base event Engine Type is the incident or deviation for which an observation is made (see Figure 40). Engine type can then be conditionalised on the observation. Since there is no other parent for Engine type than Aircraft Generation, Engine Type can only change if Aircraft Generation changes as well. This means that for observation Engine Type to be made, Aircraft Generation has to have had the corresponding value or distribution. This could for instance indicate that a high occurrence of a particular type of malfunction in a certain type of instrument is correlated with being made on a Monday – even although the majority of the Monday production is all right. Being made on Monday then can be considered the cause, or one of the causes. Since Aircraft Generation also affects Engine Make, Engine Make also has to be different from the population average. So for confir-matingthe finding of Engine Type and the relationship between Engine Type and Aircraft Generation, the value (distribution) of Engine Make can be observed. In any case the combined changes in Engine Type and Engine Make lead to changes in Engine Durability and Engine Failure.

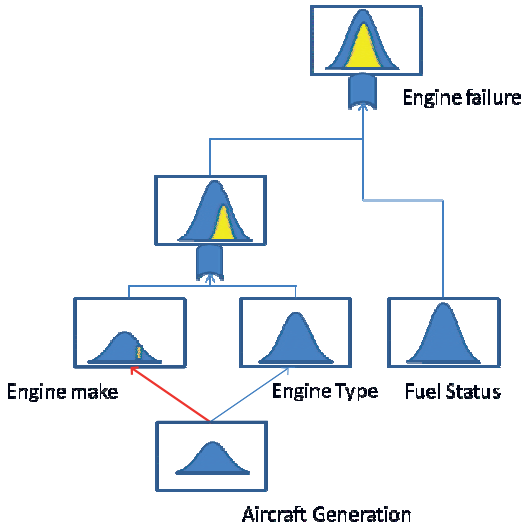


Figure 41: Known end distribution

Another use of this mode of analysis is to investigate what part of the system to address to effect a change in the value (distribution) of a node or event. Suppose one would want to change the value of Engine type, say improve the reliability of an engine. The model then first of all indicates that that can only be done by changing Aircraft generation (say replace the aircraft by another type). The model then also indicates that this will have the side effect of changing Engine make (say the lay-out of the cockpit) which may also change Engine durability. This could be advantageous, but it could also be an undesired side effect of the change made in Aircraft generation.

3.3.4 Creating a different system

In the second mode of analysis the relationships between the nodes in the model are deliberately changed. This could correspond to a situation in which a certain type of aircraft is always fitted with a certain type of engine. As was illustrated in the previous paragraph, a change in engine can only be achieved by changing the aircraft type as well. However, one could sever or change the relationship between Engine type and Aircraft generation, allowing a change in type of engine without a change in aircraft type (Figure 41). This type of

analysis has to be done with even more care than the other types of analyses, as there is no guarantee that the observations, data analyses and expert opinions on which the relationships and valuations in the model are based all continue to be valid, when this relationship is changed. Nevertheless this mode of analysis is a powerful tool in analysing the potential success and the potential side effects of system changes.

Most of the analysis however will be associated with finding the position of sub sets of systems in the whole of the population or at finding ways to change the current position into a better one. This would not make the relationships or the valuations for the whole of the population invalid and thus could be done more readily.

An alternative implementation of changes "outside" the current population would be to introduce "change nodes". The advantages and disadvantages of such a method is currently being investigated(Figure 42).

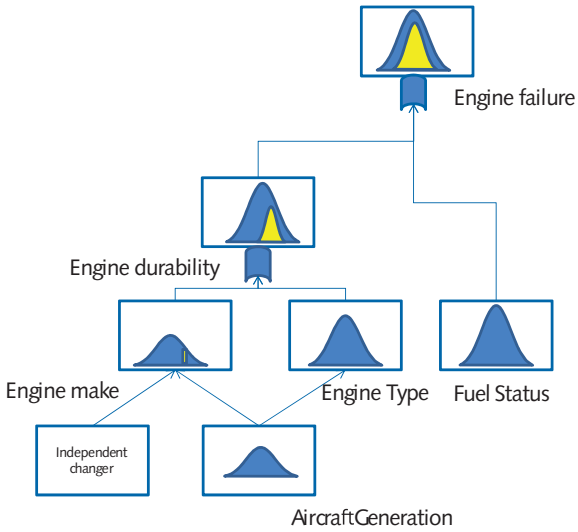


Figure 42: Unknown factor found

3.3.5 Unknown Unknowns

As described above some of the assumed unknowns are really just undetected results of internal dependencies. Therefore unexpected outcomes of model calculations should not be dismissed but investigated, as they probably indicate undetected accident pathways or causes giving an increased probability of failure. This does not take away the fact that many may be really unknown. How the results would change if these unknowns were known is not known either, and this problem cannot be solved immediately, only with time and further results.

This is where efficiency and effectiveness become parameters of the problem solving exercise. By further analyses more can be found out about the behaviour of the system when parameters of the system vary within the statistical boundaries. In a system such as CATS, this includes variations which would normally not be considered to be inside the design space. Such variations could be the fitting of wrong parts, failure to detect faults at maintenance, unspecified delay of repairs etc. These faults and deviations are part of reality and from statistical analyses of accidents and incidents the probability of occurrence can be determined or inferred. As such they become part of the model.

In turn this also would allow the analysis of the cost effectiveness of further reducing the probability of "beyond design" occurrences. Even after an accident or incident it could be systematically analysed whether spending resources on further reducing the probability of such an event is worthwhile. This could be a radical change with respect to current practices, where measures to prevent a re-occurrence are almost always taken, mainly because the probability of such a re-occurrence is rarely investigated.

4 Quantification

4.1 Requirement

The CATS model requires Fault Trees to quantify the causes of each initiating event and pivotal event in each of the 33 ESDs. The probabilities of the Fault Tree top events are equal to the probabilities of the ESD events (NLR, 2006). The Fault Trees show the breakdowns of causes of these events, to the extent that is possible within the limitations of Fault Tree Modelling. Where the causes result from human behaviour, the base events of the Fault Trees link to the BBNs of human performance. Other base events of the Fault Trees link to the user inputs of the CATS model.

In order to integrate with the BBNs of human behaviour, the Fault Tree Models are made to be implemented in a giant BBN. To achieve this, the uncertainties in the base events are described by probability distributions. The Fault Trees are sufficiently detailed and sufficiently robust to contribute towards the objectives of the overall model.

A number of data sources were used as a basis for quantification.

Airclaims and ADREP were the primary accident data sources. The time period considered was1990-2003. This period provided a dataset that is large enough for quantification and is considered representative for 'current' air transport. When only Airclaims was used the time period was slightly expanded to 1985 - 2005 to provide a larger data sample. Because of the size of most databases involved, much of the initial analysis was done by running queries, e.g. looking for particular key words. Each incident in the resulting dataset was then individually analysed to verify whether it 'fitted' the particular ESD under consideration. The quality of information in ADREP about accident causes is not sufficient to support the present analysis. Therefore original accident investigation reports have been used where available. In other cases, the summary information from Airclaims, Aviation Safety Network, Flight Safety Foundation and others explains the causes in sufficient detail to relate to the barrier model. Incident reports have also been used where available.

The primary source of data for quantification of the probability of occurrence of the initiating event are the databases of Service Difficulty Reports and Air Safety Reports but sometimes other sources of data were used if these were considered to be more accurate.

4.2 An Example ESD

The following sections use a single example (ESD12 - spatial disorientation) to illustrate the approach. The other Fault Trees use the same methodology and

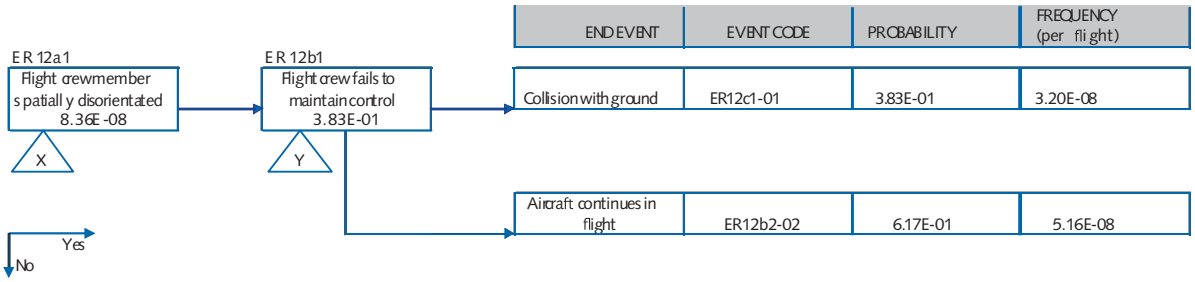


Figure 43: Quantified ESD for Spatial Disorientation

deliver results in the same format. Further details are provided in the overall DNV report on the Fault Tree modelling (Spouge, 2008).
Spatial disorientation refers to a scenario where the flight crew develop a mistaken perception of their position and motion. This may induce flight commands that place the aircraft in an extreme attitude, outside the normal flight envelope, i.e. unusual bank or pitch angles. Unless this attitude is corrected, it is virtually impossible for the flight crew to control the aircraft's trajectory, and so in effect control is lost at this point. Controlled flight into terrain (CFIT) is treated separately in ESD35.
The quantified ESD for spatial disorientation is shown in Figure 43. The labels X and Y highlight the initiating and pivotal events whose causes must be quantified through Fault Trees.

4.3 Barrier Model

The structure of the Fault Tree Model is based on a barrier model of each scenario. While the ESD shows physical events preceding the accident, the barrier model is able to show the logically necessary events for the accident to occur.
For example, the following are the major barriers against loss of control due to spatial disorientation, all of which must be unsuccessful if the accident is to occur:

- Autopilot control. Spatial disorientation occurs when the flight crew have manual control of the aircraft trajectory. Modern commercial aircraft have autopilots that can control the trajectory for almost the entire flight except take-off and touch-down. However, to maintain handling practice, pilots often retain manual control in climb to about 10,000ft, and in descent from about 3000ft or on making visual contact with the runway.
- Attitude guidance. In instrument flight rules (IFR), the flight instruments are the primary mechanism for maintaining spatial orientation. The key instrument is the attitude director indicator (ADI). Commercial aircraft are fitted with 3 ADIs, one for each pilot, and a spare that can be selected by either pilot.
- Visual orientation. In visual meteorological conditions (VMC), the pilot's ability to see the horizon assists in maintaining spatial orientation.
- Attitude monitoring. The pilot not flying (PNF) should monitor the actions of the pilot flying (PF) and challenge any incorrect attitude commands before the aircraft reaches an extreme attitude.
- Control recovery. If the aircraft reaches an extreme attitude, it remains possible for pilots to correct it, but this requires actions for which only military or aerobatic

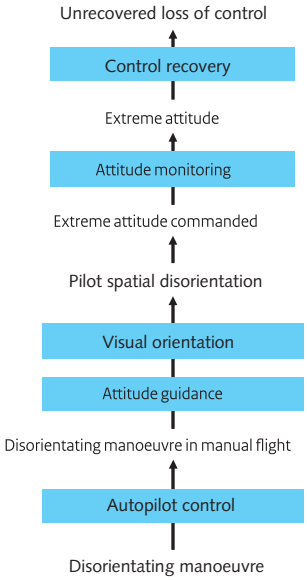


Figure 44: Barrier Model of Spatial Disorientation

pilots are trained, as there is a danger of inducing structural failure through incorrect commands.
Figure 44 shows the barrier model for spatial disorientation. It shows a sequence of accident precursors, which the different barriers attempt to prevent developing into the accident.

4.4 Causes of Barrier Failure

Each barrier may fail (i.e. be unsuccessful in preventing the scenario developing to the next precursor) for different reasons. The reasons for barrier failure are represented in the Fault Tree as the causes of the accident. Thus the model represents at least 5 causes of any accident due to spatial disorientation.
As an example, considering only the first barrier, the identified causes of lack of autopilot control are:

- Autopilot not capable of controlling the aircraft in the required manoeuvre. This typically occurs on older or smaller aircraft where the autopilot is only suitable for steady flight conditions.
- Autopilot not used by the flight crew. This may be due to:
  - Training (or maintaining familiarity) of the flight crew in manual flight.
  - Crew preference to use manual flight.
  - Crew lack of knowledge of how to use the autopilot to control the aircraft in the required manoeuvre.
- Autopilot incorrectly used by the flight crew. This is where the flight crew attempt to use the autopilot but fail to achieve control over the aircraft trajectory with it.

The model represents these as alternative causes of failure of the first barrier. The underlying human and organisational reasons for these events are not suitable for modelling in the Fault Trees. The development of causes in the Fault Tree therefore stops at the point where reasonably distinct independent events can be identified that are either necessary or sufficient to cause failure of a barrier.

4.5 Causal Data

Quantification of the Fault Tree Model uses distributions of causes obtained from accident and incident experience. The quality of information in the ADREP database about accident causes is not sufficient to support the present analysis. Therefore original accident investigation reports have been used where available. In other cases, the summary information from Airclaims, Aviation Safety Network, Flight Safety Foundation and others explains the causes in sufficient detail to relate to the barrier model. Incident reports have also been used where available. The term “event” is used below to refer to both accident and incidents.
To quantify the Fault Tree, it is not necessary to know the causes of every event that has occurred. Since the ESDs have been quantified using probability data, consisting of comprehensive counts of the numbers of events among known flight exposure, the causal breakdown in the Fault trees can be quantified from a representative sample of events. It is therefore assumed that the events whose causes are known, and which are used to quantify the causal breakdowns in the Fault Trees, are representative of the causes of the full set of accidents.
For ESDs with little or no accident experience, the Fault Trees are quantified using experience from precursor incidents. These are incidents that were

prevented from developing into the relevant accident by the success of one or more barriers. It is assumed that the causes of these incidents indicate the likely causes of initiating events in future accidents. The causes of the necessary further barrier failures can be obtained from other ESDs in which the barriers are relevant, or as a last resort from expert judgement about their relative likelihood. In general, the Fault Trees have been developed only to a level that can be quantified mainly from available accident or incident data. Pure judgements about event probabilities have been minimised.

Spatial disorientation events are difficult to identify from available databases such as ADREP. Therefore, events have been accumulated through a slow process of categorising loss of control events into the different ESDs.

Table 2: Example Spatial Disorientation Accident

Date	**23 Aug 00**	
Type	A320	
Operator	Gulf Air	
Location	Bahrain	
Flight phase	Missed approach	
	Description of failure	**Fault Tree event**
Visual orientation	Conditions were darkness with no moon but good visibility.	Darkness
Autoflight system	The aircraft was making a VOR/DME approach. At 1700ft the autopilot was disconnected when visual with the airfield.	Autopilot not capable.
Attitude guidance	The aircraft was fast on approach, and while at 600ft used a non-standard 360o turn to reduce speed, but then made a missed approach. During this, the captain became spatially disorientated, falsely perceiving the aircraft was pitching up. He commanded the aircraft into 15o nose-down pitch, while at only 1000ft altitude.	ADI not used.
Attitude monitoring	The first officer (PNF) had not alerted the captain to the non-standard elements of the approach, in contravention of operating procedures. When the nose-down pitch was commanded, the speed increased and the Master Warning indicated flap speed exceeded, which the first officer called out, but the captain did not respond to this.	Lack of monitoring.
Control recovery	GPWS pull-up warnings were received. The captain then responded only by raising the flaps.	Incorrect recovery action.
Consequences	The aircraft struck the sea at 280 knots.	

Ten spatial disorientation accidents were used for preliminary quantification, and no incidents were found. More recently, further events have been identified and could be used in future work. Significant uncertainty results from the use of such a small dataset, as considered further below.

Each event has been analysed to determine the cause of failure of each barrier. Table 2 shows this analysis for an example spatial disorientation accident. Each cause is assigned to the events identified for the Fault Tree. Analysis of sufficient events creates the required causal distributions.

4.6 Fault Tree Model

The Fault Tree provides a logical structure showing how causal factors could combine to cause an initiating or pivotal event of the ESD. The ESD shows how combinations of these events may cause an accident. Figure 45 shows the different types of logic gates used in the Fault Tree. They are explained in turn below.

AND gates are used where an event A has two independent, necessary causes B and C. The probability is:

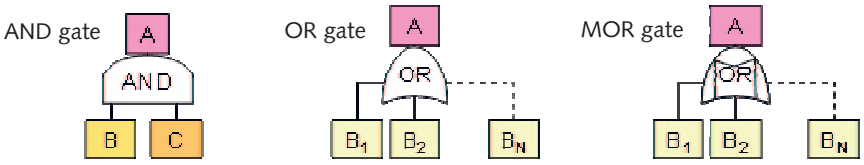


Figure 45: Schematic Fault Tree Logic Gates

$$P(A) = P(B) \times P(C)$$

OR gates are used where an event A may result from N alternative causes Bi. Assuming the causes are independent, the probability is:

$$P(A) = 1 - \prod_{i=1}^{N} (1 - P(B_i))$$

MOR gates are used where an event A may result from N alternative causes B_i that are mutually exclusive (i.e. only one can occur at once by definition). The probability is:

$$P(A) = \sum_{i=1}^{N} P(B_i)$$

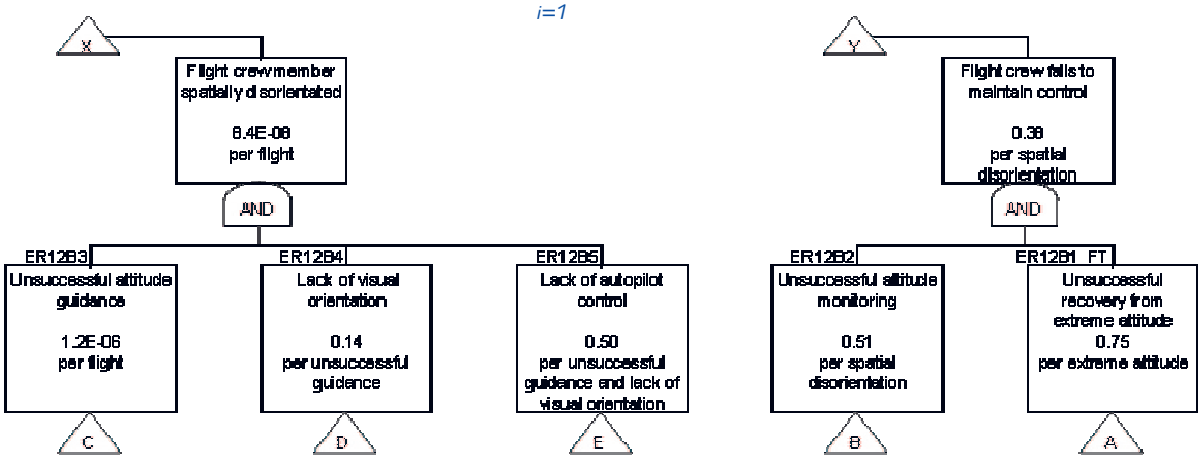


Figure 46: Top Events of Fault Tree for Spatial Disorientation

In developing the Fault Tree, a top-down approach is followed, which reverses these calculations. The top events of the Fault Trees are known from the initiating and pivotal events from the ESD. These are split into events corresponding to unsuccessful performance of each barrier. These unsuccessful barrier events are then further split into the causes of barrier failure. Each stage requires further information, which is obtained either from the causal distributions above, or from other data sources or judgements.

The same approach was used by EUROCONTROL in developing the Fault Tree Models for the Integrated Risk Picture (IRP) (Eurocontrol, 2006). By agreement with EUROCONTROL, the IRP models for collisions have been adopted from this source. Some changes have been necessary because the explicit modelling of common-cause events in the IRP is not required in CATS, since this aspect is represented by BBNs (see below).

Example sections of the Fault Tree for spatial disorientation are given in Figure 46 and Figure 47. For each event, the tree shows the failure probability per demand. In general, the events are conditional on occurrence of events to the left of them in the tree. For the events on the extreme left side, the relevant demand is a flight.

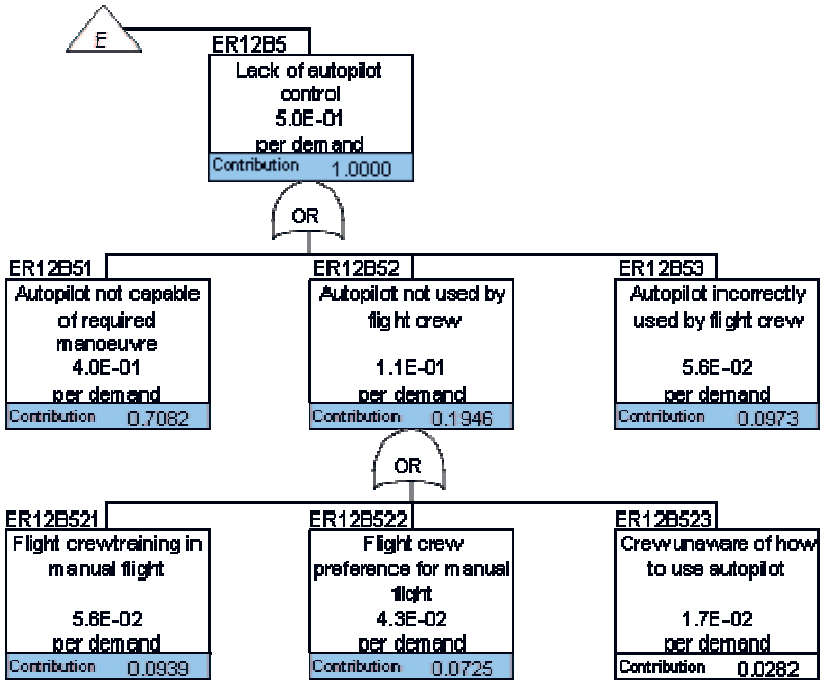


Figure 47: Example Base Events of Fault Tree for Spatial Disorientation

4.7 Event Contributions

The Fault Tree Model includes the "contribution" of each causal factor, which gives a simple indication of its relative importance to the accident frequency for that ESD. The contribution is calculated following a top-down approach, beginning from the top event for each pivotal event of the ESD, which is given a contribution of 1.

At an AND gate, the contribution of input events B and C is taken as the same as the output event A. This is the same as in the Fussell-Vesely importance measure, reflecting the fact that the output changes in direct proportion to changes in either input.

At an OR gate, the contribution of the output event A is divided among the input events B_i as follows:

$$C(Bi) = \frac{C(A) \times P(B_i)}{\sum_{j=1}^{N} P(B_j)}$$

where:

$C(B_i)$ = contribution of input event B_i

$C(A)$ = contribution of output event A

$P(B_i)$ = probability of input event B_i

This is the Differential Importance Measure (NASA, 2002) based on a uniform change for all inputs, and has the property of being additive within each OR gate, i.e.:

$$C(A) = \sum_{i=1}^{N} P(B_i)$$

The contribution gives a simple estimate of the maximum benefit, expressed as a fraction of the accident frequency that could be achieved by improve-

ments in each specific factor. A contribution of 0.1 implies that the accident risk would be reduced by 10% if the causal factor could be prevented. Due to the non-linearity of Fault Tree (notably the large probabilities), any such single measure of causal contribution is only an approximation. More accurate results could be obtained where necessary through comprehensive sensitivity testing, which is appropriate as part of the giant BBN model.

Figure 48 shows the contributions from the base events of the spatial disorientation Fault Tree. For clarity, they are expressed as contributions to the spatial disorientation accident frequency, as shown in the Fault Tree. The 90% confidence ranges shown on these results are based on epistemic uncertainties, as explained below.

Such results provide potentially useful information about the relative importance of different causal factors. For example, the results above suggest that pilot failure to use the ADI is a more common cause of spatial disorientation than failure of the ADI. It also shows that most spatial disorientation accidents involve disorientating manoeuvres (e.g. turns during missed approach) in IMC under manual flight control. This shows areas in which safety improvements could be concentrated. These are also areas in which the model could be made more detailed in future work.

4.8 Case-Specific Modifications

The Fault Tree Model represents a generic average of commercial aircraft operations. It is based on causal breakdowns drawn mainly from experience during the period 1990-2006, which provides sufficient accidents and incidents to quantify the model while also being reasonably consistent with modern operational practice. It is consistent with the data choices that were made in quantifying the ESDs. In particular, the ESD for CFIT represents only aircraft with a terrain awareness and warning system (TAWS), since this has been required for commercial aircraft since 2007. In other respects, the Fault Tree Model represents an average of commercial experience during earlier periods, and is assumed applicable to current operations.
The Fault Tree Model is able to represent specific cases that differ from the generic average. These cases are defined by user inputs, which select from the possible states of various influences on the base events of the Fault Tree. Ideally these should be continuous variables, but in practice suitable metrics and data are usually not available. Therefore most of the influences are defined as sets of discrete states. Each state has an exposure probability, defined as the proportion of world-wide flights of commercial aircraft that experience the influence in that state. The generic case represented in the Fault Tree is the average of these specific statesThe relationship (or mapping) between an influence and a Fault Tree base event is expressed as a modification factor (MF), defined as:

$$MF = \frac{\text{Base event probability in specific state}}{\text{Base event probability in generic state}}$$

The effects of these mappings on the overall risks are expressed as risk ratios (RR), defined as:

$$RR = \frac{\text{Accident frequency in specific state}}{\text{Accident frequency in generic state}}$$

MF and RR represent model inputs and outputs respectively. The difference between them arises from the non-linearity of the Fault Tree Model.

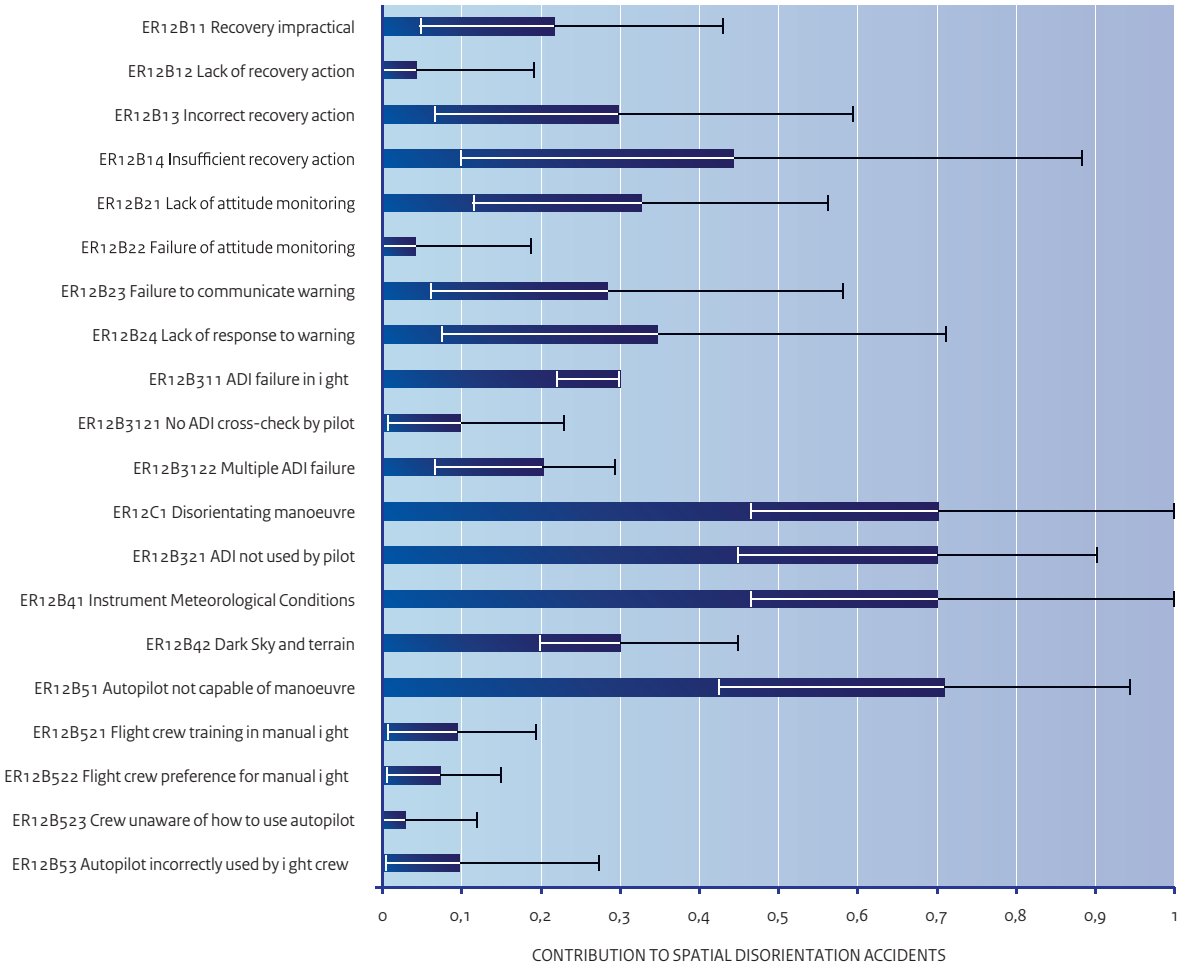


Figure 48: Contributions from Base Events of Fault Tree for Spatial Disorientation

When quantifying the MFs, the following types of influences are distinguished:

- Deterministic influences. These are where the Fault Tree event is defined so that it can only occur in one of two possible input states (e.g. wind-shear present or absent), so that the MF is determined directly by the exposure probability of the chosen state.
- Data-based influences, i.e. probabilistic influences quantified using stratified data. These are where the Fault Tree event may occur in any of a set of alternative states (e.g. aircraft generation 1, 2, 3 or 4), whose probabilities of occurrence and MFs can be obtained using accident and exposure data.
- Judged influences, i.e. probabilistic influences quantified using judgement. These are where the Fault Tree event may occur in any of a set of alternative states, whose relative probabilities are based on judgement in the absence of any useful data.
- Functional influences i.e. probabilistic influences expressed as an analytical function of a user input (e.g. airport elevation). The function may be a

judgement that the Fault Tree event is proportional to the input, or it may be a fit to data-based influences.
Table 3 lists the influences that are represented in the Fault Tree Model.

Table 3: List of Fault Tree Influences

GROUP	INFLUENCE	STATES/METRIC	TYPE
Operating environment	Geographical region Traffic level TMA complexity	ICAO region Fraction of 1990-2005 average ATC vectoring commands per flight	Data Function Function
Flight operation	Flight phase Operation type	Taxi, take-off, climb, en-route, approach/landing Passenger, cargo, non-revenue	Data Data
Aircraft	Aircraft propulsion Aircraft size Aircraft generation GPWS type ACAS PWS Autoflight use	Jet, turboprop kg MTOW 1, 2, 3, 4 None, early, standard, TAWS Installed, not installed Installed, not installed Fraction of trajectory changes via FMS	Data Data Data Data Deterministic Deterministic Function
Airport	Airport elevation Approach type Runway length Runway crossing Runway condition Runway slipperiness measurement Runway slipperiness criteria Runway maintenance criteria FOD criteria Bird management	ft above mean sea level Precision, non-precision Short, medium, long Runway crossings per flight Wet, dry Frequency Used, not used Used, not used Used, not used Used, not used	Function Data Judged Function Judged Judged Judged Judged Judged Judged
ANSP	LLWAS STCA Ground radar RIMCAS Terminal area radar MSAW	Installed, not installed Installed, not installed Installed, not installed Installed, not installed Installed, not installed Installed, not installed	Deterministic Deterministic Deterministic Deterministic Deterministic Deterministic
Ambient environment	Light condition Visibility at airport Visibility in flight Cross-wind Wind-shear Turbulence Icing at airport Precipitation at airport	Daylight, dark Restricted, unrestricted IMC, VMC Strong, weak Present, absent Strong, weak Freezing, above freezing None, light, moderate, heavy	Judged Judged Judged Deterministic Deterministic Deterministic Deterministic Data

Figure 49 illustrates data-based influences using the effect of operation type. The number of fatal accidents in each operation has been obtained from the ADREP database for Western commercial aircraft during 1990-2006. After dividing by the estimated number of flights in each operation, the results are expressed as a risk ratio relative to average. Fatal accidents have been selected to avoid bias due to uneven reporting of non-fatal accidents. The example shows that risks on passenger operations are significantly lower than average, while in cargo and non-revenue operations they are significantly higher than average. Modification factors are based on the necessary adjustments of the base events in the model to produce the risk ratios that are shown in the data. In the example, it is further assumed that operation type affects all base events influenced by flight crew and maintenance performance (see below).

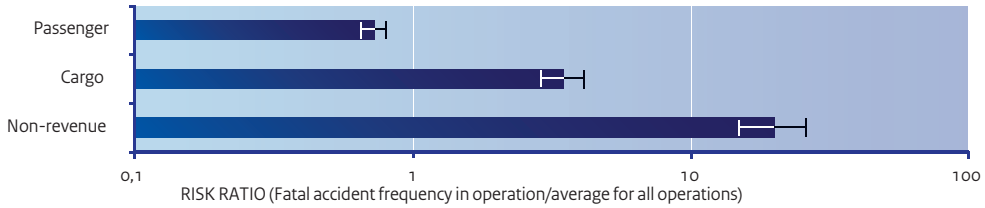


Figure 49: Example Risk Ratios from Accident Data

At present, each of these influences is considered as if its effects were independent of all other influences. In reality, many of the influences are correlated. In principle, these correlations could be modelled using BBNs, but developing suitable models is challenging, and will be considered in future work. Meanwhile, it must be recognised that combination of the MFs from different influences may be unrealistic, and may tend to over-estimate the effects (i.e. produce risk ratios that diverge excessively from 1).

Table 4 Example Base Event Influences for Spatial Disorientation

CODE	EVENT NAME	FLIGHT CREW PERFORMANCE	ATCO PERFORMANCE	MAINTENANCE PERFORMANCE	OTHER INFLUENCES
ER12B11	Recovery impractical				
ER12B12	Lack of recovery action	x			
ER12B13	Incorrect recovery action	x			
ER12B14	Insufficient recovery action	x			
ER12B21	Lack of attitude monitoring	x			
ER12B22	Failure of attitude monitoring	x			
ER12B23	Failure to communicate warning	x			
ER12B24	Lack of response to warning	x			
ER12B311	ADI failure in flight			x	
ER12B3121	No ADI cross-check by pilot	x			
ER12B3122	Multiple ADI failure			x	
ER12C1	Disorientating manoeuvre				Airport quality
ER12B321	ADI not used by pilot	x			
ER12B41	Instrument Meteorological Conditions				Visibility in flight
ER12B42	Dark Sky and terrain				Light condition
ER12B51	Autopilot not capable of manoeuvre			x	Autoflight use
ER12B521	Flight crew training in manual flight				
ER12B522	Flight crew preference for manual flight				
ER12B523	Crew unaware of how to use autopilot	x			
ER12B53	Autopilot incorrectly used by flight crew	x			

Other influences (notably the standard of performance of flight crew, ATC and maintenance personnel) are quantified by human performance BBNs elsewhere in the CATS model. The influences of these parameters are represented by conditioning the BBN to particular values within the distributions. Their effect on the Fault Tree is assumed to be the same for each base event that has a logical connection to the human performance. Example connections for base events from part of the spatial disorientation Fault Tree are shown in Table 4.

4.9 Uncertainties

The Fault Trees present best-estimates of the average probabilities of events among commercial aircraft operations. The following types of uncertainty can be distinguished in the results:

- Variability (also known as aleatory or Type A uncertainty). This is due to natural randomness. Due to many influences, some flights experience a higher probability of accidents than others.
- Epistemic uncertainty (also known as Type B uncertainty). This is due to lack of knowledge. It is impossible to know exactly what the probability of an event is, although this uncertainty can be reduced by more data collection or better

modelling. Epistemic uncertainties include:

- Model uncertainty (also known as structural uncertainty). This is due to simplifications or lack of realism in the formulation of the model. This is very difficult to quantify, unless by comparing independently produced models.
- Sampling uncertainty (also known as parametric uncertainty). This includes uncertainties due to:
  - Data quantity. This arises from the fact that relatively small datasets are available (and sometimes no accident experience at all). Standard mathematical techniques are available to quantify this type of uncertainty.
  - Data representativeness (or bias). This arises if the selected data does not match the problem of interest, e.g. it may be old or based on a few countries that investigate accidents thoroughly. Once these biases are understood, corrections can be made to minimise their effects.
  - Data interpretation. This arises because the accidents and incidents may be not fully understood or not clearly linked to the model. This is again very difficult to quantify, unless by independent evaluations of the available data.

In order to convert the Fault Trees into BBNs, it is necessary to define the complete uncertainty distribution (including variability) for each base event. This is expressed as the probability distribution for MF, as defined above, which has a mean value of 1 by definition.

The probability distribution of variability is quantified by combining the influences in the table above for each base event, assuming they are independent.

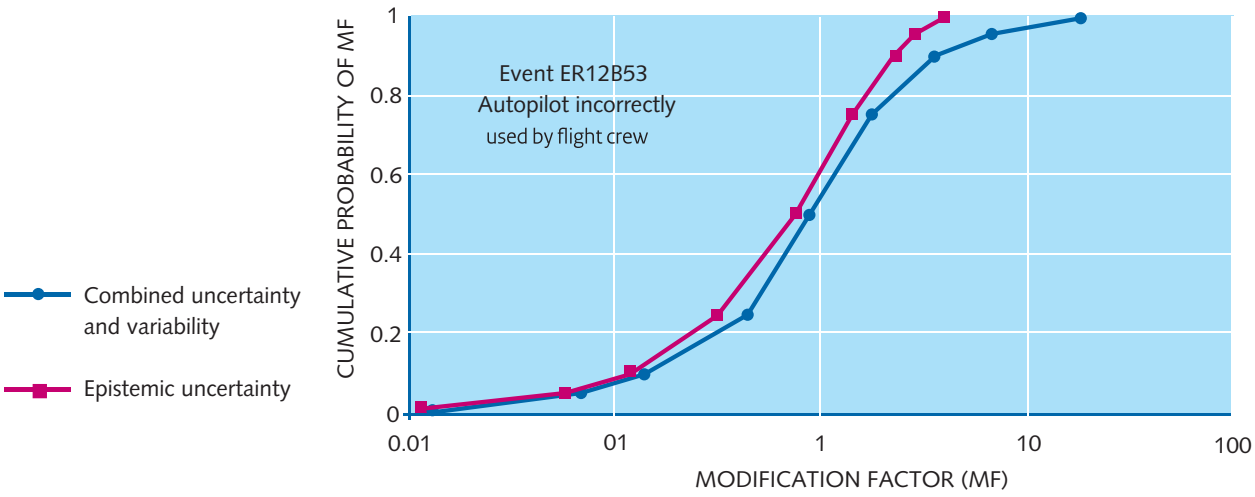


Figure 50: Uncertainty Distributions for Example Event

Comprehensive analysis of all sources of epistemic uncertainty would be resource intensive and itself extremely uncertain. In order to obtain probability distributions for the Fault Tree base events, a simplified approach is adopted, choosing the largest confidence range from the following sources:

- If the event probability is based on data, the distribution is calculated directly from the data quantity. On-demand probabilities are represented by a Beta distribution; per-flight frequencies are represented by a chi-square distribution.
- If estimates of the probabilities are available from alternative sources or using alternative judgements, the largest and smallest of these alternatives are used to define the extremes of a triangular distribution, with the best-estimate (i.e. MF=1) used as the modal value.
- In the absence of any data or alternative approaches, the uncertainty range is defined by judgement. The distribution is assumed to be lognormal if the log uncertainty range is symmetric, or triangular otherwise.

The chosen epistemic uncertainty distribution is then combined with the variability distribution, to obtain the complete uncertainty distribution for each base event. Figure 50 shows the uncertainty distributions for an example event. This event is based on only one accident, and so the epistemic uncertainty is large, with a 90% range covering nearly two orders of magnitude. The variability between flights is also large, and due to its skew it dominates the combined distribution for large MFs, even though it is truncated at a value of 18, when the event probability reaches its maximum of 1.
The effects of user inputs on these distributions can be approximated by skewing the distributions until the mean is increased by the required MF.

4.10 Dependencies

Dependent events are defined (Mosleh et al 1997) as pairs of events A and B, where the combined probability P(A and B) ≠ P(A) P(B). This may be because the events are functionally linked, or because they are associated through some extrinsic influence such as human interaction or the environment.
Dependencies are important for systems with multiple barriers such as aviation, because the overall accident probability may be very sensitive to the degree of dependency between the barriers. If this is inadequately represented in the model, substantial errors may occur in the results. The top-down approach to quantification used in the CATS Fault Trees ensures that the overall probabilities are consistent with actual accident data, but the effects of dependencies could in principle lead to errors in the predicted effects of interventions.
Fault Tree Models in general attempt to represent functional dependencies, to the maximum practical extent, in order to obtain base events that are as independent as possible. Common cause failures (CCFs) represent the residual dependencies between base events that are not explicitly modelled in the Fault Tree structure. This approach is used in Eurocontrol's IRP.
In the CATS model, the BBNs of human behaviour represent the main dependencies between base events. The implementation of the Fault Trees in the giant BBN takes account of the majority of these dependencies without requiring an additional CCF model. It is acknowledged that some errors remain, which could only be eliminated through a full probabilistic model, which would require substantial additional data and elicitation of correlations.
Another type of dependency occurs when an aircraft enters an ESD in a degraded state following a non-catastrophic event in a previous ESD. This linkage between ESDs is not modelled at present.
A full account of the dependencies modelled can be found in appendix NLR 2008-309

4.11 Validation

The Fault Tree Model is implemented in a spreadsheet, and it is appropriate to validate the results in some detail.
One key requirement is to verify that the Fault Tree has been constructed correctly. The first line of defence against errors in Fault Tree construction is DNV's quality assurance (QA) system. This requires:

- Definition of responsibility for each part of the work. This is defined in the report on each accident category and in the documentation within the Fault Tree package, which is to be included in CATSPAWS.

- Detailed self-checks by the responsible person at each stage of the work. This is an essential part of DNV's competence training for risk analysts. Some of the available checks are explained below.
- Independent review of each part of the work. The extent of this review depends on the competence of the responsible person, as judged from previous reviews. The identity of the reviewer is documented in the report on each accident category.
- Full documentation of the work through project reports.. Summaries are included in the Fault Tree package, which are copied into the parameter database CATSPAWS. (Spouge, 2008)

No QA system can guarantee that there are no errors in a model as complex as the CATS Fault Trees. Nevertheless, a high degree of error correction is achieved, through use of the following self-checks:

- Each Fault Tree Model is implemented in a gate-by-gate form, showing all intermediate results, which are also included in the project reports. This allows manual checks of each stage in the model, and this has been effective in identifying errors in the model.
- Each Fault Tree Model is implemented twice, quantified once from the top down (developing base event probabilities), and once from the bottom up (recalculating the top event probabilities). The fact that this returns numerically identical probabilities helps trap a high proportion of errors in model construction.
- The contributions of causal factors for each barrier in the Fault Tree sum to 1. This allows a simple check against numerical errors in these results. The contribution of each cause of barrier failure also allows a check against input data and subjective expectation.
- The Fault Trees have been implemented independently as BBNs, which provides a further verification that the calculations are consistent with the chosen logic gates.

In addition to checks of the Fault Trees themselves, there has been some cross-checking of the top event probabilities against the corresponding events in the ESDs, in the cases where the Fault Trees and ESDs were quantified independently. Quantification of epistemic uncertainty has also motivated consideration of alternative sources of probability estimates, which is believed to increase the quality of the result.

5 Human performance models

The human performance models were developed in BBNs from the start. As was described in chapter 20, there are three of these models currently is CATS, one for FLIGHT CREW, one for ATCOs and one for MAINTENANCE. The development of the crew model is described here in more detail to give an impression of how these models were developed. The development of the other two models has been done along similar lines and the reader is referred to the technical appendices NLR3, NLR4, NLR11 and NLR12 for more details.

5.1 FLIGHT CREW model

The human operator plays an essential role at the execution level of any risk bearing activity. In order to account for the influence of the human operator on accident causation, that role must be properly represented in the causal risk model.
The objective of this part of the study is to develop a quantified model for flight crew performance for CATS. The purpose of the human performance model is to quantify the probability of human failure in certain events in the ESDs.
It is proposed here to represent the flight crew performance model as a Bayesian Belief Net, with flight crew error as the child node and performance shaping factors as the parent nodes.

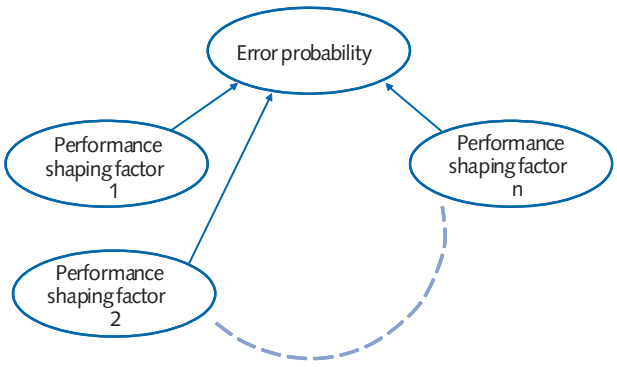


Figure 51: Human performance model schematic

By representing the performance shaping factors in a BBN, we are not limited by the assumption that the Performance Shaping Factors are independent.
If necessary, dependencies between performance shaping factors are easily introduced. We propose not to let the specific task determine the (initial) error probability, but to take the associated event in the event sequence diagram or Fault Tree as the starting point.

5.2 Linking with other parts of the CATS model

The flight crew performance model is developed to connect to the Fault Trees and event sequence diagram of the CATS model. Because the output of the CATS model is an accident probability per flight, the flight crew performance model must provide a flight crew error probability per flight. In several instances there may be a requirement for a further division into flight crew error per flight phase.

The flight crew performance model must be able to represent managerial and organisational influences on human performance. In this model, safety management is described as ensuring the provision, use, monitoring and maintenance of risk control measures. Proper provision and maintenance of the risk control measures, and monitoring whether they are used is managed by providing the necessary resources and criteria for each of those generic tasks. Resources and criteria are generically described by the following delivery systems: Procedures, Availability, Competence, Communication, Commitment (see chapter 6).

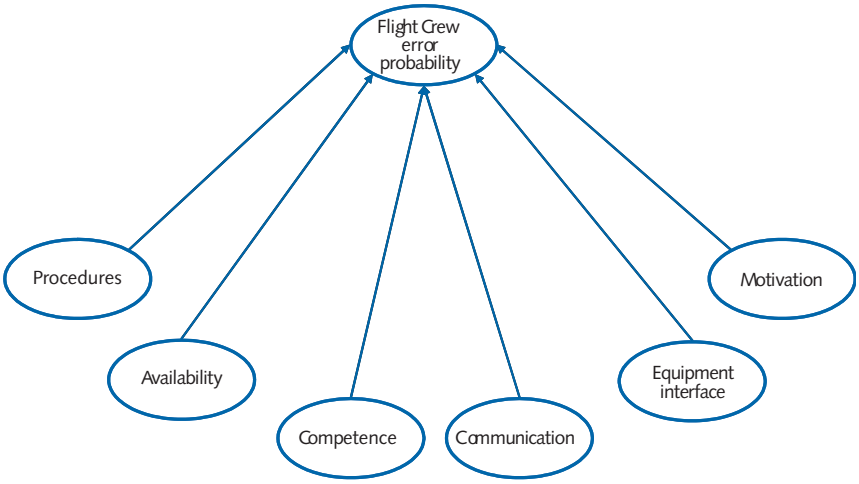


Figure 52: Linking managerial influences to flight crew error probability

In order to represent managerial influences on flight crew error probability these are described in terms of the delivery systems (see paragraph 6.2)
A basic description of the mathematics of BBNs is given in section 3.2. A number of influences have to be quantified on the basis of expert judgement.
A consequence of the use of expert judgement is that the variables under consideration must be expressed in operational terms and objectively quantifiable units. As stated before, operational definition means a set of rules or procedures that must be followed (and may be replicated by different observers) in order to determine the presence and quantity of something such as a variable, a term or an object. "Objectively" denotes that a particular value of that unit has the same meaning for expert A and expert B. This requirement demands a selection and definition of the variables that are included in the model.
As an example, consider the variable "safety culture". This is an influence that currently is often stated to be of prime importance for the safety of an operation. However: there is no operational definition of safety culture available. Then every expert will interpret it differently and results from different experts cannot be combined. The complete modelling is described in (NLR, 2005).

5.2.1 Model variables

Factors – or variables – that are considered to have a significant influence on the human error probability in air transport safety are considered here. Performance shaping factors have been selected after a review of literature Rasmussen, 1982; Speyer and Fort, 1982; Stein and Rosenberg, 1983; Swain and Guttman, 1983; Kirwan, 1994; Degani and Wiener, 1994; Hancock et al, 1995; Arbuckle, 1998; FSF, 1996, 1999, 2001; Roelen and Wever, 2002; Roelen, 2005) and preliminary analysis of a large sample of accidents and incidents.

The following list of performance shaping factors were initially selected:
experience
training
fatigue
crew composition
time pressure
workload
communication
time of day
weather
safety culture
man-machine interface
procedures
pilot attitude.

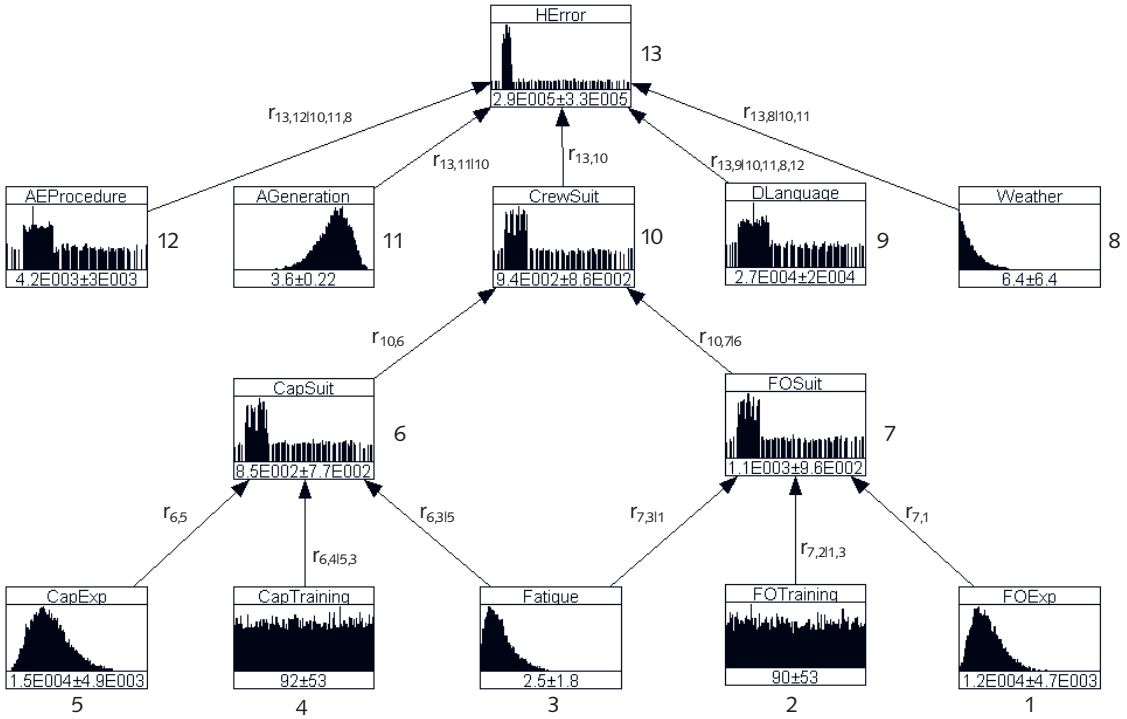


Figure 53: Flight crew performance model quantified

It was concluded that "time of day" and "fatigue" are interrelated. "Quality of procedures" is difficult to define and quantify properly. "Pilot attitude" is considered to be closely related to safety culture.
These three factors (time of day, procedures, and pilot attitude) are therefore not taken into account further. The remaining factors are accounted for – one way or another – and are discussed and defined below.

5.2.2 Model structure

Each variable described in the previous section is represented as a node in the model. In addition, nodes have been included for "captain suitability", "first officer suitability", and "crew suitability". Captain and first officer suitability are defined as the probability of successfully passing a proficiency check, if this check would be conducted at the moment of interest. Crew suitability is then defined as the probability that both the captain and the first officer are "suitable".

5.2.3 Model quantification

Quantification of the model concerns quantification of the marginal distributions of the nodes in the model as well as of the dependencies between the nodes. Table 5 lists the nodes, their definition and how the marginal distribution will be derived, either based on data or on expert judgement. All relations between nodes are quantified based on expert judgment.

Table 5: Flight crew model variables

Index	Name	Description	Source
1	FOExp	Total number of hours flown since the pilot's license obtaining by first officers.	Data
2	FOTraining	Number of days passed since last recurrence training for First Officers.	Data
3	Fatigue	Stanford Sleepiness Scale. 1signifies "feeling active and vital; wide awake" and 7 stands for "almost in reverie; sleep onset soon; struggle to remain awake".	Data
4	CapTraining	Number of days passed since last recurrence training for Captains.	Data
5	CapExp	Total number of hours flown since the pilot's license obtaining by captains.	Data
6	CapSuit	Number of Captains failing their proficiency check test per 10,000	Expert Judgment
7	FOSuit	Number of First Officers failing their proficiency check test per 10,000	Expert Judgment
8	Weather	Rainfall rate (mm/hr)	Data
9	DLanguage	Number of flights in which the pilot and first officer will have a different mother tongue per 100,000	Expert Judgment
10	CrewSuit	Number of Captains or/and First Officers failing their proficiency check test per 10,000	Expert Judgment
11	AGeneration	Aircraft generation is a scale from 1 to 4 where 4 is the most recent generation of aircrafts	Data
12	AEProcedure	Number of times the crew members have to refer to the abnormal/ emergency procedures section of the aircraft operation manual during flight per 100,000 flights	Expert Judgment
13	HError	Number of un recovered errors during any flight phase that might lead to a hazardous situation during the flight per 100,000	Expert Judgment

Quantification of the marginals will be discussed for each node separately. This again is fully described in (EWI, 2007).
The factors mentioned can rarely be expressed in quantifiable terms directly so the performance shaping factors are approximated by proxy quantities. This means that the value of a particular quantifiable entity takes the place of the original, but not quantifiable, entity. Weather is represented by rainfall and level of training by the time since last recurrence training. The rationale behind all these choices is elaborated in (Roelen, 2008).

Weather

The weather situation is characterised by the rainfall rate in mm/hr.

Type recurrent training

The marginal probability distribution of the number of days since last training is a function of the number of days between two subsequent trainings. If this number is n days, then the probability that m days have passed since the last training simply is 1/n for any $0 \leq m < n$.

Aircraft generation

The fleet is divided into 4 generations; generation 1 being the oldest jets and turboprops and generation 4 the most modern aircraft with digital systems and fly-by-wire control, such as the Boeing 777 and Airbus A-380.

Fatigue

Fatigue of the pilot is measured using the Stanford Sleepiness Scale.

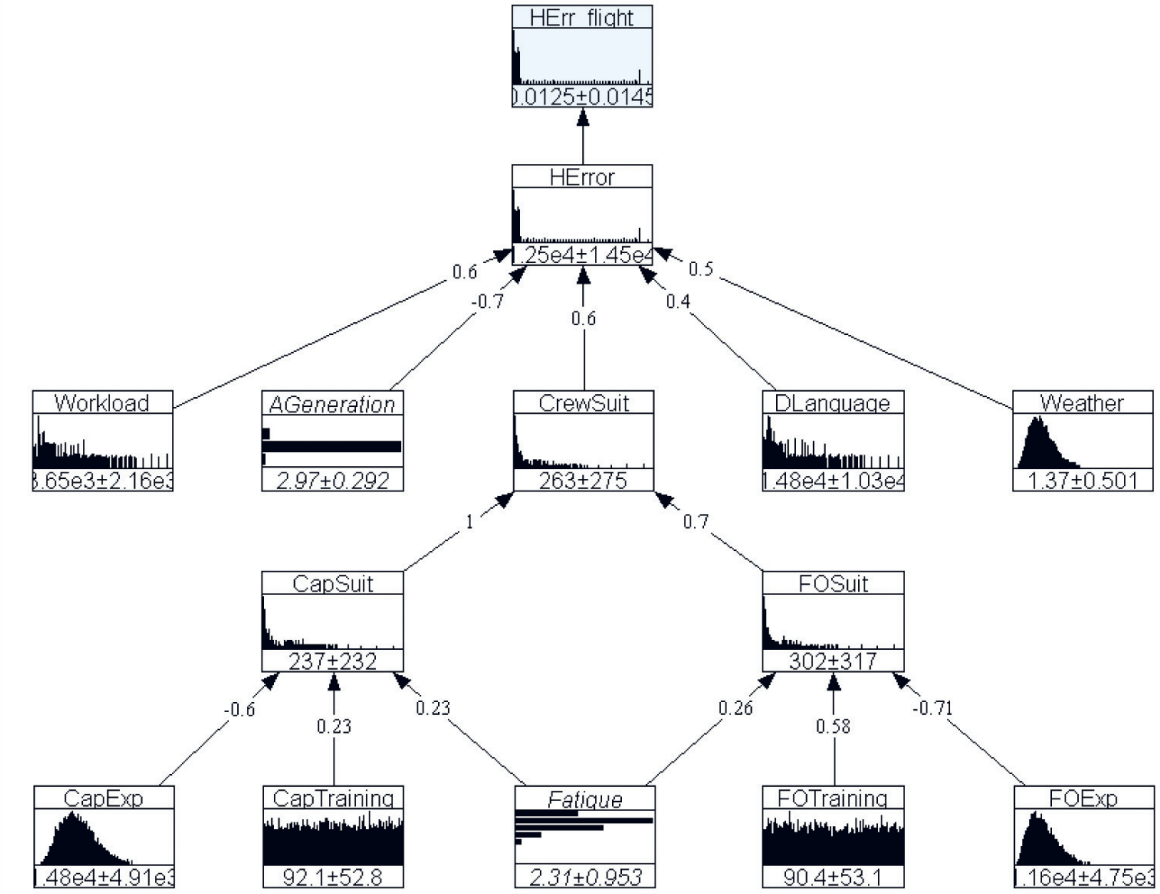


Figure 54: Flight Crew Performance Model in UniNet.(Preliminary Results)

Experience

Data regarding the experience of pilots in years is obtained from individual airlines.

Expert elicitation

Dependencies between variables were obtained from expert elicitation. An elicitation protocol has been developed by TU Delft for this purpose (Cooke and Goossens, 2000).

Figure 54 shows the Flight Crew model as represented in the software package UniNet. Each node has a histogram representing the marginal distribution of each variable. At the bottom of each histogram the expectation (and standard deviation after the ± sign) of each variable is shown. The arcs show the value of the rank.

According to Figure 54 and Figure 55 with the information obtained from *two experts*[3] one may see that in the unconditional distribution one could expect a flight crew error probability of 1.25% per flight. This particular error refers to a lack of control per inappropriate handling during landing and corresponds to the Approach/Landing phase. Other base events in the FT's representing different flight crew errors will be modelled in a similar way.

[3] In the final analysis 5 experts were used

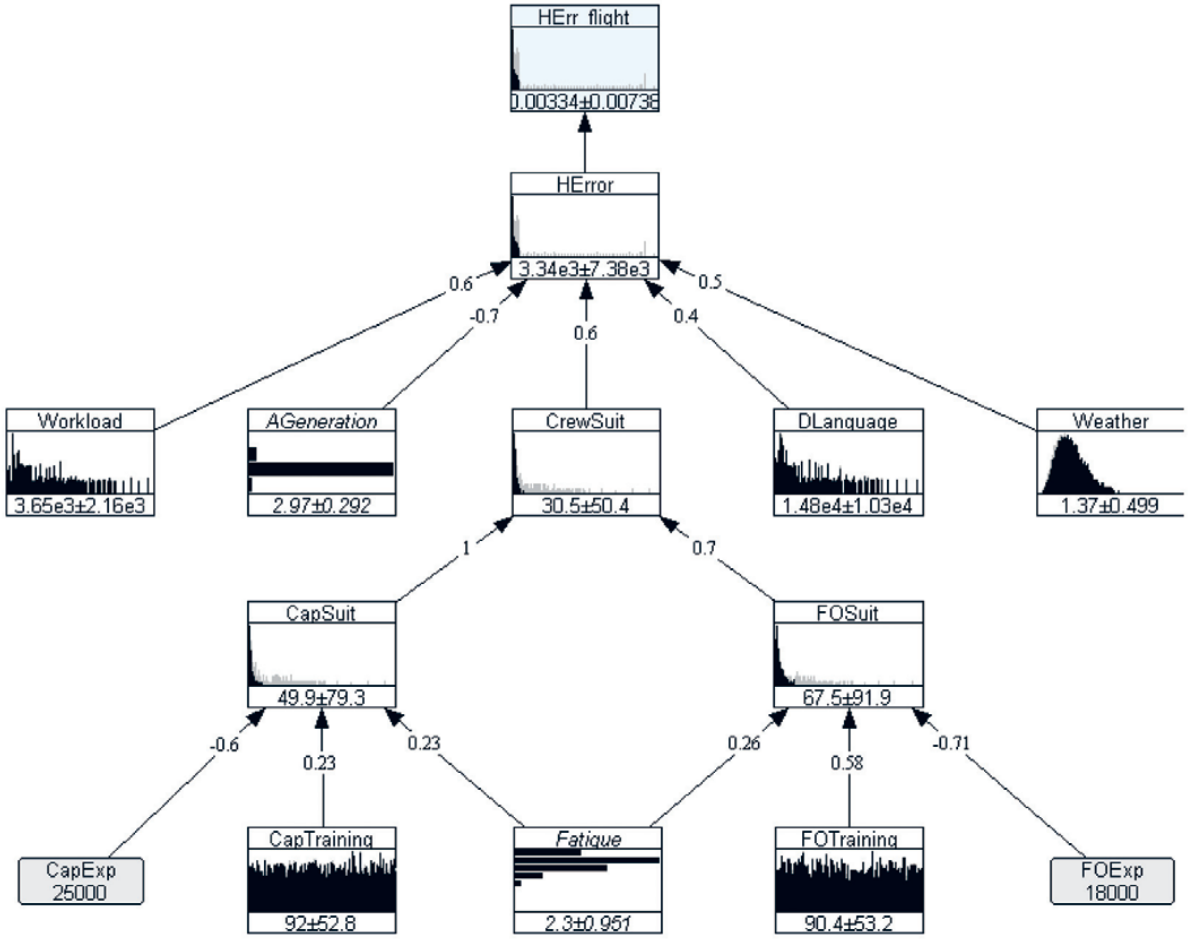


Figure 55: Flight Crew Performance Model in UniNet Conditioned on CapExp = 25,000 hours & FOExp = 18,000 hours flown since the obtaining the pilot's license. (Preliminary Results)

Figure 55 shows the conditional distribution when the captain's experience is 25,000 hours and first officer's 18,000 hours. Observe that the distributions for human error, crew suitability, first officer's suitability and captain's suitability have changed. In particular, the probability of lack of control per inappropriate handling during landing would reduce from the previous estimate of 1.25% to 0.334% in the unconditionalised distribution.

As for the variables related to suitability, the results may be observed in Figure 56. There is a significant increase (8%) in the probabilities of having a "suitable" crew, captain and first officer when each has 25,000 and 18,000 hours experience.

For the flight crew model and the air traffic controller model, 5 experts were available. Unfortunately, for the maintenance model only one expert was found to be prepared to participate in the elicitation process, which in view of the procedure described by Cooke and Goossens (2000) is insufficient.

5.2.4 Quantification

Elements of a Bayesian Belief Net with continuous variables have been quantified. Contrary to a "traditional" BBN that allows discrete variables only (usually "yes/no" or "ok/not ok" descriptions are used to limit the computational complexity), this model uses probability density functions to describe parameter values. However for the purposes of comparison a discrete version of

the model was built and results refer to this discrete version of the model. The BBN represents a 'missed approach'.
The following parameters were quantified as continuous or semi-continuous variables:

- Visibility
- Crosswind
- Aircraft fuel status
- Aircraft system status
- Flight crew alertness
- Deviation from the approach path
- Traffic separation in air.

All marginal distributions except the number of missed approach executions per 100,000 flights were quantified with real data.

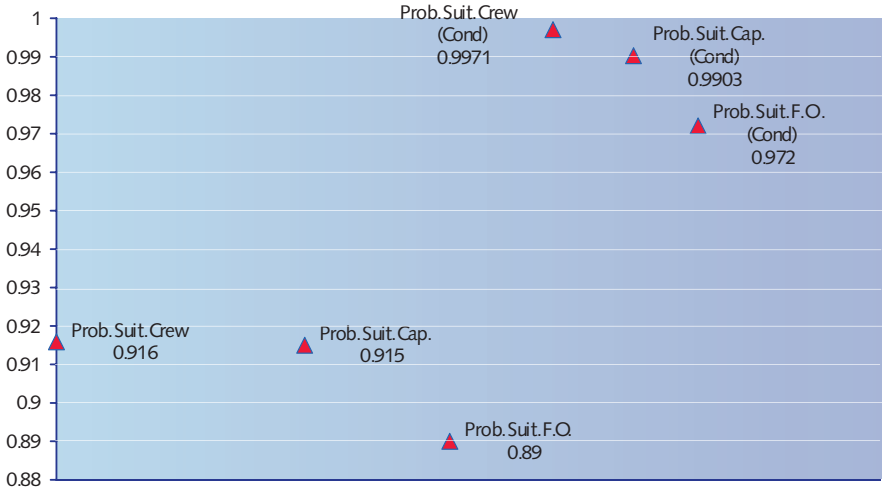


Figure 56: Probability of Suitable Crew, Captain and First Officer from the unconditional distributions and conditionalised on CapExp = 25,000 hours & FOExp = 18,000 hours flown since the obtaining the pilot's license. (Preliminary Results)

5.3 Human performance and CATS

In over 100 places in the final model human intervention is needed to avoid an accident. The probability that these actions do not result in the desired effect are described in the human performance models (HPM), which are models of influence on human error probability (HEP). Of these there are three: the crew (Roelen et al 2007) (Figure 13), the ATC controller (Roelen et al 2008a) (Figure 15) and the maintenance technician (Roelen et al, 2008) (Figure 14). The development of the Crew model is described above.
These models are linked to the base events of CATS wherever the event is designated a human action. For maintenance technicians these links are less apparent than for the crew and the ATcontroller. However each event that involves the failure of a technical component is linked to the maintenance technician model. It should be noted that the maintenance technician usually operates in a maintenance organisation and under circumstances where errors, faults or substandard performance may be corrected before an airplane is put in service. In the

current version of CATS the maintenance organisation is not further modelled. The effect of the organisation is modelled and can be changed by the user of cats through the settings of the performance shaping factors as described in annex NLR11.
This is no different than the depth of modelling of the ATC operator model. Only in the crew model a first layer of management influence on these performance shaping factors is modelled.
The quantified influence of the performance shaping factors has been derived from expert elicitation exercises with a limited number of experts. The procedure described by Cooke and Goossens (2000) requires 25 experts to get to reliable results. That many experts were just not available for the CATS development. The five experts use in for the CREW and ATC model are barely enough and the one expert available for maintenance is obviously insufficient. However this is the maximum that could be done in this project. The results of the human performance evaluation should be viewed and used with caution and be regarded as a first attempt. It is hoped that this development will encourage the industry to come forward with more expertise, in order to improve what is considered in general, and by industry in particular, one of the most important and most missed parts in modelling, estimating and evaluating risk and safety.

6 Safety Management quantification

CATS is the first model in which management influences on the performance shaping factors of human performance is explicitly modelled and quantified. In CATS this is only done for the CREW. The availability of pertinent information proved to be a major problem and the timeframe of the project did not allow for a similar exercise for the other two HPMs.

The management model used in CATS builds on the work done in I-RISK, WORM and ARAMIS (Bellamy et al 1999; Ale, 2006, Ale et al, 2006 Papazoglou and Ale, 2007), Guldemund et al, 2006).

The state of a system depends amongst other things on the way humans interfere with the system. Humans can be divided into two classes - "operators" who interfere with the system directly and "managers" who interfere with the system indirectly through a management system. In this way the management system is thought of as being separate from the technical system. The relationship between the two can then be modelled as an interface, of which operators are a part.

In the I-Risk Project (Bellamy et al, 1999; Ale et al 1998) an interface was built between the management and technical system. This concept was used later in the WORM project (Ale, 2006; Ale et al 2008; Bellamy et al, 2007) to connect the management system to safety barriers. The purpose of the management system is to have and keep safety barriers in place. Safety barriers contribute to preventing the occurrence of the unwanted event and their failures are modelled in CATS in Fault Trees and subsequently in a Bayesian Belief Net. E.g. one safety barrier might be the integrity of the aircraft pressure boundary. Another might be the separation between aircraft. These safety functions serve to keep the system operating safely for the whole of its life cycle.

Management is considered to be the process for delivering criteria and resources to keep such safety barriers intact. These resources and criteria are categorised into eight systems called DELIVERIES. These deliveries are delivered to the safety barriers by four categories of barrier TASKS.

6.1 Tasks

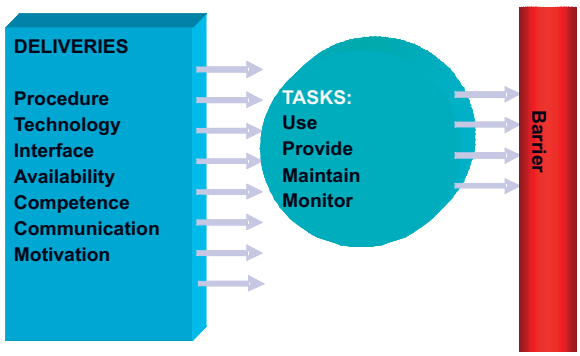


Figure 57: Deliveries and tasks of the management system

The management tasks associated with the safety functions are as follows:

- Provide = specify and design technology or human tasks and procedures
- Use = function (technology) or carry out defined actions (human)
- Monitor = check functioning of technology or people (inspect/observe)
- Maintain = restore function to designed level

This is shown in Figure 57. As using the barrier is the ultimate task it is depicted in the barrier itself.
For these tasks eight deliveries or "components of working safely" were distinguished in I-Risk. They are called deliveries because they are risk controls, criteria and resources that management should deliver to the tasks. The components are designed to be mutually exclusive and complete. That means that no additional components are needed. The eight I-RISK deliveries have been modified in CATS, combining motivation and conflict resolution and renaming 'equipment' as 'technology function' and 'ergonomics' as 'technology interface'. The definitions used in CATS are described in detail in section 6.2 below.

6.2 Deliveries

The deliveries are the components of working safely that the safety management system has to supply (deliver) to the risk control measures (technological, human or mixed). These are the following:

- *Procedures, rules, checklists and goals*: Rules and procedures are specific performance criteria which specify in detail, usually in written form, a formalised "normative" behaviour or method for carrying out an activity (checklist, task list, action steps, plan, instruction manual, fault-finding heuristic, form to be completed, etc.). Output goals are performance measures for an activity which specify what the result of the activity should be, but not how the results should be achieved. They are objectives, goals or outputs (e.g. accident/incident targets or trends, exposure of risk levels, ALARA, "safe", numbers of activities carried out, etc.). It is also convenient to regard definitions and criteria for choosing one course of action over another as output criteria. Plans refer to explicit planning of activities in time, either how frequently tasks should be done, or when and by whom they will be done within a particular time period (calendar time, flight hours, etc). They include the maintenance regime, maintenance scheduling and testing and inspection activities, which need to link to the parameters of maintenance frequency, test interval and time for maintenance and repair.
- *Availability of manpower*: allocating the necessary time (or numbers) of competent people to the safety-critical primary business tasks which have to be carried out. This factor emphasises time-criticality, i.e. people available at the moment (or within the time frame) when the tasks should be carried out. This delivery system singles out the manpower planning aspects, including the planning of work of contractors and the availability of staff for maintenance and repair work on critical equipment outside normal work hours, incl. coverage for absence and holidays. It also considers the availability of critical personnel at all times for emergency situations, coverage for peak loads, holidays, etc.
  The output of this delivery is the presence of the right people, with the right competence (next delivery system) at the right place and time, with the time to operate the risk control measures defined.
- *Competence and suitability*: the knowledge, skills and abilities in the form of first-line and/or back-up personnel who have been selected and trained for the safe execution of the critical primary business functions and activities in the organisation. This system covers the selection and training function of the company which delivers suitable staff for overall manpower planning. This competence pertains to both physical suitability (strength, and colour vision, health, etc.) and cognitive qualities of persons (situation awareness, decision making, etc.), which can be learned through training, experience and practice
  The output of this delivery is competent workers with adequate situational

awareness and decision making who can handle the safety-critical tasks they are assigned in routine/proceduralised situations or during unplanned or unexpected situations.

- *Commitment (to safety) and resolution of conflicts*: the incentives and motivation which personnel have to carry out their tasks and activities with suitable care and alertness, and according to the appropriate safety criteria and procedures specified for the activities by the organisation. Commitment also relates to non-compliance with and deviation from the procedures; violating the rule can be necessary in some situations to which standard procedures do not apply and is not of itself proof that commitment is an issue. The ability to deviate when necessary is delivered by the competence delivery system.
- *Communication and coordination*: Communications refers to exchange of information and instructions between people within the steps of any primary business activity. They are only relevant to this protocol if the activity is the on-line control of risk carried out by more than one person (or group), who may be working for more than one organisation (esp. pilot and ATC). Communications between tasks, which are represented in the other parts of the management system, are not included here, since they are represented by the continuity of activity within those delivery systems and protocols.

  Communication occurs either:
  1. verbally: face-to-face, or talk through communication channels such as (mobile) telephone, radio
  2. by (written) message emanating from: data link, e-mail, memo, briefing

  Where the communication is via instrumentation, this is covered under the delivery system for the equipment interface
  Coordination covers those mechanisms designed to ensure the smooth interaction of actions between individuals and groups working on a joint task or responsible for the correct functioning of a given risk control measure. These include plans, meetings, authorisation and communication procedures.
  This delivery links to the competence delivery, since communication is also a procedure and skill which has to be learned, e.g. CRM training. It also links to the procedure delivery, since many communication processes are formalised and subject to procedures, e.g. crosscheck and verbal confirmation.
- *Technology function*: These are the equipment & spares which are installed. This delivery covers both the correctness of the technologies for their functioning, and the availability of spares when and where needed to carry out the activities. Within this delivery are the processes of design, selection of technology, purchase, installation, adjustment, maintenance and repair, which define and ensure the good functioning.
- *Man-machine interface*: The ergonomics of all aspects of the technical systems which are used/operated by operations, inspection or maintenance. This covers ergonomic design and layout of flight decks, air traffic control rooms and manually operated equipment, location and design of inspection and test facilities, the maintenance-friendliness of equipment and the ergonomics of the tools used to maintain it. This delivery covers both the appropriateness of the interface for the activity and the user-friendliness needed to carry out the activities.
- *Resolution of conflicts*: deals with the incentives of individuals carrying out the primary business activities not to choose other criteria above safety, such as ease of working, time saving, social approval, etc. Organisational aspects of conflicts are dealt with at a higher level in the organisation, and, if not resolved, will result in choices on-line which are not optimal for safety. Conflict resolution covers the mechanisms (such as supervision, monitoring, (group) discussion, etc.) by which

potential and actual conflicts between safety and other criteria in the allocation and use of personnel, hardware and other resources are recognised, avoided or resolved if they occur. The CATS model concentrates on the on-line resolution of conflicts. The processes to avoid and resolve conflicts at levels higher in the organisation are currently not dealt with and are potential issues for later resolution.

The issue of safety culture is outstanding. On the one hand users and people working in aviation are of the opinion that safety culture is an important designator for the safety performance of organisations. On the other hand is there little scientific evidence that safety culture is an independently manageable entity. There is no clear operational definition that can be implemented into the safety management system at this point. This is required when one wants to predict the effect of change. The extent of the deliveries combined may constitute safety culture. But the issue whether safety culture is a separate entity of an overall result remains outstanding for future resolution.

A distinction is made between a hard/software implementation of a function and a peopleware implementation. Only the dominant deliveries for these implementations are modelled in this version of CATS, which means that for hardware it is the *technology-function* and *technology-interface* that are modeled and for people it is the *procedures, availability, competence, communication* and *commitment*.

The delivery systems also require resources and controls for their correct and efficient functioning. That would lead to a second layer of tasks and deliveries. In theory more layers could be added but that would lead into a multiple regression of delivery systems to delivery systems. We do not show these flows in the model, rather it is restricted to a single level. The quality of the resources and controls to run each delivery system, e.g. the competence and availability of personnel, the procedures and methods and the hard- and software for all the tasks involved in it, are therefore contained within it.

Section 6.3 is devo to the subset of the technical models of CATS which is the basis to factor in the management and organisational influences. There the principle of linking management into CATS is described. Section 6.4 shows the data analysis from two types of databases. Section 6.5 illustrates the framework of linking the complex relationship between SMS and human error in CATS and describes how the management influences in CATS is quantified by using a paired comparisons method and expert judgment on the central estimate and the range of values to consider in the calculations.

6.3 Technical model of CATS and the integration of management into CATS

Each cause of a barrier failure in a Fault Tree is a base event, which can be categorised by actor, i.e. humans and their behaviour (including flight crew performance, ATC performance, and maintenance people), technology (including ATC equipment, aircraft system, and airport infrastructure and equipment) and weather (not controllable, but where information about it can be provided and lead to decisions to apply other procedures, etc).

A base event failure can be seen as a control measure failure of the system. There are two types of control measures in CATS.

6.3.1 Human control measure

For human control measures, there are models in CATS for the behaviour of the flight crew and ATC who play an essential and interacting role at the execution level.

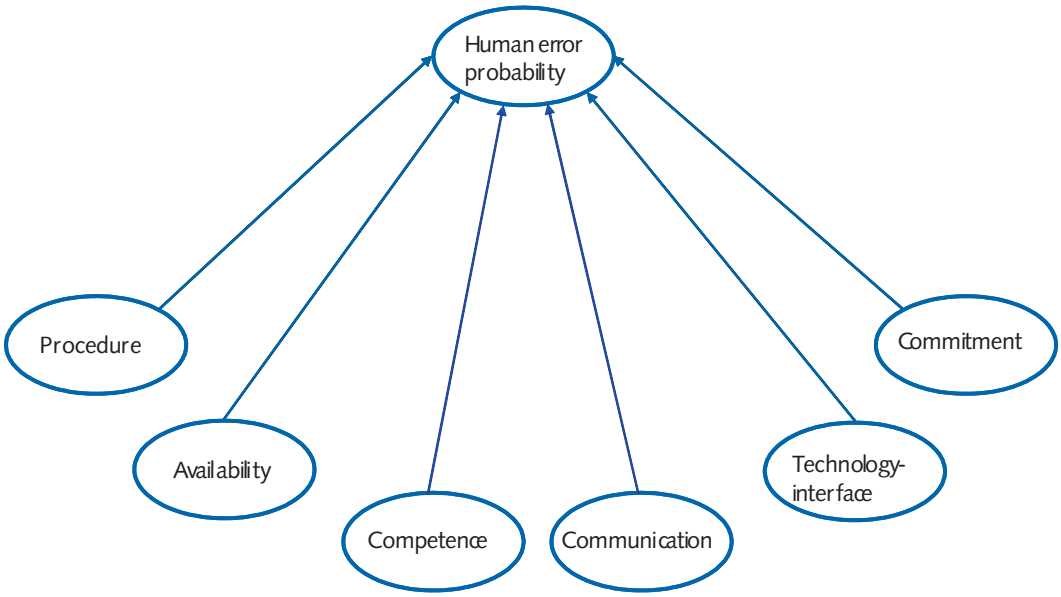


Figure 58: Linking managerial processes to human error probability

The functioning of the technology is significantly affected off-line by the maintenance personnel, for whom there is also a Human Performance Model.
The concept of the Performance Shaping Factor (PSF) in the human performance models is used to quantify the probability of human error for a certain event in the FT. The influences of the human operator on accident causation are represented in BBNs.
The human performance models represent managerial and organisational influences on human performance. These influences are depicted in terms of the delivery systems. Delivery systems provide a framework for PSF selection. The aim is to select the PSFs in such a way that each of the delivery systems is represented. See Figure 58

6.3.2 Technology control measure

For technology control measures including "ATC equipment", "aircraft system" and "airport", there are two delivery systems covering the functioning of

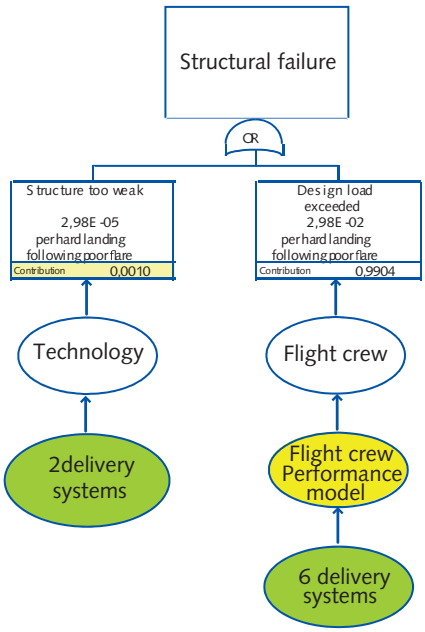


Figure 59: Delivery systems and Fault Tree

the technology and the ergonomic design of its interface, each with the tasks 'provide', monitor and 'maintain'. The interface delivery system feeds into the human performance models (Figure 58 and Figure 59), as it affects the ease and accuracy with which the person concerned can use the technology to perform the whole risk control task. The technology functioning delivery system feeds only into the hardware shows the concept of integrating management into technical model of Figure 60 shows the flight crew performance model. Each node in the flight crew performance model can be seen as an influence on the human decision making and can be supported by the 6 human delivery systems, and their tasks 'provide' 'use' 'monitor' and 'maintain'. The ATC & maintenance human performance models are described in (Roelen et al, 2008 and 2008a). Note that, a generic model like the flight crew performance model is required to represent the role of the flight crew for different situations and is not required to include every human task during a flight from A to B. Therefore, it has to be borne in mind that the model like this is a proxy for flight crew performance of air transport safety, and only focuses on the most important influencing factors. There are other influences thinkable such as the quality of training. But these influences are poorly if at all defined and currently cannot be measured on any sort of (semi) quantitative scale. The flight crew performance model contains only quantifiable influences. These are shown in Table 6.

Table 6: Initial selection of performance shaping factors

Delivery system	Selected PSF	Operationalised PSF
Procedure	left out	
Competence	Experience of captain and FO	total number of hours flown
	Training of captain and FO	the number of days since the last type recurrent training
Availability	Fatigue	Stanford Sleeping Scale
	Workload	number of times the crew members have to refer to the abnormal/ emergency procedures
Commitment/conflict resolution	left out	
Communication	Intra-cockpit communication	number of flights in which the pilot and first officer will have a different mother tongue
Man-machine interface	Technology interface	four aircraft generations

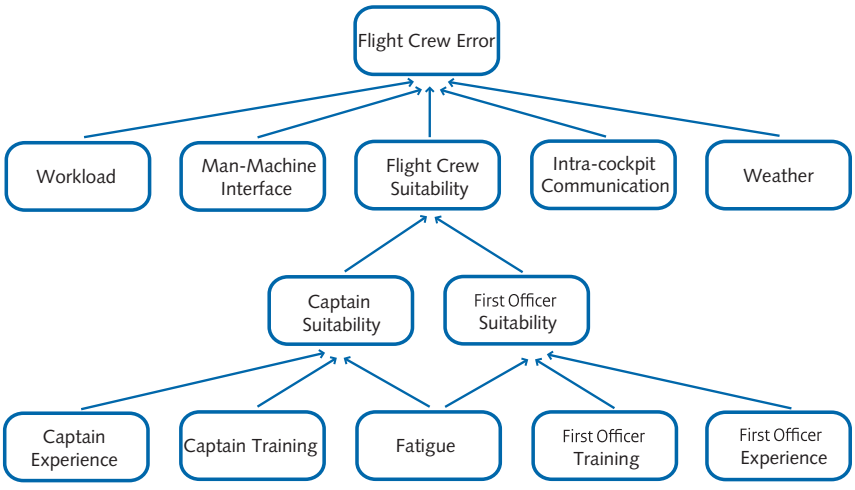


Figure 60: Flight crew performance model

An extension of these nodes to give a much fuller coverage of the delivery systems may be developed in the future.

At this stage, however, the basic influences and the representations chosen here have the advantage that the probabilities of the nodes can be quantified with objective data. The task of the management model in this implementation is to indicate and quantify how changes in the parts of the management delivery systems can influence the calculated probabilities in the CATS model. To do this data are needed about failures in the management delivery systems. These can come from accident and incident data, audit data and expert judgment. These are discussed in the next sections.

6.4 Analysis of accident and incident data for management factors

There are two types of accident and incident data collected in this research to try to quantify the relationship between the safety management system and the online failures in the CATS model.

6.4.1 Human factors in ADREP

The first source was the study of the original accident and incident reports from the ICAO Accident/Incident Reporting System (ADREP) (ICAO, 2000) from 1990 to 2006.

ICAO has a standard report format, which has been adopted by ICAO member states throughout the world. Member states are urged to submit their accident/ incident data using the standard ADREP taxonomy and report format. The detailed description of the analysis performed of ADREP data for uncovering and counting management factors is given in **Annex SSc1**.

The experience of analyzing the ADREP data shows that there is still a lack of detail about organisational factors recorded in the accident/ incident investigations all over the world. Many accident reports do not specify the organisational factors lying behind the issues of on-line risk control on which they concentrate.

There are five groups of underlying causes which have a major influence on human errors that are recorded in the ADREP data. They are

- fundamental limitations in the human sensory, cognitive and motor processes
- commitment and conflict resolution, e.g. routine violations - e.g. pilots discouraged from making a go-around due to cost implications
- online supervision, e.g. failing to notice that a task has been carried out incorrectly
- online communication and coordination problems, with ATC and between team members.
- competence of airmanship and crew resource management skills.

The data show managerial failures in delivering the functions of competence, communication, commitment, and availability dominate with 80% of all online failures. In annex SSc1 we also demonstrate that managers have to deliver different resources and controls to prevent different types of flight crew errors.

6.4.2 LOSA

The second set of incident data is from direct observation of task performance from Line Operations Safety Audit (LOSA) data (ICAO 2002).

All the analyses in Annex SSc1 are done using the ADREP accident/incident database. There is no on exposure data present there. This means that all results are about flights that ended in an accident/incident. Whether a certain management deficiency is likely to be material in producing a given accident/ incident and how many of those errors the safety management system has prevented before the accident happened can only be answered if exposure data are known. Therefore, we need information from daily operation data.

In LOSA trained observers fly in the cockpit and record the types of threat and errors committed, and how flight crews manage these situations to maintain safety during normal operations. These data can provide exposure data for the model, as they show how often particular deviations occur and are corrected. Some detailed statistical analysis on this is done in SSc2 for the LOSA dataset. The information analysed in SSc2 has been used to validate part of the Fault Tree in the CATS model by DNV.
Although the ADREP's taxonomy is much more comprehensive than the PSFs we have modeled in the flight crew performance model and we already have quantitative figures for management influences from the accident analysis, up to now it is not easy to use those quantitative data for CATS. The reason is that the ADREP data does not contain most of the exact operationalised variables defined in the flight crew performance model e.g. communication (number of flights in which the pilot and first officer will have a different mother tongue) or workload (number of times the crew members have to refer to the abnormal/ emergency procedures). The ADREP has much more generalised categories than this
Another complication is that the ADREP and LOSA datasets use different classification systems which makes it difficult to compare the occurrence of the same deviations in the two datasets. However, since it offers access to real data in a way which has not been possible in the past with management data, further research and more development work will be needed to combine these two data sets into CATS.

6.4.3 Audit data

A potential source of data directly about failures of different elements of the SMS is the Audit results from the airline audits conducted with the ICAO audit tool. However, these data could not be obtained as they are confidential to the airlines concerned.

6.5 Quantifying management factors in CATS

The approaches summarised previously have increased the understanding of failures in relation to organizational errors but have so far failed to result in quantifiable data for the management influences on the performance of human errors critical to safety. Therefore an alternative strategy has been developed to implement the management model within the CATS model. A combination of paired comparison method with direct estimation has been designed and implemented in the flight crew performance model. By this method three values are estimated: for the world average, corresponding to world average management and the two bounds of the range in CATS, which correspond to extremely poor and extremely good management respectively. The same principle can be applied in the ATC and maintenance personnel models.

6.5.1 Methods

The combination of BBNs and expert opinion has been used earlier as a quantification method for incorporating the influences of organizational factors into probabilistic safety assessment (Embrey, 1992; Oien, 2001; Mosleh and Goldfeiz, 1995; Mosleh et al, 1997). However, in the general case, as the number of parents of a node (PSF) increases the set of conditional probabilities that have to be assigned grows exponentially. This fact makes the parameter assessment by an expert or its estimation from data difficult. To bridge this gap of process complexity, a paired comparison of management policies for reducing

human errors in aviation has been developed in 4 steps. This is described in the following sections.

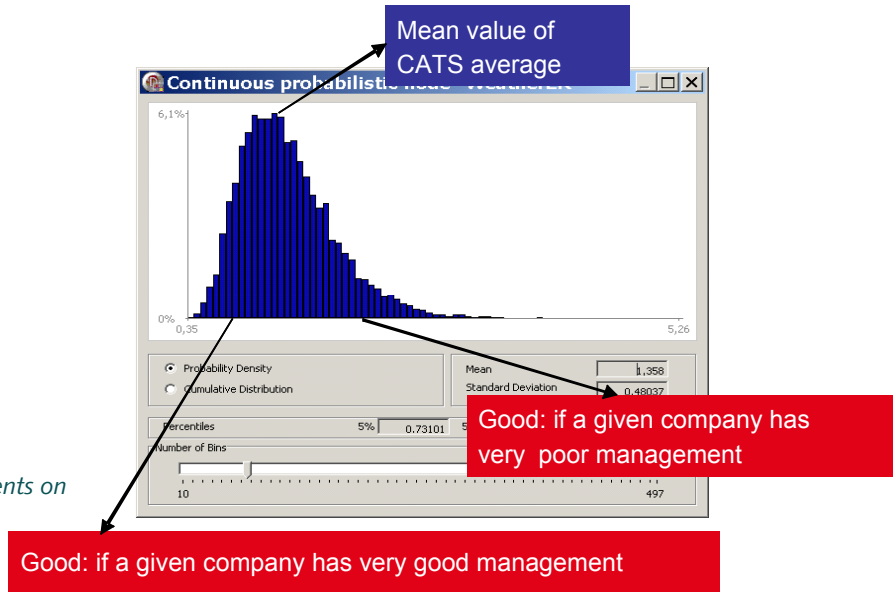


Figure 61: expert judgments on truncation points

Safety functions should constantly be used, provided, monitored and maintains. Good management can deliver and the necessary resources to maximise human performance, for example by better scheduling of processes to prevent flight crew members becoming too fatigued. Good management can also be helpful by avoiding contexts in which there are human performance is negatively influenced, such as landing in extreme weather. Hence, in principle good management can reduce the error probability of fatigue or the probability of the plane encountering the bad weather condition en route. In terms of the BBNs good management will shift the current fatigue or weather probability to the left (the red bar in the left hand side of Figure 61) by better managing the factors and conditions. Through propagation through the BBN, this will reduce human error and eventually reduce risk. On the other hand bad management can push the conditions to the other extreme, e.g. over scheduling very tired captain and first officer flying into the adverse weather. Therefore, bad management will make the situation worse and shift the current probability of the factors and conditions (PSF) to the right (the red bar in the right hand side ofFigure 61). In general the mean value of a distribution will shift. The range estimated by experts will be used as the variablity of the HPMs under the influence of management as will be described below.

6.5.2 Generating management factors

As mentioned in earlier the HPM must be able to represent managerial and organizational influences on human performance and at the same time safety management should provide the resources and criteria for the frontline workforce to operate safely. Based on literature (Smit and Slaterus, 1992; Hale et al., 1998b), the protocol of generating management factors consisted of:

- detailed definitions of the parameters, what they included and excluded
- a description of the scenarios which could lead to deviations from an optimal value for the parameter and which therefore have to be managed (e.g. fatigued pilots due to poor scheduling).
- a list of the detailed task steps directly relevant to the parameter and to the management of those scenarios. For example, fatigue is determined by the following main tasks (broken down further in the protocol): decide fatigue management concept; plan and prioritise scheduling work, etc.

- a list of management influences which could affect how well the steps and the scenarios are managed.

In Table 6 the first two columns present the initial list of selected performance shaping factors together with the associated delivery systems. Some factors which could not be represented in objectively quantifiable units, i.e. procedure and commitment, have been left out at this stage.
In the flight crew performance model, the selected PSFs have been operationalised in objectively measurable form in the third column of Table 6. As discussed earlier this is a limited operationalisation, which may not capture the full scope of the delivery systems, but it a starting point.
Based on the protocol, relevant management risk control objectives are generated and instructions and resources for each of the PSFs. The list was presented to one of the expert pilots who was asked to add any management influences he missed or eliminate any if he felt they were not appropriate.
Annex SSc3 shows how the management factors for each of the PSFs were operationalised and how they mapped onto the delivery systems. To generate those management items, each definition of the PSF node and the logic behind it has been carefully crossed referenced with its definition in the NLR report (Roelen et al 2007). There are:

- 14 management factors to prevent flight crew fatigue.
- 13 management factors to prevent the plane encountering the bad weather condition en route and reduce the weather risk.
- 4 factors for managing the aircraft system malfunction which might cause the crew members having to refer to the abnormal/ emergency procedures (A/E procedure) section of the aircraft operation manual during flight and so increase their workload.

Table 7 shows the management factors for fatigue as an example and how those management factors map with the delivery systems.

Table 7:. Management factors to reduce fatigue

Management tools	Delivery systems
Set maximum hours for per flight duty period and cumulative duty period	Availability
Set a minimum rest period after per flight and a minimum period free of all duty after a given number of consecutive days of duty	Availability
Set an average sleep requirement for 8 hours in a 24-hour period	Availability
Provide comfortable accommodation for getting good sleep at stopovers	Availability
Create a suitable crew rest environment and an appropriate placement of a nap in multicrew aircraft	Availability
Provide several days off for the flight crew to adjust to a new sleep/ wake schedule	Availability
Provide a feedback system and occurrence reporting system, which the data is used to adapt schedules	Commitment
Require crew to attend an education and training module that helps pilots to understand the cause and effect of fatigue, and teaches pilots how to minimise fatigue and its effects (e.g. NASA nap, use of bright light exposure to minimizing circadian rhythm)	Competence
Check alcohol and drug consumption for a suitable period before flying	Competence
Provide and use good fatigue assessment tools to objectively discover pilots with relatively high fatigue and performance decrement	Man-machine interface
Provide a technical alert system that informs pilots if they are falling asleep during operations (e.g. active noise production)	Man-machine interface
Provide equipment designs to improve work condition to reduce operator's on line fatigue and discomfort	Man-machine interface
Require good communication between flight crew members to openly discuss fatigue and their current ability to carry on work and, if necessary, to rotate flight tasks with other crew members	Communication
Ensure that management policy is not overridden in practice by over-scheduling tired pilots	Availability

6.5.3 Paired comparison method

Since the nodes of experience, training, intra-cockpit communication and technology interface only have one management policy for each, it is not necessary to use paired comparison techniques to elicit the relative importance of the management influences on improving any of these nodes. For fatigue, weather, and workload, the goal was to use structured expert judgement to identify the relative importance of different management factors in reducing the factors and conditions which may affect flight crew performance online. All the management factors in Annex SSc3 were transferred into questionnaires. Each questionnaire was about one of these three variables and the range of management influences on it, giving a total of three questionnaires for experts to fill in. Annex SSc4 shows the questionnaires.

Expert elicitation

The protocols so devised were presented to 7 expert pilots for a paired comparison exercise. They all had a minimum of 2,500 hours experience in flying. Each expert's opinions were elicited independently. They were run to a standard format. First the purpose of the project was introduced and the protocol was discussed step by step to allow the experts to have a chance to pose questions about definitions, scenarios, steps and influences. The majority of the experts commented on the definition of the weather node and the workload node. They remarked that the definitions may not reflect reality and also make it difficult for them to think about the management influences. However they were asked to continue and do their best. Finally, each expert was asked to work through a set of paired comparisons, in which each influence was compared to each other for each parameter. The order of presentation in each list was randomised, but was the same for each session. For each pair the expert had to indicate which influence was the more important in affecting the parameter in question.

Analysis

The data were fed into the COMPAIR program (Cooke and Solomatine, 1990) and tested for inconsistencies by analyzing the number of circular triads. Experts who could not pass the threshold for the circular triads on a given parameter were removed from the analysis for that parameter. The program assigns rank orders and weights to the influences depending on how often they are rated as the most important in a pair. Coefficients of agreement, concordance, and p value were calculated to indicate the agreement between the rank orders and between the ratings of each pair of items across all the retained experts. P-values lower than 0.05 are considered necessary before it can be concluded that there is a reasonable consensus across an expert group (Cooke, 1991) Only the results for the total group are reported here.

Table 8: Influences on fatigue (N=14)

Item name	Weighting %	Type of influence
Provide comfortable accommodation	18.6%	Availability
Management not overide tired pilots	14.6%	Commitment
Set maximum hours	10.4%	Availability
Create suitable crew rest environment	10.4%	Availability
Set a minimum rest period	9.7%	Availability
Good communication between pilots about fatigue	9.3%	Communication
Good fatigue assessment tool	6.8%	Man-machine interface
Feedback system	5.7%	Availability
Provide several dates off for new schedule	4.1%	Availability
Set an average sleep for 8 hrs	3.6%	Availability
Improve work environment	3.5%	Man-machine interface
Require crew to attend education	1.7%	Competence
Technical alert system	1.3%	Man-machine interface
Check alcohol before flight	0.3%	Competence

u=0.11, W =0.4, p=0.0009

Results

Table 8, Table 9 and Table 10 give the rank orders and weights assigned to the influences for the parameters, together with the coefficients of agreement (u), concordance (W) and p-value (p). Table 11 shows the relative weighting across the parameters of the 7 generic management influences from which the specific influences were derived.

Table 9: Influences on weather (N=13)

Item name	Weighting %	Type of influence
Equip aircraft with airborne weather radar	23.8%	Technology-function
Commitment to improve instrumentation, info provision and training for adverse weather conditions	13.5%	Commitment
Discuss weather condition before entering adverse weather	13.0%	Communication
Challenge taking unnecessary risks	11.9%	Commitment
Minimum weather criteria	11.4%	Procedure
Train flight crew to enhance decision making	7.3%	Competence
Complete a review of weather info before flight	6.5%	Procedure
Monitor weather info en route	3.8%	Procedure
Communication between pilot& dispatcher	3.1%	Communication
Create a daily strategic plan	2.4%	Procedure
Collaborate with ATC	2.0%	Communication
Provide weather information	1.3%	Communication
Rewards and taking of disciplinary action	0.1%	Commitment

u=0.15, W =0.45, p=0

6.5.4 Expert judgement on the ranges

After the importance of the management influences had been compared for each variable (fatigue, weather, workload), we were interested in how much of the variance of each variable can be explained by the (defined) management factors, the experts were asked to estimate the maxium change shift in the value of the parameter which was under the influence of the factors assessed. This was done in two steps:

- Step 1: The expert was given information on the mean and the distribution of each variable (fatigue, weather, workload) from our data set used in the BBN.
- Step 2: based on the information provided in step 1, for each variable, the expert was asked to give his judgment on a mean value range

$$Min \leq \bar{x} \leq Max;$$

where $\bar{x}$ = given mean value from our data set
Min = the estimated minimum value of the mean, given that all policies have been applied
Max = the estimated maximum value of the mean, given that none of the policies have been applied

Table 10: Influences on workload (N=4)

Item name	Weighting %	Type of influence
Aircraft Malfunction due to maintenance	64.1%	Tech-function
Aircraft Malfunction due to external factor	28.6%	None
Aircraft Malfunction due to inherent design	4.6%	Tech-function
Aircraft Malfunction due to crew action or inaction	2.7%	Competence/ commitment
u=0.40, W =0.59, p=0.0025		

Table 11: Relative weighting (%) of 7 generic influences

Parameter Influence	Fatigue	Weather	Workload
Availability	62.5%	-	-
Competence	0.3%	7.3%	2.7% (together share with commitment)
Commitment	14.6%	25.6%	2.7% (together share with competence)
Communication	9.3%	19.2%	-
Procedures	-	24.1%	-
Tech-Interface	11.6%	-	-
Tech-function	-	23.8%	68.7%
			*28.6% influenced by external factor

The range between the lower bound and upper bound of the mean is the error probability which can be explained by the given management factors. It means, for example, that the distribution of the crew fatigue as measured by the score on the Stanford Sleeping Scale in average will not go below the minimum value of the mean, given that we have efficiently implemented (provide, maintain, use, monitor) the 14 defined management policies to reduce fatigue. On the other hand, the average Stanford Sleepiness Scale will not go beyond the maximum value of the mean, given that we have managed all the policies poorly.
Based on these steps each expert was asked to work through the estimation of the shift of the mean at the end of each questionnaire. The equal weighted mean value is given in Table 12.

Table 12: Estimated minimum/maximum value of the mean influenced by management

	Minimum value of the mean	Baseline (World average) mean value	maximum value of the mean
Fatigue	1.73 (SSS)	2.31 (SSS)	3.2 (SSS)
Weather	1.02 (mm/hrs)	1.37 (mm/hrs)	1.877 (mm/hrs)
Workload	2025 (flights /100,00 flight)	2910 (flights /100,00 flight)	4475 (flights /100,00 flight)

The experts who did not pass the consistency test in the pair comparison were also taken out at this step. By cross referencing all the answers given by the experts, we found that those experts who have poor consistency in their paired comparison also deviate the most in their value estimation.
This method gives interesting and useful results, but it is open to some criticism and more work should be done to fine-tune the method for the cut-offs of the distribution and design a better mechanism to calibrate the truncation points that are given by the expert groups. This fine-tuned method can be considered as an improved way to link the management into HRA.

6.5.5 Transfer the relative weighting scale into true scale value

Using the methods described above we arrived at the relative values of management influences on improving flight crew error, but we need to conditionalise the absolute values on a certain value of the PSFs nodes and see how those management factors influence error probabilities. In order to transform the paired comparison values (the results we got from section 4.6.3) into absolute values, reference values must be supplied.
The status of the management policies differs between airlines and we do not know the baseline management information of how many airlines implement how many of the policies described. Hence setting the default from the world average does not increase the user's insight in the effectiveness of policies. To be able to do so properly, we would need data about how many airlines implement each of the range of different policies. We would need to know what "the average" airline is already doing in terms of the defined policies and therefore we would have to collect data from audits or survey on their safety management system. Since the information is unknown to us (i.e. no reference values were available), we assume that the default of management influences start from the worst case that is none of the management policies have been efficiently applied in the airlines. In other words, we use the truncation points (the maximum value of the mean which we got from section 6.6.4) to generate reference values.
The experts are likely to be judging which factors would give a relatively greater improvement in the parameter if they were to be improved, or which could give a greater degradation if neglected. Relative importance in these two directions may not be symmetrical. But in this exercise, we have had to assume that they are symmetrical and asked the experts to judge which factors would give a relatively greater improvement in the parameter if they were to be implemented. So it is logical to set the default from the management worst case baseline (no policies implemented), and improve the management of e.g. fatigue by adding different poliTherefore the transformation of the relative weighting scale into a true scale value by implementing one policy will be given by the equation:

$$=Max_{\bar{x}} - [(Max_{\bar{x}} - Min_{\bar{x}}) * \text{relative weighting scale (shown in Table: 8 - Table 10)}]$$

The true scale value of implementing multiple policies *n* will be:

$$=Max_{\bar{x}} - [(Max_{\bar{x}} - Min_{\bar{x}}) * \sum_{i=1}^{n} \text{relative weighting scale}_i]$$

By these two equations we will obtain a new value of average e.g. fatigue level for a company under condition of n policies has been applied. Through conditionalisation and propagation via BBN, we will obtain a new value of human error probability.
The advantage of this method is that manager can calculate the error probability of any influence nodes under his own setting and compare to the world average

(the mean). For example, if he selects on 5 polices which are now in place in his company but he finds out that the world average mean is smaller than he is now, it means that his company is doing less well in controlling fatigue than the world average companies do. He can decide to add more policies and see if he can reach the world average or beyond the world average.

6.6 Additional observations

The coefficients of agreement and concordance for the influences on fatigue, weather and workload are above the level at which the rank orders and weights can be reliably used as a consensus. The paired comparison method using the pilot experts in judging the relative importance of management influences has therefore been successful in this exercise. Availability has the highest relative weighting of 62% out of the 7 generic management influences for fatigue. The results for fatigue shown in Table 12confirm that fatigue is originally a PSF selected for representing Availability. The relative weighting of the management influences for weather are more complex and diffuse. They mainly distribute between procedure, commitment, and technology function. Workload as defined here is mainly influenced by how well management supports an effective policy (provide, use, maintain, monitor) to reduce aircraft systems malfunction (that require the A/E procedures to be used) due to the maintenance.
Care must always be taken in generating the operationalised performance shaping factors, as we have emphasised in the previous sections. Some nodes are too complicated to represent at this stage or we are unable to quantify them in numerical units. Therefore the nodes have been limited in their definition and modelled in a way that can be quantified by the BBN. However, this does not necessarily tell the whole story because important influences may have been lost. This makes it difficult for the expert groups to make connections with their perceptions of reality and to judge the relative importance of the management influences on the nodes. Most experts indicated that it was hard to concentrate their minds on workload defined only as the number of times the crew members have to refer to the abnormal/ emergency procedures, and a few experts had similar comments on the definition of weather as only rainfall rate in mm/hr.
Care must also be taken in interpreting the results of the expert judgements. The judgement is crucially determined by the original list of possible influences and their phrasing. Ideally more experts should have been involved at this stage to make sure no relevant factors were left out. More work should be also done to fine-tune the method for the cut-offs of the distribution (See annex SSc1).

6.7 Use of inspection data

In the course of the project it was investigated whether it would be feasible to use the results of airworthiness compliance audits (maintenance Part 145) as a route to the quantification of the management influences. Three delivery systems have been analysed: competence, procedures and communications.

6.7.1 Approach

In the BReS (Bedrijf registratie system) database the results of airworthiness (Part 145)[4] compliance audits can be found. Of interest to CATS were:

a The possibility to find management related data

b Considering how these data could be coupled to CATS to assist the regulator in risk prioritising

BReS contains the results of a number of inspections per company, specifying who is the auditor, the date of the audit, what has been audited and the findings. These results are confidential. About 200 audits a year are conducted. It was considered that these data could be analysed on the basis of results per article of the regulation, which in turn could be classified according to the delivery system scheme of the management influences described earlier.

Since the BReS contain company specific data, an anonymous set of data with company names replaced with ciphers could also be used to look at whether there were company differences in audit findings per article of the regulation. This could be useful in contemplating whether management, as a factor to be modelled in CATS, could have relevant input variables identified from inspections which could link to the delivery system part of the CATS model.

It has already been described that management is considered to be the process for delivering controls, criteria and resources to keep safety barriers intact. These resources and criteria are categorised into seven systems called DELIVERIES. These deliveries are delivered to the safety barriers by four categories of barrier TASKS. These are Provide, Use, Maintain, Monitor. For application to classification of the regulations the idea is to determine what aspect of the management of a safety barrier(s) the regulation/article is intended to support. As was mentioned earlier, the delivery motivation can be seen as consisting of two separate deliveries: motivation and conflict resolution. This was the way the delivery systems were set up originally (Bellamy et al 1999), making the number of deliveries eight rather then seven. Initially Part 145 was classified according to the eight delivery system scheme. The regulations were then analysed according to the CATS delivery system scheme definitions and the CATS task definitions) where appropriate.

Figure 62 gives the classification for Part 145 and Figure 63 for part of JAR-OPS. This classification is provisional and needs further development.

[4] Part 145 of the regulation COMMISSION REGULATION (EC) No 2042/2003 of 20 November 2003 on the continuing airworthiness of aircraft and aeronautical products, parts and appliances, and on the approval of organisations and personnel involved in these tasks

		Delivery Systems							Processes				
		Procedures	Competence	Availability	Commitment	Communication	Technology function	Man-machine system	Provision	Use	Monitoring	Maintenance / Improvement	
section A													
145.A.10	NR												Scope
145.A.15	NR												Application
145.A.20			X			X							Terms of approval
145.A.25a)							X						Facility requirements
145.A.25b)								X					Facility requirements
145.A.25c)								X					Facility requirements
145.A.25d)							X						Facility requirements
145.A.30a)			X2						X				Personnel requirements
145.A.30b)			X2	X2						X	X	X	Personnel requirements
145.A.30c)			X2								X		Personnel requirements
145.A.30d)				X									Personnel requirements
145.A.30e)			X										Personnel requirements
145.A.30f)			X										Personnel requirements
145.A.30g)			X										Personnel requirements
145.A.30h)			X	X									Personnel requirements
145.A.30i)			X										Personnel requirements
145.A.30j)			X										Personnel requirements
145.A.35			X										Certifying staff and category B1 and B2 support staff
145.A.40a)							X		X	X			Equipment, tools and material
145.A.40b)							X				X	X	Equipment, tools and material
145.A.42							X						Acceptance of components
145.A.45		X											Maintenance data
145.A.47a)		X		X			X						Production planning
145.A.47b)				X									Production planning
145.A.47c)						X							Production planning
145.A.50		X											Certification of maintenance
145.A.55		X											Maintenance records
145.A.60						X							Occurrence reporting
145.A.66a)					X								Safety and quality policy, maintenance procedures and quality system
145.A.66b)		X			X								Safety and quality policy, maintenance procedures and quality system
145.A.66c)1					X2						X		Safety and quality policy, maintenance procedures and quality system
145.A.66c)2					X2							X	Safety and quality policy, maintenance procedures and quality system
145.A.70					X2	X2							Maintenance organisation exposition
145.A.75	NR												Privileges of the organisation
145.A.80					X2								Limitations of the organisation
145.A.85					X2	X2							Changes to the organisation
145.A.90	NR												Continued validity
145.A.95					X								Findings

Figure 62: Part 145 classification. X = identified in the article,X2= second level of delivery (higher level), NR=Not relevant to delivery system classification

		Delivery systems								
		Procedures	Competence	Availability	Commitment	Communication	Technology function	Man-machine system		
SUBPART N										Flight crew
SUBPART N	OPS 1.940(a)(1)				X					Composition of Flight Crew
SUBPART N	OPS 1.940(a)(2)				X					Composition of Flight Crew
SUBPART N	OPS 1.940(a)(3)			X						Composition of Flight Crew
SUBPART N	OPS 1.940(a)(4)				X					Composition of Flight Crew
SUBPART N	OPS 1.940(a)(5)						X			Composition of Flight Crew
SUBPART N	OPS 1.940(a)(6)				X					Composition of Flight Crew
SUBPART N	OPS 1.940(a)(7)			X						Composition of Flight Crew
SUBPART N	OPS 1.940b)				X	X				Composition of Flight Crew
SUBPART N	OPS 1.943			X						Initial Operator's Crew Resource Management (CRM) training
SUBPART N	OPS 1.945			X						Conversion Training and checking
SUBPART N	OPS 1.950			X						Differences Training and Familiarisation Training
SUBPART N	OPS 1.955			X						Nomination as commander
SUBPART N	OPS 1.960			X						Commanders holding a Commercial Pilot Licence
SUBPART N	OPS 1.965			X						Recurrent Training and Checking
SUBPART N	OPS 1.968			X						Pilot qualification to operate in either pilot's seat
SUBPART N	OPS 1.970			X						Recent experience
SUBPART N	OPS 1.975			X						Route and Aerodrome Competence Quali fication
SUBPART N	OPS 1.978			X						Alternative Training and Qualification Programme
SUBPART N	OPS 1.980			X						Operation on more than one type or variant
SUBPART N	OPS 1.981								X	Operation of helicopter and aeroplane
SUBPART N	OPS 1.985			X						Training Records
	SUBPART Q									Flight and duty time limitations and rest requirements
SUBPART Q	OPS 1.1090 1.				X					Objective and scope
SUBPART Q	OPS 1.1090 2.				X					Objective and scope
SUBPART Q	OPS 1.1090 3.				X					Objective and scope
SUBPART Q	OPS 1.1090 4					X				Objective and scope
SUBPART Q	OPS 1.1090 5.				X					Objective and scope
SUBPART Q	OPS 1.1095	NR								Definitions
SUBPART Q	OPS 1.1100				X					Flight and duty limitations
SUBPART Q	OPS 1.1105				X					Maximum daily flight duty period (FDP)
SUBPART Q	OPS 1.1110				X					Rest
SUBPART Q	OPS 1.1115				X					Extension of flight duty period due to in-flight rest
SUBPART Q	OPS 1.1120				X					Unforeseen circumstances in actual flight operations — commander's discretion
SUBPART Q	OPS 1.1125				X					Standby
SUBPART Q	OPS 1.1130				X					Nutrition
SUBPART Q	OPS 1.1135				X					Flight duty, duty and rest period records

Figure 63: JAR-OPS classification

6.7.2 Analysis of the audit data

BReS Audit data supplied by IVW had to be pieced together to create a suitable table of data for the purpose of analysis. From the original data supply only the Part 145 audits from 2004 (when the regulation changed) were used. There were 687 such audits carried out across 73 companies of which 383 audits had findings.

The results were imported into and analysed in storybuilder, preserving all the information provided by IVW.

The method of quantifying the company results was to measure:

1 The frequency of an audit per company
2 The frequency of an audit with findings per company
3 Frequency of regulation articles with findings, classified as delivery systems and per company
4 Probability of a specific delivery system (eg. Competence) finding per audit report per company
5 Frequency of levels (1, 2 or 3) of seriousness (with 1 being most serious) per company

Only delivery systems competence, communications and procedures were analysed in detail for illustrative purposes

6.7.3 Results

The results of these analysis are depicted in Figure 64 and Figure 65. Figure 64 shows the probability of there being a specific delivery system finding (one or more relevant articles) per audit report. It can be quite clearly seen that there are differences between companies.- While the probability per audit of a procedure or competence finding appears to decrease with decreasing company size, that for communications appears to increase with a decrease in company size. The important point is that the companies look very different in the patterns of delivery system results.

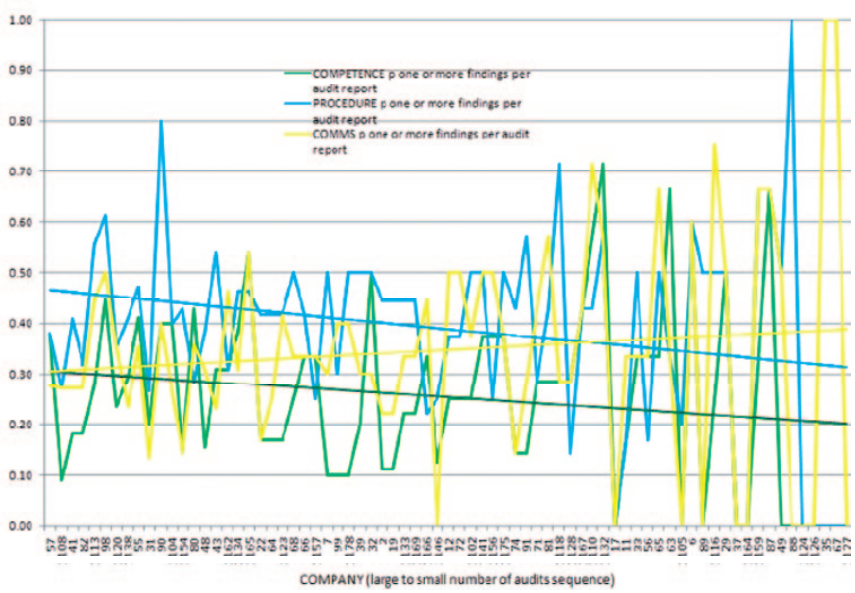


Figure 64: Probability (vertical axis) of a delivery system article finding per audit report per company (horizontal axis) showing also trendlines.

Figure 65 shows that the delivery system results are not correlated.

Companies produce different results which appear to be partly size related and partly company specific. These differences could be a mix between inspection bias (deliberately focusing on specific areas) and company performance. Delivery systems do not seem to have correlated results across companies although no statistical tests have been performed.
This is the first time delivery system results have been measured for specific companies with such a rich source of data. If maintenance incident data for aircraft serviced by these companies could be identified a very important source of data for assisting with the management model part of CATS would be provided. In addition, it assists with the mapping of the regulation onto CATS in order to support a more risk based approach to inspections.

6.8 Management in CATS

It has been possible to implement a structured method of accounting for the quality of safety management in CATS. However, the quantified effects of management on human performance shaping factors is for now completely based on expert judgement. Although the LOSA data could be used to verify probabilities in the Fault Trees, an estimate of the range of influence of management in quantitative form was not possible, neither using LOSA or ADREP or any of the other databases.
It will require an extensive analysis of the underlying data and field observations to acquire documented quantifiable evidence of the relationship between management influences and human performance.

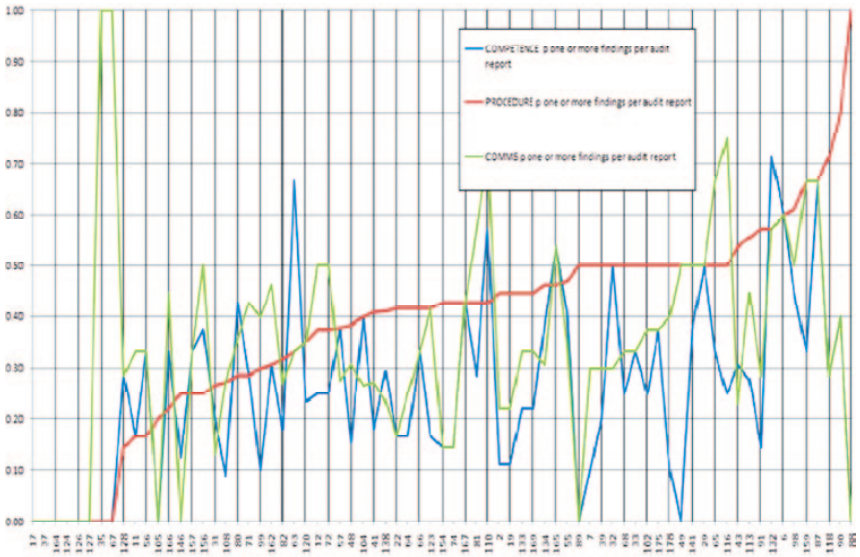


Figure 65: Companies (horizontal axis) placed in order of smallest to highest probability (vertical axis) of a procedure delivery system finding showing how competence and communications vary around this line.

7 Consequences

7.1 Requirement

In order to show the overall importance of causal factors to accident risks, the CATS model is be able to combine the results of the different ESDs into a suitable risk measure. This takes account of the consequences of each accident from each ESD. The consequence model provides fatality and damage distributions, capable of being combined with the frequencies of accidents for each ESD to give an overall risk measure.

7.2 Consequence Types

Aircraft accidents may result in diverse consequences, including injuries and fatalities to people, damage to property, disruption to business and impacts on the environment. The dominant effects are fatalities to people on-board and damage to the aircraft itself. These are therefore represented in the consequence model. In future work, it would be possible to add other consequence types, such as injuries, external (third-party) fatalities, and costs of delay and disruption resulting from accidents.

7.3 Aircraft Damage Profile

Aircraft damage profiles have been supplied for each accident from each ESD (NLR, 2008). These consist of distributions of insurance loss as a percentage of the aircraft hull value. They are based on the insurance loss recorded by Airclaims for accidents that were used to quantify the ESD.
In order to be able to combine the damage from different ESDs, the profiles have been converted to excedence probabilities for standard damage fractions of 20%, 40%, 80% and 100% of aircraft value. These can be combined after weighting the profile for each ESD by its accident frequency.
Figure 66 shows the overall damage profile for all modelled events (excluding

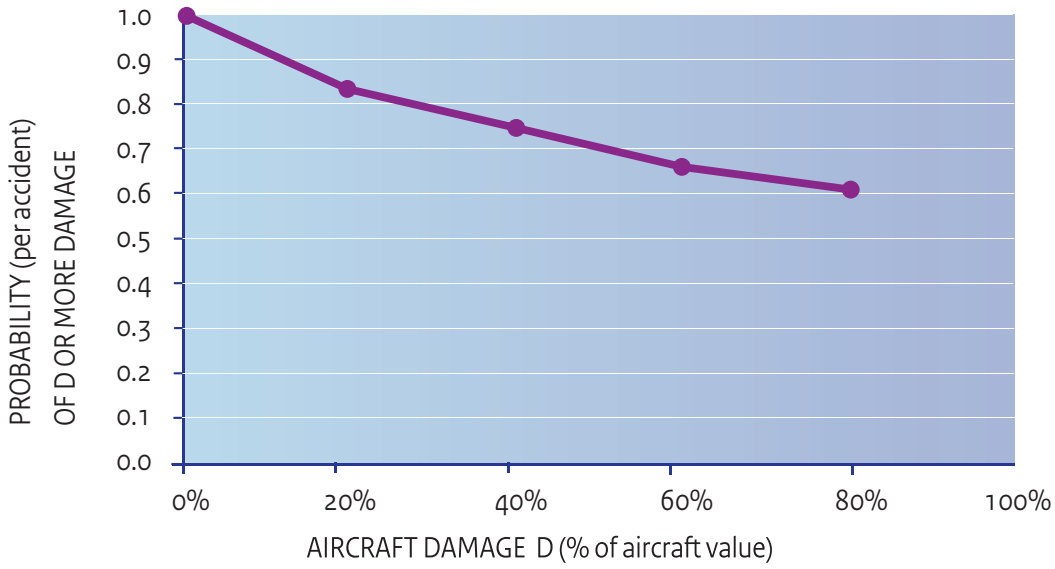


Figure 66: Overall Aircraft Damage Profile

ground collisions, ESD 36, for which damage profiles have not yet been obtained). The shape of the curve reflects the fact that most insurance policies treat the aircraft as a total loss (100%) if the repair cost exceeds 75% of the hull value.

7.4 Fatal Accident Probability

Fatalities only occur on a sub-set of accidents. The metric of "fatal accident probability" shows the relative severity of ESD end events. It is the conditional probability that an event will result in one or more fatalities.
Fatality consequences have been obtained from an analysis of fatal accidents in the ADREP database for Western commercial aircraft during 1990-2006. From a dataset of 327 accidents with fatalities on-board that can be represented by the ESDs in the CATS model, a generic fatal accident probability of 0.22 is obtained. In other words, 22% of accidents represented in the ESDs involve one or more fatalities.

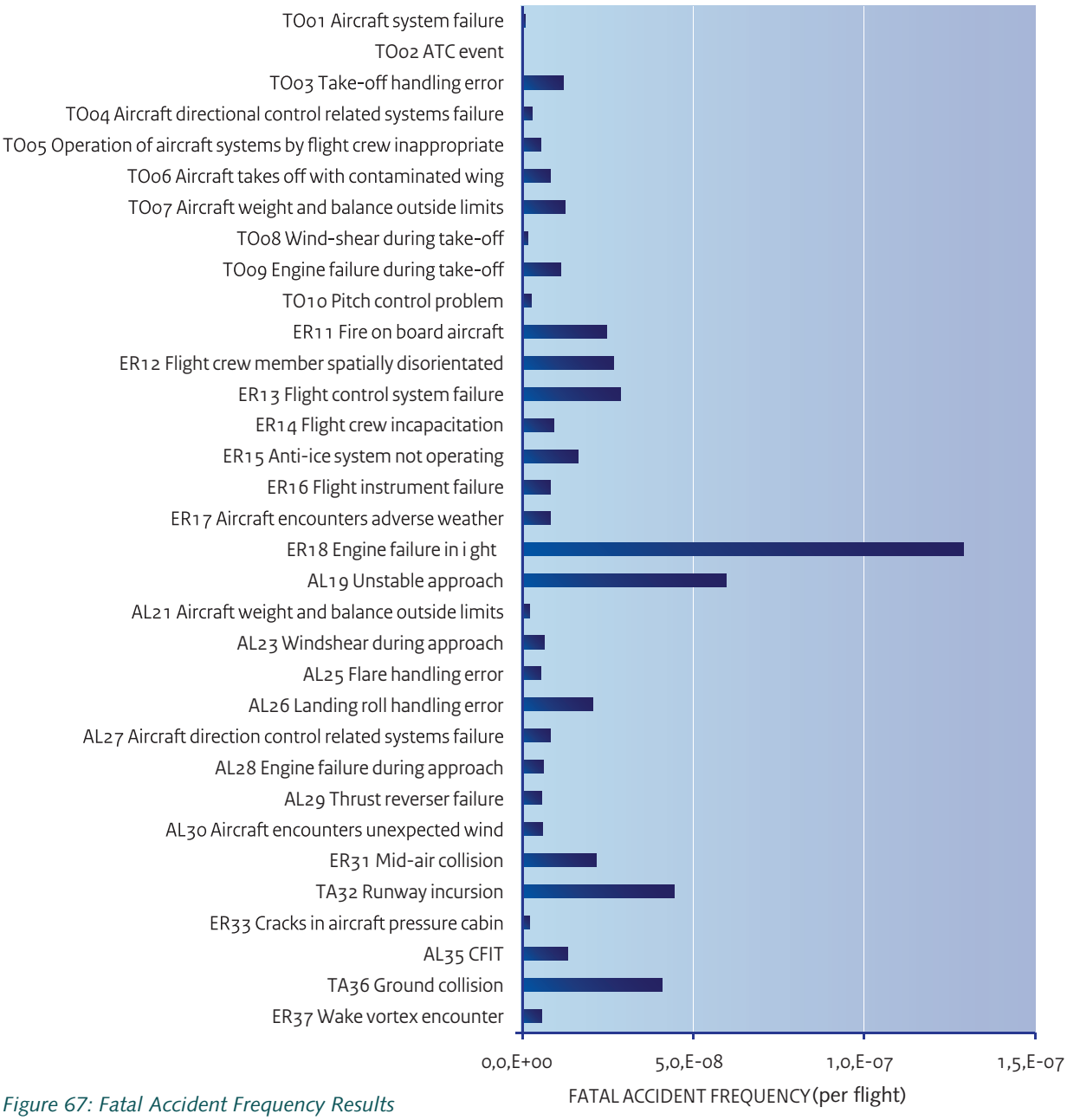


Figure 67: Fatal Accident Frequency Results

In order to estimate this probability for each ESD end event, three alternative methods have been used:

A. Where the ADREP database includes at least one fatal accident represented by the ESD end event, the probability is obtained from this experience and this ESD alone.
B. Where the ADREP database has no accidents corresponding to the ESD end event, but there are accidents in the same end state of other ESDs, the probability is obtained from this combined experience.
C. Where the ESD has used a dataset that does not match the period 1990-2006 used in the analysis of the ADREP database, the probability is obtained from a suitable alternative source in which both fatal and non-fatal accidents are known.

These approaches give a fatal accident probability for each ESD end event, even in cases where there is no fatal accident experience.

7.5 Fatal Accident Frequency

The metric of "fatal accident frequency" shows the relative fatality risk from ESD end events. It is the frequency per flight of accidents involving one or more fatalities.

The total fatal accident frequency estimated by combining the consequence model with the ESDs is 5.6 x 10-7 per flight. This is a preliminary estimate, as a full uncertainty distribution will be determined from the completed BBN model. Figure 67 shows the breakdown for each ESD. It shows that risks are dominated by engine failure and unstable approach. The relatively low contribution from CFIT results from the assumption that all commercial aircraft are fitted with TAWS. This illustrates a prediction that can be made by the CATS model, but which is not available from accident data.

7.6 On-Board Fatality Profile

The profile of fatalities in an accident is conventionally shown as a frequency-fatality (FN) curve, which is a complementary cumulative distribution function (CCDF) of the number of fatalities (N). In order to show the consequences in isolation, this can be expressed as a distribution of the probability of fatalities, given that a fatal accident has occurred. In order to make the distribution independent of the number of people on-board (POB), the fatalities can be expressed as a fatality ratio (FR), defined as the fraction of POB that are killed. For consistency with aircraft damage, the resulting CCDF of FR is termed a "fatality profile" below.

On-board fatalities have been obtained from the above dataset of 327 accidents from ADREP with fatalities on-board that can be represented by the ESDs in the CATS model. Figure 68 shows the fatality profile in this data, from which a generic average fatality ratio of 0.787 is obtained. In other words, an average of 79% of people on-board are killed, given a fatal accident represented in the ESDs.

In order to allow the fatality profile to be modified in a consistent way, it is represented by a Beta distribution. In future work, it would be desirable to improve the fit in the region of 100% fatalities.

There are insufficient fatal accidents to obtain a fatality profile for each ESD end event. Therefore, the fatality profile is obtained for the following accident types:

- CFIT (Controlled Flight into Terrain) – ESD 35
- Collision (Mid-Air Collision and Runway Collision) - ESD 2, 31, 32, 36, 37
- Fire (Fire on Board) – ESD 11
- LOCF (Loss of Control in Flight) – ESD 12 – 16, 18, 37
- LOCL (Loss of Control in Landing) – ESD 19, 21, 23, 25 – 30, 37
- LOCT (Loss of Control in Take-off) – ESD 1, 3-10, 37
- Structural (Structural Damage) – ESD 17, 33

These profiles can then be modified to represent each end event type. The modified distribution is selected so that its expectation value reflects the FR in the accident data for the end event type, while its standard deviation responds to this modification in a way that is characteristic of the accident type. The parameters of the Beta distribution are determined from this. The result is an on-board fatality profile for each ESD end event that is a smooth and consistent adjustment of the data for each accident type. This can be combined with the accident frequencies and fatal accident probabilities, and cumulated for all ESD end events. The resulting fatality profile can be combined with the POB to give the overall FN curve.

7.7 Consequence Factor Model

The analytical fatality profile above allows the FN curve to be modified in response to influences on the accident consequences. As with the influences on the accident frequencies above, modelling of the dependencies between these consequence factors is challenging and can be addressed in future work. At present, a simple model of independent factors is available. Although the effects of these factors on overall risks have been established directly from the data, it should be noted that most aircraft accident investigations do not analyse consequence factors in detail, and so data interpretation uncertainty is large at present. Furthermore, the effects of these factors on individual ESDs are not yet validated.

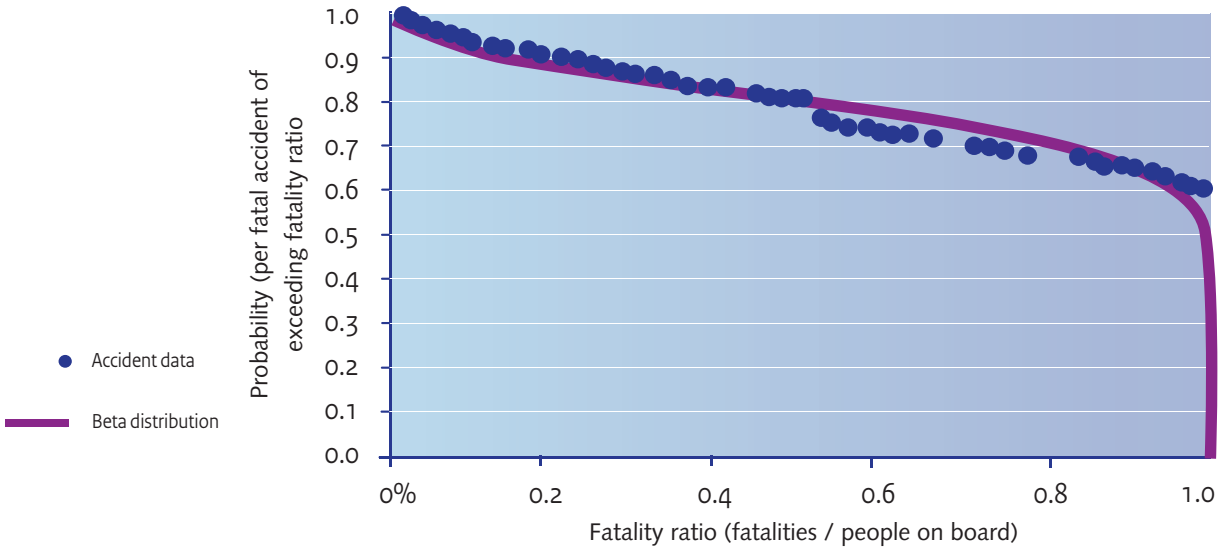


Figure 68: Overall On-Board Fatality Profile

7.8 Overall Accident Costs

A metric of "accident cost" combines the overall damage and fatality risk from ESD end events, in the form of the expected accident cost per flight.
Cost estimates accidents on commercial aircraft have been based on the ASICBA project [71]. These include an average aircraft value of €70 million; an average of 9 crew and 85 passengers on board, an average of 0.3 third party fatalities per accident, and a value of preventing a statistical fatality of €2.5 million. These are combined with the average damage fraction and the average on-board fatality ratio, to obtain the expected total cost of each ESD end state. The average accident cost estimated from the consequence model with the inputs above is €8.0 million per accident, which is an average cost of €240 per flight when spread over all commercial flights world-wide. This is a preliminary estimate, as above, and excludes ground collision, whose costs have not yet been accurately estimated. The breakdown by cost type is:

- 59% from physical damage to the aircraft
- 38% from on-board fatalities
- 4% from airport closure
- 10% from other modelled components

Figure z shows the results for each ESD. It shows that accident costs are dominated by engine failure, unstable approach and ground collision, with a relatively low contribution from CFIT results as above. This illustrates the type of information that could be used as an input to cost-benefit analysis of risk reduction measures.

8 Validation

There are a number of steps that will be taken to satisfy the user that CATS produces an adequate description of reality and that the results of any analysis or prediction using CATS can be used as a valuable input to decision making. In Table 13 ways of performing validation are summarised.

Table 13: Different ways to verify a model

Verification	Check whether the model works as specified
Calibration	Make the model give the known answers when known inputs are given to the program
Validation	Check that the model gives the right anwers when given an input for cases not used in calibration
Face validity	Check whether experts are of the opinion that the model corresponds to the real world
Case validity	Check whether the model gives answers for a specific case which correspond to reality
Sensitivity analysis	Check whether the sensitivity of the model to changes in the inputs correspond to reality
Convergent validity	Check whether the model gives answers for a specific case which correspond to then results of **another model, which describes the same case**
Scientific Peer Review	Put the model in front of scientific peers for comment

The model has been validated as far as was possible given the available data, as is described in the next sections.

8.1 Verification

There has been extensive checking of the correct workings of the program and the correct quantification has been verified by several methods
The Fault Tree Model is first implemented in a spreadsheet, and the results are validated in some detail.
It has been verified that the Fault Tree has been constructed correctly. The first line of defence against errors in Fault Tree construction is DNV's quality assurance (QA) system. This requires:

- Definition of responsibility for each part of the work. This is defined in the report on each accident category and in the documentation within the Fault Tree package, which is to be included in CATSPAWS.
- Detailed self-checks by the responsible person at each stage of the work. This is an essential part of DNV's competence training for risk analysts. Some of the available checks are explained below.
- Independent review of each part of the work. The extent of this review depends on the competence of the responsible person, as judged from previous reviews. The identity of the reviewer is documented in the report on each accident category.
- Full documentation of the work through project reports. The project report (Spouge, 2008) provides detailed documentation. Summaries are included in the Fault Tree package, which are copied into the parameter database CATSPAWS.

No QA system can guarantee that there are no errors in a model as complex as the CATS Fault Trees. Nevertheless, a high degree of error correction is achieved, through use of the following self-checks:

- Each Fault Tree Model is implemented in a gate-by-gate form, showing all intermediate results, which are also included in the project reports. This allows manual checks of each stage in the model, and this has been effective in identifying errors in the model.

- Each Fault Tree Model is implemented twice, quantified once from the top down (developing base event probabilities), and once from the bottom up (recalculating the top event probabilities). The fact that this returns numerically identical probabilities helps trap a high proportion of errors in model construction.
- The contributions of causal factors for each barrier in the Fault Tree sum to 1. This allows a simple check against numerical errors in these results. The contribution of each cause of barrier failure also allows a check against input data and subjective expectation.
- The Fault Trees have been implemented independently as BBNs, which provides a further verification that the calculations are consistent with the chosen logic gates.

In addition to checks of the Fault Trees themselves, there has been some cross-checking of the top event probabilities against the corresponding events in the ESDs, in the cases where the Fault Trees and ESDs were quantified independently between DNV and NLR. Quantification of epistemic uncertainty has also motivated consideration of alternative sources of probability estimates, which increases the quality of the result.

8.2 Calibration

The model was calibrated against the available data. This was done in two stages. First the model was calibrated with the interdependencies implicit. In this calibration a direct comparison wit the data for all accident types is still possible. Subsequently the model was recalibrated after the dependencies were made explicit. It should be noted that the actual numbers in section of the BBN can no longer be directly compared with historical data, because the historical data have the dependencies in them. A mechanism was designed to achieve frequency conservation of total accident by modifying the major discrepancy contributors in the model. Therefore, large fractional discrepancies on other related end events are unavoidable (changing frequency of one pivotal event will change the frequencies of at least two end events, some increase and others decrease). Effort has been made to ensure large discrepancy fractions only occur on small contributors.

After the influence of management has been made explicit a further calibration is necessary, but could not be finished within the scope of this project.

It will be necessary to develop a more automated way of recalibrating the model, as it is to be expected that more implicit mechanisms, that cause interdependency, will be made explicit, now that the modelling technique allows it. As described in section 2.20 an extensive bookkeeping system was developed in order to make it possible that every accident or incident used in the analysis could be traced. In addition in CATSPAWS the source of each number used in the analysis can be retrieved. This also makes it possible to amend numbers quickly if further data warrant this.

8.3 Face validity

Several example calculations have been discussed with the experts in the various advisory groups associated with the project. No real anomalies were found. But further validation needs to be done.

8.4 Validation

Validation against an independent data set was unfortunately impossible, because all available data were used to populate the model with numbers and calibrate the model on the historical accident rates.

8.5 Case validity

The application of the model on specific case studies has been postponed to a later project.

8.6 Sensitivity analysis

A tool to perform sensitivity analysis has been developed and no effects of changing the input were detected that could not be explained by a closer look at reality. Further testing remains necessary and is envisaged for a follow up project.

8.7 Convergent validity

Comparison with other models of this type have been postponed to a later project.

8.8 Scientific Peer Review

Validity of the approach can also be verified by invoking peer review by the scientific community. This is done by publishing papers in the open literature. The following papers have been published

2006 B.J.M. Ale, L.J. Bellamy, R.M. Cooke, L.H.J.Goossens, A.R. Hale, A.L.C.Roelen, E. Smith, Towards a causal model for air transport safety – an ongoing research project, SAFETY SCIENCE, Volume 44, Issue 8, October 2006, Pages 657-673.

2007 B.J.M. Ale, L.J. Bellamy, R van der Boom, R.M. Cooke, L.H.J.Goossens, A.R. Hale, D. Kurowicka, O. Morales, A.L.C.Roelen, J. Spouge, Futher Development of a Causal model for Air Transport Safety in Risk, Reliability and Social Safety – Aven & Vinnem (eds), Taylor and Francis, London, ISBN 978-0-415-44786-7

2008 B.J.M. Ale, L.J. Bellamy, R. van der Boom, R.M. Cooke, L.H.J. Goossens, D. Kurowicka, P.H. Lin, A.L.C. Roelen, H. Cooper, J. Spouge, Further development of a Causal model for Air Transport Safety (CATS); The complete model., PSAM9, Hongkong, May 2008

2008 B.J.M. Ale, L.J. Bellamy, R.P. van der Boom, J. Cooper, R.M. Cooke, D. Kurowicka, P.H. Lin, O. Morales, A.L.C. Roelen, J. Spouge, Using a Causal model for Air Transport Safety (CATS) for the evaluation of alternatives, ESREL2008 (ACCEPTED)

2008 B.J.M. Ale, L.J. Bellamy, R van der Boom, R.M. Cooke, L.H.J.Goossens, A.R. Hale, D. Kurowicka, O. Morales, A.L.C.Roelen, J. Spouge, Futher Development of a Causal model for Air Transport Safety in Risk, Reliability and Social Safety – RELIABILITY ENGINEERING AND SYSTEM SAFETY, (ACCEPTED).

2008 Lin, P.H., Hale, A.R., Gulijk, C. van, Ale, B.J.M., Roelen, A.L.C., Bellamy, L.J. 2008. Testing a Safety Management System in aviation, paper presented at the Ninth International Probabilistic Safety Assessment and Management Conference, PSAM 9, Hong Kong, China.

2008 O. Morales, D. Kurowicka, A. Roelen. 2008. Eliciting Conditional and Unconditional Rank Correlations from Conditional Probabilities, Reliability Engineering and System Safety, Vol. 93, Issue 5, p. 699-710.

9 User wishes

In this chapter the requirements of potential users and the wishes of the client are summarised. These are compared with what could be achieved in the current project.

An initial overview of model requirements was obtained from two rounds of interviews with representatives of possible users of the model. The interviewees were members of the CATS expert group and the CATS advisory group and were therefore already familiar with the high level objectives and the modelling approach. Seven interviews were conducted in the first round, and nine additional interviews were held in the second round.

The first round of interviews was aimed at obtaining general requirements for the model. An important response was that the goal of CATS was not entirely clear. It was felt that, as different goals and scopes ask for different choices in model development, the absence of a clear goal was hampering model development. According to the interviewees, CATS should be a tool to determine the safety effects of future changes, thereby supporting decision making and policy development. A secondary use of the CATS model could be as a tool to provide understanding on how safety depends on the operation, indicating main risks and providing strategic directions for safety improvement. The results of CATS should be quantified, reliable and understandable. It should not be a black box. The model should allow analysis of (the safety effects of changes to) processes that support the operation, such as management, organisation and the environment. CATS is seen as a tool for analysing strategic changes, rather than tactical or operational.

The second round of interviews intended to make the initial requirements more specific by addressing the following three topics:

- How should safety culture be made tangible in the model
- Which variables should be included in the model
- How can the model be made transparent

There was no unanimous vision on how to express safety culture in the CATS model. Therefore, according to the interviewees, more knowledge on the subject should be acquired. It was also felt however that this may be beyond the scope of the CATS project.

On the level of detail of the models and the variables to be included in the model, the general opinion of the interviewees was that this is difficult to define a priori. Instead it was suggested to deal with this topic in a series of user test sessions. During these sessions the users should be allowed to apply the model and provide comments and feedback to the model developers.

Transparency of the model is considered to be important but it is also acknowledged that this is different for different groups. In order to be transparent and understood, the results of CATS should be explained in operational terms, depending on the user. It is important that details of the CATS model are fully accessible to specialists so that it can be peer-reviewed.

The model described in the previous sections of this report is implemented in a computer based system. The model and the system together are designed to fulfil the requirements of the client and the user base identified by the client. These requirements have been developing over time and even at the end of the project were not completely established. The requirements are listed in the following table, together with the way these requirements are implemented in CATS. Many of the requirements demand a priori validation of CATS, which because of reasons given earlier could only in part be done and which will be the main purpose of a continuation project.

Main element	Client wishes	Remarks	CATS implementation
Scope			
1.	The model will describe commercial aviation as a whole.	It should give a Global picture of the risks involved in commercial fixed wing aviation from gate to gate, including aerodrome operations, ATS operations, aircraft operations in all flight phases for all aircraft involved. The model shall reflect the risks deriving from the aircraft design and aircraft maintenance as well. Specific attention will be given to modelling of Western and European Aviation (EU/EASA)	See chapter 3 and 4
2.	The model will be as detailed as possible, taking into account the available incident and accident data as well as figures deriving from expert opinion.		Fulfilled
3.	Transparent model	The model will be transparent to the user as well as to the (future) model developer. It will show the background of the data and the chain of events that leads to accidents	There is extensive documentation and a repository for data, giving all sources and derivations
4.	The model will lead to a good understanding of causes of accidents as well as the aftermaths of accidents		Fulfilled
5.	The model may be used as a device to monitor and evaluate developments in aviation safety		Fulfilled. It is noted that continued use of the model is only possible if the model and the underlying data are continuously maintained.
6.	The model will be scientifically valid		Assured by having the model description and the approach published in peer reviewed scientific journals. However additional validation is necessary (Chapter 8 of this report)
7.	The model will be set up in a way to make it useful on the long term.	The model will be suitable for later adoption of changes with respect to new data, a higher level of detail and incorporation of changes in aviation	The BBN approach chosen assures this
8.	The model will be useful for all potential users.	Users from the aviation sectors may be found in air operators, aerodrome operators and Air Trafic Service providers. Users from the Ministry of Transport may be safety inspectors as well as safety policy developers The users are to be found in Europe and the developed aviation countries	There has been an extensive discussion with the users. Further discussion with the users outside pilot sessions was curtailed by explicit instruction of client not so spend any more resources on this issue
9.	The model will show or facilitate prioritisation of risk mitigation measures		Priority may be determined by risk but also by the cost effectiveness of potential measures. The costs of accidents can be determined using CATS. But since the costs of measures is unknown, CATS only facilitates the prioritisation but in itself supplies insufficient information
10.	The model will categorise aircraft and aircraft operations	Aircraft may be categorised by the propulsion system. The operation categories may be pax, freight, air taxi.	Fulfilled
11.	The model will show the risks of accidents that occurred or may occur	These risks will be calculated from accident and incident data as well as expert opinion. The model will show the background/contribution to the overall calculated risk	Fulfilled
12.	Relate to ICAO definitions		The ICAO definition of accidents is taken. However accidents on board, not involving the state of the aircraft, and occupational accidents in ground handling are not part of the model

Main element	Client wishes	Remarks	CATS implementation
Model			
13.	The model will show the safety of aviation operations including the related personnel and the relevant technical elements as well as the way it is managed from the safety perspective		Fulfilled
14.	The model will show causes, consequences and interactions.		Fulfilled
15.	The model will define the relation between the nodes in the model and relevant regulations	Where international regulations are applicable these will be defined. Where national regulations apply a facility will assure a choice of national regulations	This is an outstanding issue. It was not part of the original scope of work. The feasibility is established and coding has started, bit could not be finished within the timeframe of the current project
16.	Uncertainty bands of input and output parameters will be available		There is a method for the user to explore uncertainty. In all parameters and variables.
17.	Interdependencies will be taken into account		The BBN structure allows taking these interdependencies. Interdependencies that are derived from the analysis of accident reports and data are explicitly modelled.
18.	The model will contain a cost benefit trade off sub-model		Due to the absence of data regarding the costs of measures the client has decided to not implement this feature. The costs of accidents can be evaluated using CATS.
Requirements for using the model			
19.	For ease of use of the model a user friendly user interface is needed.	The interface may show categorised input variables. These variables have to be set consistently throughout the model. The options for settings will reflect specific operational, technical conditions and time periods (f.i. last year). Safety indicators may be shown	The user interface has been developed in cooperation with the client. The wishes of the client have been incorporated as far as possible given the available information and the available budget.
20.	The user interface needs a professional appearance	Settings may be memorised, including the related output. The output range of an input variable may be shown A comparison of different settings will be possible Selections can be made in the overview of session outcomes A printout of the results will be possible.	Fulfilled.
21.	The model will be set up in a way to facilitate a variety of users	One may consider users from the side of authorities as well as aviation sector. These users my be either capable to use the standard build in tools or be sufficiently known to the model to introduce a specific input, new data, new model elements or new working methods in handling CATS	The client has decided to postpone further developments until a later phase of the CATS development.
22.	The user interface shall support the user in finding and ranking the most risky activities as well as the related causes		Fulfilled. A special toolkit is supplied for performing detailed analyses (UNISENS).
23.	The user interface shall support the user in finding common causes		CATS can do this, however also CATS has limitations in the sense that there always are unknown unknowns
24.	The risk contribution of a user organisation may be shown		If it is meant that the results may be evaluated per airline, this is possible given that certain data on the airline are known. It is not possible to evaluate the contribution of the Ministry of VenW to safety.

Main element	Client wishes	Remarks	CATS implementation
25.	Safety implications of changes in aviation as a result of decision making and policy changes may be shown		This requirement is fulfilled, It should be noted though that the analysis is as good as the supporting information.
26.	Risks related to safety oversight may be shown	Relation between the model and rulemaking (of the country involved), including the risk figures; Contribution of oversight(element) to the overall risks	For this an estimate of the effect on the probability of failure on intensity and quality of oversight is necessary. If this is known CATS can be used to evaluate the resulting risks
27.	Impact of management of safety may be shown		A first level of modelling is implemented.
Methodes of analyses			
28.	Safety implications of strategic decisions can be analysed	These decisions may be safety oriented or otherwise like economic or environmental.	In as far they can be translated into model parameters
29.	Safety implications of non compliance to rulemaking can be analysed	This may include legal rules and requirements as well as company rules	The relationship between rules, compliance and technological elements in the model is uner development outside this project.
30.	Lack of rulemaking in risk causes including the importance of the risks can be analysed		This is currently not possible.
31.	Safety implications of operational and technical aspects can be analysed	Issues to consider may be for instance: Runway layout, Air space configuration, Traffic flow, means of communication, Technological changes.	Implemented
32.	Impact of variables upon operational decision making can be analysed	E.g. Impact of time pressure on decision making	Implemented. However care should be taken when using CATS on a detailed level that all parameter settings are tailored to this detail.
Output requirements			
33.	Predefined safety indicators		Fulfilled
34.	Compliance ratio		Unknown quantity
35.	Friction in interfaces		Unknown quantity
36.	Risk contributing factors	Problem Statements	This is the essence of a causal model and fulfilled
37.	Identification of precursors	This may be derived from the relation between accident and incident data	Fulfilled
38.	Selection of time intervals		When this means that the settings are part of the output, this is fulfilled
39.	Selection of operational conditions		Idem
40.	Selection of technical conditions		Idem
41.	Selection of areas in the world	E.g. Western Operators, EASA operators, African operators, Regions.	Idem
Presentation of results			
42.	Risk contribution of causes may be shown		Can be done through UNISENS
43.	Risk Contribution of participants/ participants categories in aviation can be shown		Can be done through UNISENS.
44.	The model may support communication to the public on policy decisions and safety performance		No, models do not do this. The results may support the decision making.

From this table it can be concluded that most of the requirements of the client and the wishes of potential users could be realised. However there are a number of issues that could not be resolved. Many of these issues remain because of the limitations of the available data or because the information became available too late in the project. The technology in CATS also allows further development in areas were previously tools for analysis just were absent. Management influences is one of them. In the last chapter of his report the potential for further development will be discussed as part of the overall conclusion of this work.

10 Conclusion

A computer based Causal model for Air Transport Safety has been developed. In its present state it can be employed for the quantitative analysis of air transport risks.

The Causal model for Air Transport Safety (CATS) integrates models for technical failures such as event sequence diagrams, Fault Trees, event trees and models for human behaviour in a single BBN.

All potential accidents are divided into accident categories, which collect similar types of accidents with similar groups of causal factors for analysis in one part of the model. For each phase in a journey – taxi, take-off, en route, approach, landing and taxi – these categories of accidents are developed in event sequences, as preparation for inclusion into the BBN. These Event Sequences form the "backbone" of the model development. The events in the event sequences are the broad parts of accident scenarios such as the loss of control or the decision to abort a take-off.

Using Bayesian Belief Nets as the modelling vehicle proved to have several advantages over using Fault Trees and event trees.

First of all in BBNs the events do not have to be linked deterministically as in the trees. In trees the state of an event can only be a binary quantity: yes of no, true of false. BBNs support both functional and probabilistic nodes. Roughly, this means that they can capture all functional relations and also dependences between probabilities of occurrence of base events.

The BBN structure also allows analysis of the correlation of accidents with the underlying causes. As was discussed earlier, in a system with a highly reliable system such as the air transport system there are not many accidents for which a single defined cause can be established. Correlation analysis may give a lead to combinations of more extreme values of parameters is the system, that could cause an accident. A system was developed which displays the distributions of parameters associated with a certain selection of values of other parameters or variables.

The final outcome is the probability of an accident. In this BBN the interdependencies between different sections of the model, such as the relationship between engine failure, fuel starvation and go-around manoeuvres are already introduced. Here the real power of using a BBN over the event and Fault Trees starts to manifest itself. The effects of interdependencies on the final result can be modelled directly.

No less useful is the fact that the states of the nodes can be distributed over many values and that this distribution can be continuous rather than discrete and that the edges of the BBN are – conditional – correlations.

A model such as CATS has large data requirements, the major problem being the exposure data. It is not sufficient to know how many failures of a certain piece of equipment are recorded in an accident database. It is necessary to also know how many failures of that same instrument occurred without an accident and in how many flights the equipment did not fail at all.

Data are gathered from ICAOs ADREP database, from data made available by airlines and by airports. In addition work data is used from the Line Operation Safety Audit (LOSA) database to establish the performance of pilots with and without accidents. If the performance was – in part – influenced by the equipment or by circumstances these underlying causes were taken into the model whenever possible.

For each number in the model, the uncertainty in the estimate was expressed by a standard deviation. The estimate and the 5% and 95% quantiles are used in

the BBN to define the variability in the values used.
When no data could be found, expert judgement was employed. For this the European Standard method developed was employed to maximise the chance of unbiased estimates.
For several entities in the model proxy entities needed to be established.
For the development of CATS in all a few thousand numbers needed to be extracted or estimated. The origin and a characterisation of the quality of the data are held in a separate database. This not only helps future users of CATS in interpreting the results of an analysis, but also forms a basis for recording data in the future. By targeted recording, weaknesses and holes in the data structure can gradually be remedied.
The software to drive the BBN is open source The software to build the model is developed by TU-Delft especially for the project. Full documentation can be found on www2.
In many cases experts use aggregate notions such as the complexity of an airport, the complexity of airspace, good or adverse runway conditions and aircraft generation. These notions translate into changes in probabilities of many of the model constituents. Therefore a translation or mapping has to be made of the variables or notions common in the industry and the base events of the BBN. In CATS the estimates from experts and the estimates from data are brought together in one system. Calculations are performed to establish a consistent picture between all the "known" quantities in the BBN by adjusting the "unknown" quantities.

10.1 Uncertainty

In the course of the development and testing several occasions have been identified where the total of the information is inconsistent. In this stage of development this issue has to remain unresolved. In the next stage of the development CATS will be used to explore discrepancies between expectations, judgments and reported facts. Even when the model is kept relatively simple there are many layers in the model when safety management systems are taken into account. Differences of a factor of 1.5 build up quickly to orders of magnitude. This may be seen as an argument against quantitative modelling as the accuracy of these models then cannot be better than orders of magnitude. It should be borne in mind though that the estimates of experts are equally loaded with uncertainties. The currently dominant way of making decisions on the cost effectiveness of investments in safety, safety measures and safety management is mainly based on expert opinion. The deception in time that measures did not bring what was expected is the unavoidable result, if these opinions consistently overestimate effects of change.

10.2 Care

Validation of the CATS model has only been possible to the extent that changes in past safety performance resulting from design decisions are calculated correctly. The available data are barely enough to populate the model with the required initial set. Independent quantitative validation is impossible. Therefore other approaches will be used to maximise the validity of the model, such as comparison with other existing models, expert and peer review on the equations, probabilities and distributions used. Once this validation has been

done, the model will be used first as an additional input to safety decisions in the Netherlands air transport industries. It took about 20 years between the conception of a causal model for chemical plants and the introduction into the legal system in the Netherlands. A similar cautious introduction of these sorts of techniques in the air transport industry should be expected.
Care must always be taken in generating the operationalised performance shaping factors of human beings. Some influences are too complicated to represent at this stage or we are unable to quantify them in numerical units. Therefore the nodes have been limited in their definition and modelled in a way that can be quantified by the BBN. However, this does not necessarily tell the whole story because important influences may have been lost.
Care must also be taken in interpreting the results of the expert judgements. The judgement is crucially determined by the original list of possible influences and their phrasing. Ideally more experts should have been involved to make sure no relevant factors were left out. More work should be also done to fine-tune the method for the cut-offs of the distribution.
The work that started three years ago resulted in a single Bayesian Belief Net structure to describe the probability of an air transport accident. The first applications indicate that the model functions correctly and produces results that are in accordance with observations and expert insight.
However CATS or similar quantitative methods, which bring together reports, observations, facts, opinions, judgments and expectations can help to improve our insight into what can make air traffic safer. It also suggests a pathway to a further development of methodology in other strands of quantified risk analysis. Expected and unexpected outcomes will need to be carefully evaluated in the next period to gain confidence in this new way of building a causal model.
By virtue of the use of a single BBN, interdependencies could be rigorously modelled and the human performance models could be integrated. For this reason alone the results and performance of the model already exceed the initial expectations.

10.3 Further work

Further work needs to be done on validation and on human response and management modelling.
The current workings of the model were carefully checked and rechecked to avoid errors. The inputs were reproduced and a few preliminary case studies showed good behaviour. Nevertheless full validation against an independent dataset was impossible and therefore trust in the model can only be gained by applying the model to a series of test and real cases if possible comparing the results with results from other modelling efforts elsewhere in the world, even if the latter are much less comprehensive.
There is an obvious need for further data. It would be of great help is company specific data on incidents could be used to get an even better estimate of the probabilities of events earlier in the causal chain.
The human response modelling although much improved when compared to models in other fields still needs much improvement. The most important of these is to get a better understanding of the relationship between qualitative generally understood notions and the translation of these in real observable and thus quantifiable influence on risk and risk reduction.
Maintenance is an underdeveloped area in CATS. Although the maintenance technician is modelled, he has a much more indirect influence on the system

than crew and ATCs, whose decisions and actions are directly in the causal chain. It would be advisable to investigate whether a simplified version of the maintenance model developed for the FAA could be attached to CATS, when the FAA is in a position to release the model.

In summary the following further development would be advisable:

- Thorough testing of the model. CATS is a complicated model and it has been challenging to complete the model within the constraints of the current project. A model of this complexity needs extensive testing of the technology.
- A procedure and a mechanism needs to be developed to re-calibrate the model when implicit interdependencies are made explicit.
- Further validation by using the model for studies of various problems in the air transport system, such as changing air traffic management or the changing pattern of air traffic caused by the changing behaviour of the air traveller. This could involve further detailing of the model e.g. replacing "airport complexity" by parameters that constitute complexity.
- Further analysis of data, especially data on human performance. Amongst others these analyses should comprise acquiring and analysis of audit data.
- Improvement of the human response model, particularly the model for maintenance and the actions of management in conjunction with the modelling of his operational environment as shaped by company management. This involves replacing the current proxy variables by the real variables and modelling the management organisation.
- Improvement of the management influence model. This includes a further enrichment of the – abstract – deliveries in terms of actual actions and development of the management influences on ATC and maintenance. The latter involves acquiring data, for which no source could be identified in this project.
- Further acquisition of (confidential) data. This could involve co-operation with database owners to facilitate the extraction of information and the transfer of this information into the CATS data set.
- The consequence modelling could be enriched by adding injuries and operational costs such as those resulting from delays and other disturbances in the air transport process. The model could in principle also be linked to existing models for third party risk to obtain a complete picture of the risk involved in air transport operations.
- Finally a few preliminary results indicate that fatigue should be a major concern and further research into what constitutes workload, the onset of saturation, boredom and sleep deprivation and the relationship could be beneficial for future safety.

10.4 Finally

The current model is – as one member of the international expert committee put it – the second best representation of the reality, reality itself being the best. It provides a much safer testing ground for extreme and unexpected circumstances and new developments than reality. But is remains a model. Therefore caution with the results is always a good strategy.

Glossary

ACAS	Airborne Collision and Avoidance System
ADI	Attitude Director Indicator
ADREP	Accident/Incident Data Reporting
AIRCLAIMS	Accident database of insurance companies
AL	Approach and Landing
ATC	Air Traffic Control
ATHEANA	A Technique for Human Event Analysis
ATIS	Automatic Terminal Information System
BBN	Bayesian Belief Network
CAST	Commercial Aviation Safety Team
CATS	Causal Model for Air Transport Safety
CCDF	Complementary Cumulative Distribution Function
CFTT	Controlled Flight Towards Terrain
CL	Climb
CREAM	Cognitive Reliability and Analysis Method
CRM	Crew Resource Management
DAG	Directed Acyclic Graph
DH	Decision Height
DME	Distance Measuring Equipment
DNV	Det Norske Veritas
ECCAIRS	European Co-ordination Centre for Aviation Incident Reporting Systems
EPC	Error Producing Condition
ER	En Route
ESD	Event Sequence Diagram
F/O	First Officer
FA	Fatigue
FAA	Federal Aviation Administration
FAF	Final Approach Fix
FAS	Final Approach Speed
FMS	Flight Management System
FN	Frequency-Number of fatalities
FOD	Foreign Object Debris
FR	Fatality Ratio
FT	Fault Tree
GPWS	Ground Proximity Warning System
HEART	Human Error Assessment and Reduction Technique
HEP	Human Error Probability
ICAO	International Civil Aviation Organization
ILS	Instrument Landing System
IMC	Instrument Meteorological Conditions
IRP	Integrated Risk Picture
LLWAS	Low-Level Wind-shear Alert System
LOC	Loss Of Control
LOCF	Loss Of Control in Flight
LOCL	Loss Of Control in Landing
LOCT	Loss Of Control in Take-off
MF	Modification Factor
MSAW	Minimum Safe Altitude Warning system
MTOW	Maximum Take-Off Weight
NOTAM	Notice To Airmen
NLR	National Aerospace Laboratory

PF	Pilot Flying
PNF	Pilot Not Flying
PSF	Performance Shaping Factor
PWS	Predictive Wind-Shear
RIMCAS	Runway Incursion Monitoring and Collision Avoidance System
RR	Risk Ratio
SD	Spatial Disorientation
SDR	Service Difficulty Reports
SID	Standard Instrument Departure
SOP	Standard Operating Procedure
SSR	Secondary Surveillance radar
STAR	Standard Arrival Route
TA	Taxi
TAWS	Terrain Awareness & Warning System
THERP	Technique for Human Error Rate Prediction
TMA	Terminal Manoeuvring Area
TO	Take Off
TR	Training
TUD	Technical University Delft
VHF	Very High Frequency
VOR	VHF Omni-directional Radio range
V1	Take off decision speed
VR	Rotation speed
V&W	Ministry of Transport and Water management
WQ	White Queen

References

Ale, B.J.M. and R. Whitehouse (1984), A computer based system for risk analysis of process plants. In Heavy Gas and Risk Assessment III, 5. Hartwig (Ed) D. Reidel, Dordrecht, The Netherlands. November 1984)

Ale, B.J.M. (1998), J.G.Post, L.J. Bellamy, The interface between the technical and the managment model for use in quantified risk assesment, in A. Mosleh and R.A. Bari (eds) Probablistic Safety Analysis and Managment 4, Springer 1998

Ale, B.J.M. and M. Piers (2000), The assessment and management of third party risk around a major airport, JHazMat, vol 71 nos 1-3,pag 1-16, ISSN 0304 3894

Ale, B.J.M. (2006), L.J. Bellamy, I.A. Papazoglou, A.Hale L.Goossens, J. Post, H. Baksteen, M.L. Mud, J.I.H Oh, A. Bloemhoff, J.Y.Whiston, ORM: Development of an integrated method to assess occupational risk, International Conference on Probabilistic Safety Assessment and Management, May 13-19, 2006, New Orleans, ASME, New York, 2006, ISBN 0-7918-0244

Ale, B.J.M., L.J. Bellamy, R.M. Cooke, L.H.J.Goossens, A.R. Hale, A.L.C.Roelen, E. Smith (2006a), Towards a causal model for air transport safety – an ongoing research project, SAFETY SCIENCE, Volume 44, Issue 8, October 2006, Pages 657-673.

Ale, B.J.M. (2006), The Occupational Risk Model, Final report of the Workgroup on ORM, TU Delft, the Netherlands, TU-Delft/TBM RC 20060731, ISBN 90-5638-157-1, Delft, 2006

Ale, B.J.M., L.J. Bellamy, R van der Boom, R.M. Cooke, L.H.J.Goossens, A.R. Hale, D. Kurowicka, O. Morales, A.L.C.Roelen, J. Spouge (2007), Futher Development of a Causal model for Air Transport Safety in Risk, Reliability and Social Safety – Aven & Vinnem (eds), Taylor and Francis, London, ISBN 978-0-415-44786-7

Ale, B.J.M., L.J. Bellamy, R. van der Boom, R.M. Cooke, L.H.J. Goossens, D. Kurowicka, P.H. Lin, A.L.C. Roelen, H. Cooper, J. Spouge (2008), Further development of a Causal model for Air Transport Safety (CATS); The complete model., PSAM9, Hongkong, May 2008

Ale, B.J.M., L.J. Bellamy, R.P. van der Boom, J. Cooper, R.M. Cooke, D. Kurowicka, P.H. Lin, O. Morales, A.L.C. Roelen, J. Spouge (2008), Using a Causal model for Air Transport Safety (CATS) for the evaluation of alternatives, ESREL2008 Valencia, September 22-25

Ale, B.J.M., L.J. Bellamy, R van der Boom, R.M. Cooke, L.H.J.Goossens, A.R. Hale, D. Kurowicka, O. Morales, A.L.C.Roelen, J. Spouge (2008), Futher Development of a Causal model for Air Transport Safety in Risk, Reliability and Social Safety – RELIABILITY ENGINEERING AND SYSTEM SAFETY, (ACCEPTED).

Ale, B.J.M. (2008), H. Baksteen, L.J. Bellamy, A. Bloemhof, L. Goossens, A. Hale, M.L. Mud, J.I.H. Oh, I.A. Papazoglou, J. Post, Quantifying occupational risk: The development of an occupational risk model Safety Science, Volume 46, Issue 2, February 2008, Pages 176-185

Arbuckle, P.D. (1998), Abbott K.H., Abbott T.S. & Schutte P.C., Future Flight Decks, Paper presented at the 21st Congress of the International Council of the Aeronautical Sciences, Melbourne, Australia.

Arshinov, Vladimir (2003), and Christian Fuchs (eds) , Causality, Emergence, Self-Organisation, http://www.self-organization.org/results/book/EmergenceCausalitySelf-Organisation.pdf

Bellamy L.J. (2006), Oh J.I.H., Ale B.J.M., Whiston J.Y., Mud. M.L, Baksteen H., Hale, A.R., Papazoglou, I.A., Storybuilder: The new interface for accident analysis, International Conference onProbabilistic Safety Assessment and Management, May 13-19, 2006, New Orleans, ASME, New York, ISBN 0-7918-0244

Bellamy, L.J. (1999), Papazoglou I.A., Hale A.R., Aneziris O.N., Ale B.J.M., Morris

M.I., Oh J.I.H., (1999), I-Risk: Development of an integrated technical and management risk control and monitoring methodology for managing and quantifying on-site and off-site risks. Contract ENV4-CT96-0243. Report to European Union. Ministry of Social Affairs and Employment, Den Haag, (D612-D)
Bellamy, L.J. (2007), Ale B.J.M., Geyer T.A.W., Goossens L.H.J., Hale A.R., Oh J.I.H., Mud M.L., Bloemhoff, A., Papazoglou I.A., Whiston J.Y., 2007. Storybuilder—A tool for the analysis of accident reports, Reliability Engineering and System Safety 92 (2007) 735–744
Bellamy, L.J. (2007), Experiences with using ECCAIRS ADREP Database in the CATS Project White Queen report 070423-07 22 April 2007.
Bellamy, L.J. (2008) B.J.M. Ale, J.Y. Whiston, M.L. Mud, H. Baksteen, A.R. Hale, I.A. Papazoglou, A. Bloemhoff, M. Damen and J.I.H. Oh, The software tool story builder and the analysis of the horrible stories of occupational accidents Safety Science, Volume 46, Issue 2, February 2008, Pages 186-197
Bellamy, L.J. (2008), Ale B.J.M., Whiston, J.Y., Mud M.L., Baksteen H., Hale A.R., Papazoglou I.A., Bloemhoff A., Damen M. and Oh J.I.H., 2008. The software tool Storybuilder and the analysis of the horrible stories of occupational accidents. Safety Science 46 (2008) 186-197
Bellamy, L.J. and Roelen, A., (2006), Integrated ESD model SB40_a. White Queen report 060714-07, July 2006
Bellamy, L.J. and Roelen, A.,(2006), Integrated ESD model SB40_a. White Queen report 060714-07, July 2006
Bruno Debray (2004), Christian Delvosalle, Cécile Fiévez, Aurore Pipart, Henry Londiche, Emmanuel Hubert , Defining safety functions and safety barriers from fault and event trees analysis of major industrial hazards, Proceedings of PSAM7 (C.Spitzer , U.Smocker and V.N. Dang eds), ISBN 185233827X, Springer, 2004
Chalmers, D. J. (1996), The Conscious Mind: In Search of a Fundamental Theory (Oxford: Oxford University Press)
Cooke, R.M. (1991), Experts in Uncertainty – Opinion and Subjective Probability in Science. Environmental Ethics and Science Policy. Oxford University Press. Oxford.
Cooke, R.M. and L.J.H. Goosens (2000), Procedures guide for structured expert judgement, European Commission, EUR 18820 EN
Cooke, R.M., Solomatine, D. (1990), EXPO - Integrated System for processing expert judgements: User's Manual - Version 1.2. Report to the Dutch Ministry of Housing, Physical Planning and Environment. Delft University of Technology.
Degani, A., Wiener, E. (1994), On the design of flight-deck procedures, NASA Contractor Report 177642.
DNV (2001) Causal model for Air Transport CM-DNV-018
Duijm, Nijs Jan, (2004), Andrew Hale, Louis Goossens David Hourtolou, Evaluating and Managing Safety Barriers in Major Hazard Plants, Proceedings of PSAM7 (C.Spitzer , U.Smocker and V.N. Dang eds), ISBN 185233827X, Springer, 2004
Duijm, Nijs Jan, (2004a), Henning B. Andersen, Louis Goossens, Andrew Hale, Frank Guldenmund, David Hourtolou, ARAMIS project: Effect of safety management's structural and cultural factors on barrier performance, 11th International Symposium on Loss Prevention and Safety Promotion in the Process Industries, 31 May - 3 June 2004, Prague
Embrey, D.E. (1992). Incorporating management and organisational factors into probabilistic safety assessment, Reliability Engineering and System Safety 38.
EUROCONTROL (2006) , "Main Report for the 2005/2012 Integrated Risk Picture for Air Traffic Management in Europe", EEC Note 05/06, March 2006.
EWI (2007) interim report

FSF (1996), Flight Safety Foundation, FAA Human Factors Team Report on the interfaces between flightcrews and modern flight deck systems. Published in FSF Flight Safety Digest Vol. 15 No. 9/10, September-October 1996, Flight Safety Foundation, Alexandria, VA, USA.
FSF (1999), Flight Safety Foundation, Propulsion System Malfunction Plus Inappropriate Crew Response (PSM+ICR), Flight Safety Digest, Vol. 18, No 11-12, November-December 1999, Flight Safety Foundation, Alexandria, VA, USA.
FSF (2001). Flight Safety Foundation, Understanding Airplane Turbofan Engine Operation Helps Flight Crews Respond to Malfunctions, Published in FSF Flight Safety Digest Vol. 20 No. 3, March 2001, Flight Safety Foundation, Alexandria, VA, USA.
Groeneweg, J. (1998), "Controlling the Controllable", DSWO press , Leiden, 1998, ISBN 90 6695 140 0
Guldenmund, F. (2006), A.R. Hale, L.H.J. Goossens, J. Betten and N.J. Duijm, 2006, The development of an audit technique to assess the quality of safety barrier management, Journal of Hazardous Materials, Volume 130, Issue 3, pp. 234-241.
Hale, A.R. (1998), Heming, B.H.J., Smit, K., Rodenburg, F.G.Th., van Leeuwen, N.D. (1998b) Evaluating safety in the management of maintenance activities in the chemical process industry. Safety Science 28(1), 21-44
Hale, A.R. (2007), B.J.M. Ale, L.H.J. Goossens, T. Heijer, L.J Bellamy, M.L. Mud, A. Roelen, H. Baksteen, J. Post, I.A. Papazoglou, et al. Modeling accidents for prioritizing prevention Reliability Engineering & System Safety, vol 92 nr 6 pp 735-744
Hancock, P.A., Williams, G. & Manning, C.M. (1995), Influence of Task Demand Characteristics on Workload and Performance, The International Journal of Aviation Psychology, 5 (1), 63-86.
ICAO (2000), International Civil Aviation Organization, Accident/Incident Reporting Manual (ADREP), ICAO, Montreal, Canada.
Hollnagel, E. (2006) BarriersAndAccident Prevention, ISBN 0-7546-4301-8, Ashgate, Burlington, USA.
ICAO (2002), International Civil Aviation Organization, Line Operations Safety Audit (LOSA), DOC 9803-AN/761, Montreal, Canada.
ICAO (2005) Accident Prevention Programme, ICAO, MONTREAL, 2005
Kirwan, B. (1994), A practical guide to human reliability assessment, Taylor and Francis, London, UK.
Mosleh A, Goldfeiz E.B. (1995), An approach for assessing the impact of organizational factors on risk. Technical Research Report CTRS-B3-08. Center for Technology Risk Studies, University of Maryland, 1995.
Mosleh A, Goldfeiz E.B., Shen S (1997), The _-factor approach for modeling the influence of organizational factors in probabilistic safety assessment. In: Proceedings of the IEEE sixth conference on human factors and power plants, Orlando, FL: IEEE; 1997. p. 9/18–9/23.
Lin, P.H., Hale, A.R., Gulijk, C. van, Ale, B.J.M., Roelen, A.L.C., Bellamy, L.J. 2008. Testing a Safety Management System in aviation, paper presented at the Ninth International Probabilistic Safety Assessment and Management Conference, PSAM 9, Hong Kong, China.
Morales, O., D. Kurowicka, A. Roelen. 2008. Eliciting Conditional and Unconditional Rank Correlations from Conditional Probabilities, Reliability Engineering and System Safety, Vol. 93, Issue 5, p. 699-710.
Mosleh, A. et al (1997), "Guidelines on Modeling Common-Cause Failures in Probabilistic Risk Assessment", NUREG/CR-5485
NASA (2002), NASA Office of Safety and Mission Assurance, "Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners", August 2002

NLR (2005) Q4 report
NLR (2008a), "Quantification of Event Sequence Diagrams for a Causal Risk Model of Commercial Air Transport", NLR-CR-2008-464, October 2008.
NLR (2008), "Aircraft Damage and Occupant Fatality Profiles for a Causal Risk Model of Air Transport Safety", NLR-CR-2008-231, May 2008.
NN (1992) Wet van 18 juni 1992, houdende algemene regeling met betrekking tot het luchtverkeer (Stb. 2002, 593))
NN (2003) Bevi; staatsblad The Netherlands 2004 nr 250
NUREG '2005, Good practices for implementing human reliability analysis U.S. Nuclear Regulatory Commission NUREG 1792.
Oh, J.H. (1998), W.G.J. Brouwer, L.J.Bellamy, H.R. Hale, B.J.M. Ale, J.A. Papazoglou, The Irisk Project. Development of an Integrated Technical and Managment Risk Controll and Monitoring Methodology for Managing and Quantifying On-site and Off-site Risk in A. Mosleh and R.A. Bari (eds) Probablistic Safety Analysis and Managment 4, Springer 1998
Oien, K.A. (2001), A framework for the establishment of organizational risk indicators. Reliability Engineering and System Safety 74
Papazoglou, I.A. (2002), O.N. Aneziris, J.G. Post, B.J.M. Ale, The Technical Modeling in Integrated Risk Assessment of chemical installations, Journal of Loss Prevention, 15 (2002) 545-554
Papazoglou, I.A. (2003), L.J. Bellamy, A.R. Hale, O.N. Aneziris, B.J.M. Ale, J.G. Post, J.I.H. Oh, I-Risk: development of an integrated technical and management risk methodology for chemical installations, Journal of Loss Prevention, Journal of Loss Prevention in the Process Industries 16 (2003) 575–591
Perrow, C, 1994, Normal Accidents: Living with High-Risk Technologies, Basic Books, NY, 1984, ISBN 0-691-00412-9
Papazoglou, I.A. (2007), B.J.M. Ale, Logical Model for Quantification of Occupational Risk, Reliability Engineering & System Safety, Volume 92, Issue 6, June 2007, Pages 785-803.
Rasmussen J. (1997), Risk management in a dynamic society: a modelling problem. Safety Science 27(2/3) 183-213.
Rasmussen, J. (1982), Human errors: A taxonomy for describing human malfunctions in industrial installations, Journal of Occupational Accidents, no 4.
Reason, James (1990), Human Error, ISBN 0-521-31419-4 Cambridge University Press, 1990
Roelen (2008) Rationale behind PSFs (ANNEX NLR12)
Roelen, A.L.C (2007), G.B.van Baren, J.W. Smeltink, P.H. Lin and O.Morales (2007) A Generic Flight Crew Performance Model for Application in a Causal Model of Air Transport, , NLR-CR-2007-562, National Aerospace Laboratory, 2007.
Roelen, A.L.C., G.B. van Baren, P.H. Lin, O. Morales, D. Kurowicka and R.M. Cooke (2007), A generic air traffic controller performance model for application in a causal model of air transport, NLR-CR-2007-593
Roelen,A.L.C., G.B. van Baren, O. Morales, K. Krugla (2008), A generic maintenance technician performance model for application in a Causal Model of Air Transport NLR-CR-2008-445)
Roelen, A.L.C. (2002) R. Wever. A.R. Hale, L.H.J. Goossens, R.M. Cooke, R. Lopuhaä, M. Simons and P.J.L. Valk, Causal modelling of air safetyDemonstration model, NLR-CR-2002-662
Roelen, A.L.C., G.B. van Baren, O. Morales, K. Krugla, (2008) A Generic Maintenance Technician Performance Model for Application in a Causal Model of Air Transport, NLR-CR-2008-445
Roelen, A.L.C. (2005), Huson, W.J., De Reus, A.J.C. , Key design characteristics related to inappropriate response events, NLR-CR-2005-587, NLR Amsterdam.

Roelen, A.L.C. (2006), B,A. van Doorn, J.W. Smeltink, M.J. Verbeek, R. Wever, (2006) Quantification of Event Sequence Diagrams for a causal risk model of commercial air transport, NLR-CR-2006-520, NLR, Amsterdam, October 2006
Roelen, A.L.C. and R. Wever (2005), Accident scenarios for an integrated aviation safety model, NLR-CR-2005-560, NLR, November 2005.
Roelen, A.L.C., and R. Wever, (2002), An analysis of flight crew response to system failures, PT-1 Flight crew intervention credit in system safety assessment phase 1 report, NLR-CR-2002-547-PT-1, NLR Amsterdam.
Roelen, A.L.C. (2008). Causal risk model of air transport - comparison of user needs and model capabilities, PhD dissertation, Delft University of Technology.
Russel, Bertrand, (1946), History of Western Philosophy, George Allen & Unwinn, London.
Schupp, (2004), B.A., S.P. Smith, P. Wright, and L.H.J. Goossens. Integrating Human Factors in the Design of Safety Critical Systems; a Barrier Based Approach; Human Error, Safety and Systems Development: 2004.Kluwer Academic Publishers: p. 285-300.
Smit K., Slaterus W.H. (1992), Information model for maintenance management. Cap Gemini Publishing. Rijswijk. ISBN 90-71996-56-5
Speyer, J.J., Fort, A.P. (1982), Workload Assessment for A 300 FF Certification, in Proceedings of the international conference on cybernetics and society, IEEE-82CH-1840-8, pp 608-612.
Spouge, John (2008) Fault Tree Modelling for the Causal Model of Air Transport Safety Final Report, DNV London, DNV Project No. C21004587/3, June 2008
Stein, E.S., Rosenberg, B.L. (1983), The measurement of pilot workload, Report No. DOT/FAA/EM-81/14, FAA Technical Center, Atlantic City Airport, New Jersey, USA.
Swain, A.D., Guttmann, H.E. (1983), Handbook of human reliability analysis with emphasis on nuclear power plant applications, Final report, NUREG/CR-1278-F, SAND80-0200, Sandia National Laboratories, Albuquerque, NM, USA.
Visser, J.P., (1998), "Developments in HSE Management in oil and gas exploration and production", in A.R. Hale and M. Baram, Safety Management, the challenge of change, Pergamon, 1998, ISBN 0-08-043075-9

List of annexes

#	From	Title	# pgs
1	DNV 1	Modification of Fault Tree Model in response to user inputs	30
2	DNV 2	Fault Tree Modelling for the Causal Model of Air Transport Safety - final report (J. Spouge, DNV project no. C21004587/3)	55
2a	DNV 3	Collected Fault Trees	
3	NLR 1	Accident scenarios for an integrated aviation safety model (A.L.C. Roelen & R. Wever, NLR-CR-2005-560)	68
4	NLR 2	Quantification of Event Sequence Diagrams for a causal risk model of commercial air transport (A.L.C. Roelen, B.A. van Doorn, J.W. Smeltink, M.J. Verbeek and R. Wever, NLR-CR-2006-520)	178
4a	NLR 13	Quantification of Event Sequence Diagrams for a causal risk model of commercial air transport (A.L.C. Roelen, B.A. van Doorn, J.W. Smeltink, M.J. Verbeek and R. Wever, NLR-CR-2008-646)	187
5	NLR 3	A generic flight crew performance model for application in a causal model of air transport (A.L.C. Roelen, G.B. van Baren, J.W. Smeltink, P.H. Lin and O. Morales, NLR-CR-2007-562)	65
6	NLR 4	A generic air traffic controller performance model for application in a causal model of air transport (A.L.C. Roelen, G.B. van Baren, P.H. Lin, O. Morales, D. Kurowicka and R.M. Cooke, NLR-CR-2007-593)	64
7	NLR 5	Aircraft damage and occupant fatality profiles for a Causal model of Air Transport Safety (A.L.C. Roelen & J.W. Smeltink, NLR-CR-2008-231)	79
8	NLR 6	Dependencies between Event Sequence Diagrams for a causal risk model of commercial air transport (A.D. Balk, B.A. van Doorn & A.L.C. Roelen, NLR-CR-2008-309)	79
9	NLR 7	Identification of user requirement for CATS Internal CATS document, NLR-memorandum ATSF-2006-096, (H. de Jong)	65
10	NLR 8	Second round of interviews to identify CATS User Requirements Internal CATS document, NLR-memorandum ATSI-2007-085 (H. de Jong, C. Montijn & B. van Doorn)	94
11	NLR 9	Accident costs for a Causal Model of Air Transport Safety (A.L.C. Roelen & J.W. Smeltink, NLR-CR-2008-307)	32
12	NLR 10	Quantification of Fault Trees for a Causal model of Air Transport Safety (B.A. van Doorn, J.G. Verstraeten, A. Kurlanc, A.D. Balk, J.A. Coelho, H.T.H. van der Zee & A.L.C. Roelen, NLR-CR-2008-406)	153
13	NLR 11	A generic maintenance technician performance model for application in a Causal Model of Air Transport (A.L.C. Roelen, G.B. van Baren, O. Morales, K. Krugla, NLR-CR-2008-445)	56
14	NLR12	Rationale behind Performance Shaping Factors for generic human operator models (A. Roelen, NLR Air Transport safety Institute, 2008)	12
15	EWI 1	Report on Phase 1 Causal modeling for Schipol Airport (Roger Cooke, Oswaldo Morales, Dorota Kurowicka, 2006)	9
16	EWI 2	Report on Phase 2 Causal modeling for Schipol Airport (Oswaldo Morales, Roger Cooke, Dorota Kurowicka, 2006)	8
17	EWI 3	Report on Phase 3 Causal modeling for Schipol Airport (Oswaldo Morales, Roger Cooke, Dorota Kurowicka, 2006)	9
18	EWI 4	Report on Phase 4 Causal Modeling for Schipol Airport (Oswaldo Morales, Dorota Kurowicka, Roger Cooke, 2006)	11
19	EWI 5	Report on Phase 5 Causal modeling for Schipol Airport (Oswaldo Morales, Dorota Kurowicka, Roger Cooke, 2006)	12
20	EWI 6	Report on Phase 6 Causal modeling for Schipol Airport (Oswaldo Morales, Dorota Kurowicka, Roger Cooke, I Jagielska, 2007)	21
21	EWI 7	Description of the expert elicitation results for the Flight Crew Error model (Oswaldo Morales, 2007)	16
22	EWI 8	Description of the expert elicitation results for the ATC Performance model (O Morales-Napoles, D Kurowicka, RM Cooke, 2007)	17
23	EWI 9	Quantification of non-parametric continuous BBNs with expert judgment (I.D. Jagielska, master thesis, 2007)	101
23a	EWI 10	System Level Risk Analysis of New Emerging and Spacing Protocols (G.F. Singuran, Master Thesis, 2008)	114
24	SSc 1	Analysis of ADREP data for management factors	9
25	SSc 2	LOSA data	4
26	SSc 3	Operationalized management factors for paired comparisons	15
27	SSc 4	Protocol for elicitation of relative importance of management influences on improving Flight Crew Error	3
28	SSc 5	Implement the management number into CATS-paw	2
29	SSc 6	Management mapping	
30	WQ 1	CATS Parameters With Sources [CATSPAWS] (WQ 080712-07)	18
31	WQ 2	Evaluation of the feasibility of using IVW inspection results as an input to CATS management model (WQ 080713-07)	28
32	WQ 3	CATS Software Manual (WQ 080816-07)	46
33	WQ 4	Data and Bookkeeping Summary Report (WQ 080817-07)	167

Appendix
Description of the Air Traffic System

The aviation system is characterised by a complex arrangement of organisations and activities that has virtually no geographical boundaries. In the model the system boundaries must be defined and the system must be split into a reasonable amount of sub-systems. There are many ways of conducting this subdivision but for this project we have based it upon the type of organisation, distinguishing between airline, air traffic control organisation and airport. To further limit the size of the model only the processes that are directly related to the primary process of flying from A to B are considered. The result is the schematic of aviation system processes displayed in Figure 69.
Central to this process description is the primary process of 'flying from A to B', which is subdivided into several flight phases. This primary process is supported by air traffic control, subdivided into aerodrome control, approach control and area control, and linked to the respective flight phases. Between flights there are activities related to aircraft servicing, like refuelling. In addition to the servicing activities related to the aircraft, there are continuous airport operations aimed at providing and maintaining a safe infrastructure, such as runway maintenance and bird control. Maintenance to keep the aircraft airworthy can be performed on the apron while the aircraft is being serviced between two flights or in the hangar when the aircraft requires more extensive maintenance activities.
The whole system is embedded in international and national regulation.

Flight Operations
The description in the next paragraphs regards a typical flight.

Pre-flight, crew centre
The cockpit crew normally consists of a Captain and a First Officer. On long flights an additional First Officer or Cruise Relieve Pilot (CRP) may be scheduled for the flight. The cockpit crew reports for duty at the Crew Centre at the airport typically 90 minutes before scheduled departure time.
For larger airlines much of the preparation has been done already by the Dispatcher. A briefing package has been prepared with anticipated loading figures and weights, special loads (if any), aircraft deficiencies relevant for flight planning (if any), flight plan route including speed and altitude schedule, a copy of the ATC (air Traffic Control) filed flight plan, departure slot-time (if any), actual and forecasted weather and NOTAMS (NOTice to AirMen to alert of any hazards en route or at a specific location.) for departure airport, destination, alternates and en-route airports, temperature and wind charts and significant weather charts for en-route and a minimum block fuel for the planned flight. Based on the information available the cockpit crew may choose to take extra fuel prior after which the flight plan is accepted. The final fuel figure is then forwarded to the fuelling department by the Dispatcher.
The cockpit crew then proceeds through airport customs and security checks to the aircraft and normally arrives at the aircraft approximately 40 minutes prior to departure, just before passenger boarding commences. If they have not already met in the crew centre, the cockpit crew meets the cabin crew on-board the aircraft and discusses any relevant issues with the Senior Flight Attendant, such as anticipated delays, flight time, forecasted turbulence and security procedures.

Pre-flight, aircraft at the gate

Any deficiencies to the aircraft, cockpit or cabin are stated in the Hold-Item-List (HIL) in the Aircraft Technical Log (ATL) for aircraft and cockpit matters and Cabin Technical Log (CTL) for cabin matters. The release of the aircraft by the maintenance department is done after the pre-flight inspection by the ground engineer and signed in the ATL. The cockpit crew verifies that all open items recorded on the previous flight are either repaired or added to the HIL and verifies release of the aircraft. The aircraft may be released to service with aircraft or cockpit system deficiencies if these deficiencies are stated in the Minimum Equipment List (MEL), any contingency procedures for unserviceable equipment is mentioned in the MEL as well. Acceptance of any cabin deficiencies is done by the cockpit crew in concert with the Senior Flight Attendant.

When ice accumulation on the critical surfaces is present, the ground engineer will have scheduled the aircraft for de-icing either at the gate, if ice is present on the engine fan blades, or at a remote de-icing site if the engines can be started and the aircraft can taxi on own power. The cockpit crew can override this decision by the ground engineer upon actual inspection.

During the boarding of the passengers, loading of the cargo and fuelling of the aircraft the cockpit crew prepares the aircraft for departure. On the flight deck duties are normally divided between Pilot Flying (PF) and Pilot Not Flying (PNF) but some actions are specifically assigned to the Captain or First Officer irrespective which of the two is PF on the stretch. The decision who is going to be PF on the stretch is taken either at flight planning or when arriving in the cockpit.

The time that the aircraft is parked at the gate is a busy time for the cockpit crew. One pilot (normally the PF but this also may be assigned to the First Officer) enters the flight plan data in the Flight Management System (FMS) and prepares all other aircraft systems for departure; the other pilot communicates with the gate agent, the airline maintenance department, the cabin crew, airline hub control and the Dispatcher to keep track of latest developments. When fuelling is complete the fuelling department will bring the fuelling ticket to the cockpit, the flight crew then checks the actual fuel loading with the flight plan fuel.

The latest weather observation for the departure airport is recorded from the Automatic Terminal Information System (ATIS). Approximately 20 minutes prior to departure ATC Departure Clearance Control is contacted to obtain a departure clearance for the departure runway, Standard Instrument Departure (SID) route, transponder code and ATC slot-time (if any), this is done either by voice or via datalink.

The ATIS weather observation, the latest estimate of the take-off weight and the expected departure runway are used to calculate take-off performance figures (weight limitation, engine duration, bleed-air demand and take-off speeds). When all systems settings and take-off performance calculations are examined by both pilots, these are verified by reading the pre-departure checklist. The PF briefs the taxi- and departure route, special weather conditions, applicable NOTAMS and other relevant issues such as discussing high terrain in the departure route, related safe altitudes, discuss the engine-failure procedure and highlighting some Standard Operating Procedures (SOPs).

Approximately 5 minutes prior to departure the gate agent (or alternatively the Dispatcher through datalink) hands over the cargo manifest for special cargo, passenger information list, load sheet and any last-minute changes to this data, all based on the actual loading figures. The latest data are entered in the FMS and checked against the take-off performance data that was based on pre-flight estimates.

Push-back and engine-start

When all passenger and cargo doors are closed and the push-back truck has arrived, start-up and pushback clearance is requested from ATC Start-Up Control and Ground Control successively. The direction of the push-back and possible subsequent pull-out is prescribed per gate or overruled by ATC. During the push-back or after the pull-out the engines are started according to SOP. During the push-back the PF normally remains in contact with the push-back driver, the PNF remains in contact with ATC Ground Control.

Taxi-out

After completion of the push-back the push-back truck will disconnect. When the engines have successfully been started and relevant cockpit systems have been set for taxiing, the taxi clearance is obtained from ATC Ground Control. The taxi clearance indicates the cleared taxi route from the gate to the departure runway. This may be a standard route if this route is published in the airport information or else a detailed instruction on which route to take. Runway signs and markings as well as a map of the airport layout are used by the cockpit crew to navigate on the airport. During taxi-out aircraft systems and configuration are set and verified for departure by reading the taxi-out checklist. The cabin crew reports the cabin is secured and ready for departure. The flight crew will report their position and ready for departure when the aircraft has reached the runway where ATC Ground Control will hand the aircraft over to ATC Tower Control; Tower must then check the position of the aircraft either visually or by using the ground radar system, after which a line up approval will be given when the preceding aircraft has initiated the take-off roll.

Take-off

ATC Tower Control will issue a take-off clearance when the preceding aircraft has lifted off and no further airspace restrictions apply. When the before take-off checklist is completed, the take-off roll is commenced by the PF. All actions and call-outs during take-off and initial climb are prescribed and performed according SOPs. As the aircraft accelerates down the runway, the PNF calls out airspeed, typically at 80 kts, V1 (take-off decision speed) and VR (rotation speed). If a problem occurs before V1, the take-off will be aborted, after passing V1 the aircraft is committed to take-off. Immediately after lift-off the gear is retracted, followed later by thrust reduction, flap retraction and acceleration to climb speed. Also the autopilot will be engaged.

Climb

During flight, duties are assigned according to PF/PNF roles. The PF will monitor the flight path and adjust where required by selecting autopilot modes or changing data in the FMS. The PNF will assist the PF where required, will do the communication with ATC and perform administrative duties.
The departure is flown according the applicable SID, but sometimes a flight may have to deviate from the planned departure route because of weather or by ATC request. As the aircraft climbs to the assigned cruising altitude, and during cruise flight, the flight crew will have to switch to different ATC centres. The flight proceeds to its destination by following assigned airways. Navigation is predominantly done using the FMS that obtains its information from a range of different sources (air data measurements, fuel calculation, GPS and radio position information, aircraft system status).
An on-board weather radar provides information on actual cumulative weather

conditions ahead of the aircraft, useable range of the weather radar is approximately 100 NM or 15 minutes flying time ahead.

Cruise/descent

During the flight updated actual weather observations and forecasts of key airports (destination, en-route alternates) can be obtained directly through datalink, via dedicated radio transmissions or as request to ATC. Also the flight crew can obtain any information as required from the Dispatcher who may be reached via datalink, radio (when in reach) or satellite phone (if installed on the aircraft). With the information available the flight crew can continuously update contingency planning for an en-route diversion or emergency.
The flight is normally flown along the airways that were planned in the pre-flight phase, but ATC may issue short-cuts where appropriate. The FMS will continuously calculate important parameters such as distance to/time over/fuel remaining at the waypoints that make up the flight plan and will calculate optimum speed and altitude. The cockpit crew will negotiate the flight profile with ATC, when able the optimal altitude- and speed schedule will be flown, subject to traffic restrictions.
The flight crew will start preliminary landing preparations approximately 45 minutes prior to landing. When in range the destination airport ATIS is copied. Landing calculations are made using the latest weather data, expected landing runway, estimated landing weight and planned landing configuration. Navigation systems (FMS programming, radio beacon pre-tuning) and flight deck preparation for descent, approach and landing are performed well in time and correct settings verified according to the descent checklist. The company handling agent may be contacted on a dedicated radio frequency to obtain the gate assignment and discuss items such as special handling requests for passengers or cargo and discussion of possible maintenance actions upon arrival. The PF will give the crew briefing on expected arrival route and approach type, cockpit settings, aircraft landing configuration, weather conditions, applicable NOTAMS, expected taxi-route after landing to the gate, instructions for a possible missed approach and other relevant issues such as high terrain during arrival, approach and missed approach, related safe altitudes and highlighting applicable SOPs.
The descent and arrival may be conducted according to a Standard Arrival Route (STAR) until radar vectoring to final approach (on the extended centreline of the landing runway) is applied by ATC Arrival or Approach Control. During descent and arrival speed is reduced from cruise speed to final approach speed according to the optimal schedule, as prescribed by the STAR or by ATC. Approximately 10 minutes before landing the cabin crew is alerted to prepare the cabin for landing and the approach checklist is read.

Approach/landing

Weather permitting, non-precision or CAT I ILS approached can be flown manually following the Flight Director (FD) or with the autopilot (AP) engaged. The PF is either hand flying the aircraft and following the FD or is monitoring autopilot operation, being ready to take over if required. Meanwhile, the PNF scans alternatively inside and outside the cockpit announcing flight parameter deviations and standard call-outs according to SOPs. CAT II automatic approaches can be followed by a manual landing, although usually SOPs will prescribe an automatic landing with the autopilot engaged. In CAT III weather conditions an automatic landing is mandatory.

When established on final approach ATC Approach Control will hand over to ATC Tower Control who will provide the latest weather and runways conditions update. When the preceding aircraft has vacated the landing runway Tower Control will issue a landing clearance, this may be only the case just seconds prior to anticipated touchdown.

On final approach the aircraft is configured for landing: flaps are extended in steps, gear is extended and speed is reduced to Final Approach Speed (FAS), which is a function of aircraft weight, aircraft configuration and weather conditions. Before reaching the applicable stabilisation height (usually 1000 or 500 ft, depending on weather conditions), the aircraft must be on the correct lateral and vertical flight path (based on radio navigation guidance or visual reference), the aircraft must be in the desired landing configuration and at the desired speed (FAS) and the landing checklist has been accomplished.

During the approach, at any time the cockpit crew may abandon the approach if the visibility is below the required minimums, criteria for stabilised approach are not achieved, or when the flight crew is not convinced that safe landing is possible for whatever reason. When a missed approach is initiated, the thrust is advanced, gear and flaps are retracted. The missed approach path is followed using FMS or radio navigation or additional instructions may be obtained from ATC.

After touchdown, the following braking devices are used to decelerate the aircraft and bring it to a complete stop: ground spoilers, wheel brakes (including anti-skid and autobrake systems) and the thrust reverser system. Typically at 80 kts the thrust reverse levers are returned to the reverse idle position and then to the stow position when reaching taxi speed. ATC Ground Control will provide a taxi clearance from the runway to the gate or parking spot on the apron. When the engines are shut down the doors may be opened and cargo and passengers are off-loaded. During disembarkation the cockpit crew shuts down the different aircraft systems and completed administrative duties. If on an outstation and the same crew will perform the return flight, initial preparations may be started for the return flight.

Colofon

mei 2008

Causal Model for Air Transport Safety is een uitgave van het
Ministerie van Verkeer en Waterstaat,
Directoraat-Generaal Luchtvaart en Maritieme zaken

Vormgeving:	Mijs Cartografie en Vormgeving Rotterdam
Bestelnummer:	ISBN 10: 90 369 1724-7 ISBN 13: 978 90 369 1724-7
Bestellen:	Ministerie van Verkeer en Waterstaat, afdeling publieksvoorlichting telefoon 070-351 7086