Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics

Aviation Risks with continuous/discrete non parametric BBNs

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

for the degree

MASTER OF SCIENCE
in
APPLIED MATHEMATICS

by

KATARZYNA AGATA KRUGŁA

Delft, the Netherlands
July 2008

MSc THESIS APPLIED MATHEMATICS

"Aviation Risks with continuous/discrete non parametric BBN"

KATARZYNA AGATA KRUGŁA

Delft University of Technology

Daily supervisor

Dr. Dorota Kurowicka

Responsible professor

Prof. dr. Roger M. Cooke

Committee members

Prof. dr. Roger M. Cooke

Dr. Dorota Kurowicka

Prof. dr. Ben J.M. Ale

Ir. Oswaldo Morales-Nápoles

July 2008

Delft, the Netherlands

To my family and friends

Contents

Contents i

1 Introduction 1
1.1 Goal of the thesis 2
1.2 Outline if the thesis 2

2 Methodology 5
2.1 Bayesian Belief Nets (BBNs) 5
2.1.1 Discrete BBNs 6
2.1.2 Normal BBNs 6
2.1.3 Non parametric BBNs 7
2.2 Expert Judgment 9
2.2.1 Marginal distribution 9
2.2.2 Elicitation of unconditional and conditional rank correlations 10
2.3 Sensitivity Analysis 14

3 Human Reliability Models 19
3.1 Flight Crew Performance Model 20
3.1.1 Description of the model 20
3.1.2 Analysis of FC performance model 22
3.2 Air Traffic Control Performance Model 27
3.2.1 Description of the model 27
3.2.2 Analysis of ATC Performance Model 28
3.3 The Total Transmission Time node 32
3.4 Maintenance Technician Performance Model 34
3.4.1 Building the model 34
3.4.2 Model variables 36
3.4.3 Quantification of the model 41
3.4.4 Analysis of the Maintenance Technician Performance Model . 48
3.4.5 Sensitivity Analysis of Maintenance the Technician Performance Model 50
3.5 Final Remarks 51

4 Causal Model for Air Transport Safety **55**
4.1 Introduction to CATS model 55
4.1.1 Event Sequence Diagram (ESD) 56
4.1.2 Fault Tree (FT) 60
4.1.3 Human reliability models in CATS model 62
4.2 Development of CATS model 63
4.3 Procedure to obtain the underlying error distribution 74

5 Analysis of the CATS BBN **79**
5.1 Conditional and Unconditional distribution of accident probability . . 79
5.2 Sensitivity analysis of MT performance on accident probability 85
5.3 Sensitivity analysis of human performance models on TO accident probability 88
5.4 Sensitivity analysis of base events representing FC, ATC, MT and non-human errors on accident probability 90
5.5 Sensitivity analysis of base events representing FC, ATC, MT and non-human errors on different accident scenarios probability 93

6 Summary and conclusions **97**

Bibliography **103**

A Excitation protocol **105**

List of Figures **115**

List of Tables **118**

Chapter 1

Introduction

This thesis has been evaluated in the context of the CATS (Causal Model for Air Transport Safety) project. CATS is a project commissioned by the Dutch Ministry of Water Management and Transport to different European organizations to investigate risks in the aviation industry [17]. The full model is being developed by a consortium including the Delft University of Technology (TUD), Det Norske Veritas (DNV), the National Aerospace Laboratory (NLR) and White Queen (WQ). The purpose of this model is to describe and understand better the air traffic system and its safety functions in such a way that it is possible to analyze the risk reduction alternatives and that serves as means of communication between experts and managers within the industry.

The aviation industry is considered as one of the safest ways of traveling nowadays. However, accidents and incidents may still occur. These accidents tend to result from a combination of different casual factors (e.g. human errors, technical failures, environmental and management influences) in certain accident scenarios (e.g. loss of control, fire, collision, etc.) and the causes and consequences differ according to the phase of flight in which they may occur (e.g. take-off, en route approach and landing, etc.). As it was already mentioned, human errors may have big influence on the accidents and incidents scenarios. Therefore, it is really important to define those human factors in a proper way and understand the associated risks of those factors. This effort may be used later on to improve the safety as effectively as possible. For this reason, 3 human reliability models were developed - Flight Crew performance model, Air Traffic performance model and Maintenance Technician performance model whose evaluation is a part of this master thesis.

The CATS project approaches the complexity by evaluating separate causal models for each accident scenario in each flight phase. This are represented by Event Sequence Diagrams (ESDs), and Fault Trees (FTs) and all of these separate elements are converted later on to a single structure - Bayesian Belief Net (BBN). In practice, to build and analyze the CATS model a software package UniNet is used. UniNet is being developed at the Delft Institute of Applied Mathematics of the Delft University of Technology for dealing with large continuous/discrete non parametric BBNs. UniNet is able to update a joint distribution (assuming a normal copula) in real time. Inputs for UniNet are marginal distributions, and rank and conditional rank correlations as described in [1]. A description of some features of the program may be found in [4].

1.1 Goal of the thesis

The goal of the thesis is to:

- develop and build the Maintenance Technician performance model.
- explain the structure of the CATS model giving an overview of its different parts.
- describe different steps required for constructing the full model.
- present analysis with the version of the model that is available at the moment of writing this thesis.
- recommend possible directions for exploiting research resources in the future.

The model as stands at the present is not a final version. In particular 2 event sequence diagrams ESD36 - Ground collision imminent and ESD37 - Wake vortex encounter will be added to the model. Also the distribution of some human errors (base events in the FTs) might change as new data becomes available. More details will be presented along the thesis.

1.2 Outline if the thesis

This thesis has been structured as follows. In this first chapter we present an introduction to the subject of study and an overview of the structure of the thesis. In Chapter 2 we will present the theoretical "tools" used later on in the thesis, like a description of different types of the Bayesian Belief Nets and the elicitation procedures for marginal distributions and the (conditional) rank correlations.

Chapter 3 will present the 3 human reliability models. We will give a brief description and analysis of the Flight Crew (FC) performance model and Air Traffic Control (ATC) performance model. Next, in more details we will present the Maintenance Technician (MT) performance model. The description of the development of the model, the most influential contribution factors on the MT error, the quantification of that model and as well as analysis using software UniNet will be presented.

In Chapter 4 we will focus on the CATS model. We will present a brief description of the Event Sequence Diagrams (ESDs) and Fault Trees (FTs). Next, we will present the progress of developing the CATS model. This chapter will finish with the description of the procedure to obtain the underlying error distribution.

In Chapter 5 we will present the analysis of the CATS BBN with respect to MT performance model and accident scenario representing three flight phases, takeoff, en route and approach and landing. We will present also the sensitivity analysis.

Finally some conclusions and recommendations for the future work will be presented in the Chapter 6.

Chapter 2

Methodology

This chapter provides the main information about the theoretical "tools" used later on in this thesis. We present a brief description of Bayesian Belief Nets models with emphasis on non parametric BBNs. Also, the elicitation procedures for marginal distributions and the (conditional) rank correlations will be presented. We will finish this chapter presenting the basic ideas of the sensitivity analysis measures that will be used later in chapters 3 and 5.

2.1 Bayesian Belief Nets (BBNs)

Bayesian Belief Nets (BBNs) (also called Bayesian Networks) are graphical tools used to represent and model high-dimensional uncertainty distributions.

A BBN is a directed acyclic graph. The nodes of a BBN represent random variables, which can be either discrete or continuous, and the causal arcs represent relationships between them.

Figure 2.1 presents an example of a Bayesian Belief Net on four variables. Variable 4 is called child node and variables 1,2 and 3 are called parent nodes.

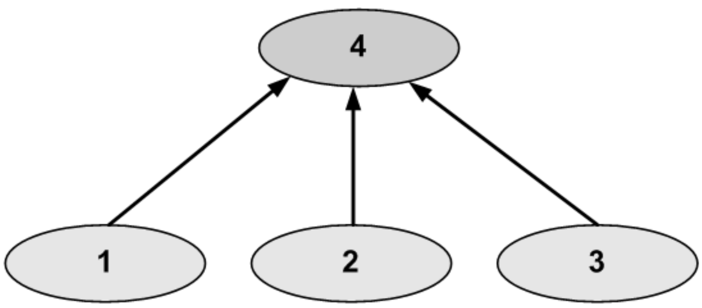


Figure 2.1: BBN on 4 variables.

In order to specify the BBN form figure 2.1 one needs to specify the marginal distributions of 1,2 and 3 and the conditional distribution of 4 given 1,2,and 3. In general, to specify a BBN, the marginal distributions of the variables without parents (source nodes) are needed, and the conditional distributions of each other variable given its parents in the graph.

In the literature, 3 types of BBNs are discussed: discrete, normal and non parametric, [1, 2]. Each of the mentioned types will be shortly presented, along with its advantages and disadvantages.

2.1.1 Discrete BBNs

The nodes of a discrete BBN represent discrete varaibles. There are some advantages and disadvantages of this kind of BBNs.

A discrete BBN requires the specification of the marginal distributions for the source nodes and conditional probability tables for all child nodes. Discrete BBNs have become a very popular tool in modeling risk and reliability. If the variables involved in the analysis are truly discrete, and data are available for calculating the conditional probability tables needed, there are advantages in working with this kind of models. One of them is that fast updating algorithms are available. Moreover, commercial tools with an advanced graphical interface that support discrete BBNs construction and inference are also available. Nevertheless if the models are large and complex, the discrete BBNs suffer the disadvantage of a high assessment burden. For example, if a child node has 7 parents and all nodes have 3 states, then 6561 (3^8) conditional probabilities are required [1]. Moreover, if data is available for the marginal distribution of a child node and not available for its conditional distribution, the latter information should be gathered from experts. Specifying the conditional distribution such that it complies with the marginal retrieved from data can be a difficult task for any expert. If the variables involved in the model are continuous, modeling their dependence with discrete BBNs requires a simplification of the model or a drastic discretization of the nodes. For these reasons, working with discrete BBNs is sometimes inappropriate. In these cases the continuous approach seems more suitable.

2.1.2 Normal BBNs

Continuous BBNs were first developed for the joint normal variables. 'Influences' of the parents on a child were interpreted as partial regression coefficients, when the child is regressed on the parents [2]. For each normal variable, the unconditional mean and constant conditional variance must be assessed together with a set of partial regression coefficients [1]. Discrete nodes can be also incorporated and modeled together with normal nodes (discrete-normal BBNs). However, there is the restriction that continuous nodes can have discrete parents but not discrete children. One of the advantages of normal BBNs is that the means, conditional variances and partial regression coefficients can be assessed algebraically independent. As in the case of discrete BBNs, this approach is suitable if the variables are truly normal.

If the normality assumption does not hold, then the individual variables must be transformed to normal variables. The conditional variance in normal units must be constant. Moreover, the partial regression coefficients apply to the normal units of the transformed variables, not to the original units, which places a heavy burden of any expert elicitation[1]. Also, if the parent node is added or removed after quantification, then the previously assessed partial regression coefficients must be re-assessed. Hence, if the normality assumption does not hold, all mentioned requirements make a normal BBNs unappealing for modeling high dimensional distributions.

2.1.3 Non parametric BBNs

In [2], a non parametric approach was proposed for the continuous BBNs. No parametric form for the joint distribution is assumed. In order to quantify BBNs using these approach, one needs to specify all one-dimensional distributions and a number of (conditional) rank correlations equal to the number of arcs in the BBN.

Each node of the graph is assigned with a continuous invertible univariate distribution. The dependence between variables is described via (conditional) rank correlations. This measure of dependence was chosen for several reasons: it always exists; does not depend on marginal distributions; measures monotonic relationships and successful expert elicitation procedures have been developed for rank correlation. The (conditional) rank correlations on a BBN are algebraically independent. Any number between [-1,1] can be attached to the arcs of a continuous non parametric BBNs. For example, for the BBN structure from figure 1.1 one needs to specify (conditional) rank correlations between variable 4 and its parents. A (non unique) order of the parents should be chosen. Let us assume we choose the order 1,2 and 3. The rank correlation between 4 and its first parent is the unconditional rank correlation $r_{4,1}$. The rank correlations between 4 and its next parent 2 will be conditioned on the value of the previous parent 1, i.e. $r_{4,2|1}$. In the same manner the arc between 4 and 3 is associated with the conditional rank correlation $r_{4,3|2,1}$.

Using these quantification nodes and arcs can be added or delated from a BBN without re-assessing previously specified correlations. Furthermore, the dependence structure is significant for any such quantification and there is no need to revise it if univariate distributions are changed. When data are not available the conditional rank correlations are elicited from experts. The elicitation procedure is presented

[1]When data is not available.

in the next sub-section.

The theory for non parametric continuous BBNs is extended in [18] to include discrete ordinal variables. These are variables that can be written as monotone transforms of uniform variables. The rank correlation of two discrete variables is translated in terms of the rank correlation of their underlying uniforms.

The (conditional) rank correlations assigned to the edges of a BBN are realized using copulae. Any copula with invertible conditional cumulative distribution function, that realizes all correlations between -1 and 1, may be used as long as it represents (conditional) independence as zero (conditional) correlation. Nevertheless we choose the normal copula for realizing the dependence structure. The biggest advantage of the normal copula is the conditioning/updating can be done analytically.

The present discussion concerns only probabilistic nodes. However, a node in the BBN can be also functional, i.e. a function of other variables. This function captures all the dependence between the parents (arguments of the function) and their child. Therefore, there is no need to assess the (conditional) rank correlations to the arcs connecting functional nodes with their parents. It is worth mentioning that the influence defined by the (conditional) rank correlation can be considered as "softer" than the influence determined by the functional relationship.

A restriction in using functional nodes is that a functional node can not have probabilistic children.

In order to fully quantify the model we need to obtain the marginal distributions together with (un)conditional rank correlations specified by the edges. When data are not available, we can use expert judgment for that purpose.

2.2 Expert Judgment

Expert opinion has been used in many fields. If some information about parameters are not known and cannot be estimated from any experiments or observations expert judgment is used instead.

In this project expert judgment is used for two purposes:

- to provide information about unknown marginal distribution for variables of interests in the Maintenance Technician performance model;
- to obtain information about the strength of the dependence relationships between variables in the Maintenance Technician performance model.

In order to estimate the marginal distributions we used the so-called classical model of expert judgment[2].

2.2.1 Marginal distribution

In order to obtain marginal distribution the classical method of expert judgment is used. The classical model constructs the weighted combination of expert' probability assessments. Experts are asked to provide their subjective probability distribution in the form of a number of specified quantiles. Most of the times, an expert is asked to specify his/her 5%, 50% and 95% quantile of uncertainty distribution for each of the variables of interest. These are defined as follows:

- The 5% percentile value means that there is 5% chance that realization of the variable is lower than this value.
- The 50% percentile (median) of the distribution, i.e. the value that has 50% chance of observing values higher or lower.
- The 95% percentile value means that there is 5% chance that realization of the variable is higher than this value.

Weights are derived based on some performance measures and satisfy the strictly proper scoring rule. There exist two quantitative measures of the performance of the experts - calibration and information (or informativeness). They are assessed based on experts' estimates on so-called seed or calibration variables. These are variables from the experts field of experience and their true values are unknown to the experts when they give their opinions, but known to the analyst. Calibration measures the statistical likelihood that a set of experimental results correspond, in a statistical sense, with the experts assessments. Information represents the degree to which the distribution provided by an expert is concentrated. For more details, see [7].

[2]In the elicitation only one expert was used

2.2.2 Elicitation of unconditional and conditional rank correlations

In this section we explain the methods used in the elicitation of unconditional and conditional rank correlations. The BNN in figure 2.2 will be used as an example.

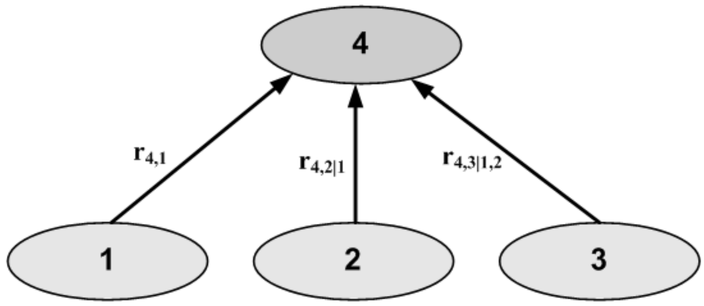


Figure 2.2: A BBN on 4 variables with associated set of the (conditional) rank correlations.

First, we present the conditional probability method for estimating rank correlations and next, we present how to extend the elicitation procedure from unconditional to conditional rank correlations.

Conditional probabilities of exceedance and rank correlations

To elicit the rank correlation $r_{4,1}$ between variables X_4 and X_1, for the BBN presented in figure 2.2, we ask expert the following question:

- *Suppose that the variable X_1 was observed above its q^{th} quantile. What is the probability that also X_4 will be observed above its q^{th} quantile?*

An answer to that question is equivalent to an estimate of $P_1 = P(F_{X_4}(X_4) > q|F_{X_1}(X_1) > q)$. In practise we can use any quantile value. However, the resulting rank correlation is dependent of the choice of copula[3].

To calculate the exceedence probability we can integrate numerically the bivariate normal density $\phi(x_1, x_4, \rho_{4,1})$ over the region corresponding to the quantile's exceedance region $[\Phi^{-1}(q), \infty)^2$, where Φ^{-1} is the inverse standard normal cumulative distribution function.

$$P_1 = \frac{1}{1-q} \int_{\Phi^{-1}(q)}^{\infty} \int_{\Phi^{-1}(q)}^{\infty} \phi(x_1, x_4, \rho_{4,1}) dx_1 dx_4. \tag{2.1}$$

[3]The normal copula is desired because the choice of other copulas imposes computational restrictions in terms of speed and numerical accuracy. For more information see[12].

After that, the analyst finds the $\rho_{4,1}$ which satisfies the expert's conditional probability assessment and transforms this to the corresponding rank correlation using the relationship in equation 2.2[4]

$$\rho_{4,1} = 2sin(\frac{\pi}{6} r_{4,1}). \tag{2.2}$$

The relationship between the conditional probability (equation 2.1) and the rank correlation $r_{4,1}$ is presented in figure 2.3. In figure 2.3 we use the 50^{th} percentile while eliciting exceedence probabilities and the normal copula to find relationship between probability of exceedence and the rank correlations.

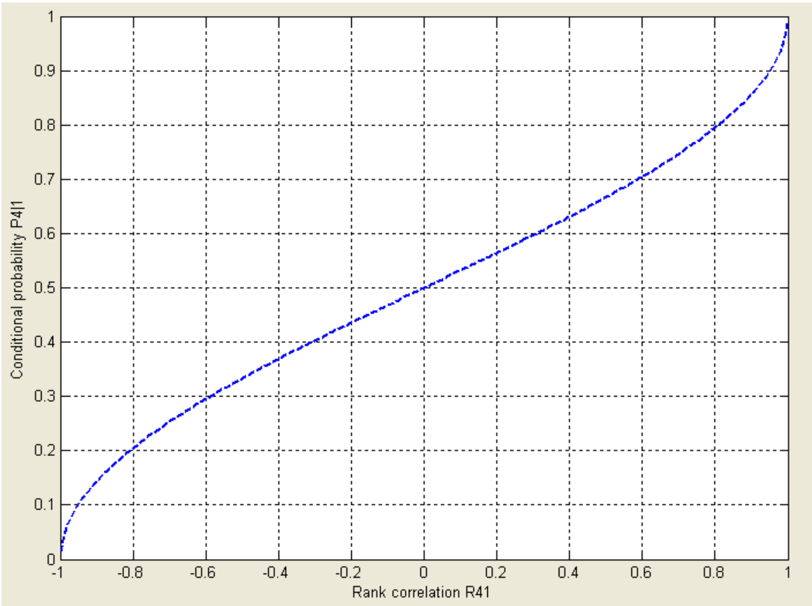


Figure 2.3: Relationship between $P(F_{X_4}(X_4) \geqslant 0.5|F_{X_1}(X_1) \geqslant 0.5)$ and the rank correlation $r_{4,1}$.

When we know the value of P_1 the corresponding value of rank correlation $r_{4,1}$ can be read from figure 2.3. It is worth mentioning that:

- if P_1=0 then $r_{4,1} = -1$
- if P_1=0.5 then $r_{4,1} = 0$
- if P_1=1 then $r_{4,1} = 1$

For example, if expert tells us that $P_1 = P(F_{X_4}(X_4) > q|F_{X_1}(X_1) > q) = 0.7$ then form figure 2.3 we can read that $r_{4,1} = 0.57$.

Next, we will present a procedure of eliciting conditional rank correlations. We will continue with an example of BBN presented in figure 2.2.

[4]If (X, Y) is a random vector with joint normal distribution then $\rho_{X,Y} = 2sin(\frac{\pi}{6} r_{X,Y})$. Pearson, 1904 in [2]

Conditional probabilities of exceedance and conditional rank correlations

As it was already mentioned we will use q=0.5 while eliciting exceedance probabilities and the normal copula to find relationship between probability of exceedance and the (conditional) rank correlations. To assess the conditional rank correlation $r_{4,2|1}$, we ask the expert the following question:

- *Suppose that not only variable X_1 but also X_2 were observed above their medians. What is now the probability that also X_4 will be observed above its median value?*

This question requires expert's estimate of $P_2 = P(F_{X_4}(X_4) > 0.5|F_{X_1}(X_1) > 0.5, F_{X_2}(X_2) > 0.5)$. The probability that the expert can provide in this situation will depend on the estimate given for the question presented in previous subsection. For example, if expert thinks that variables X_2 and X_4 are independent given X_1, then the answer to second question is identical to the answer to first question. If the expert regards variables X_1 and X_4 as completely positively (negatively) correlated then he/she would have answered $P_1 = 1$ ($P1 = 0$) and the second question would not have been necessary at all, as X_4 would be completely explained by X_1. Any answer for P_1 different than 0, 0.5 or 1 means that the expert believes that X_1 explains at least in part X_4 and hence X_2 can only explain part of the dependence that was not explained already by X_1.

In the BBN presented in figure 2.2, variables X_1 and X_2 are independent, therefore the rank correlation ($r_{1,2}$) between them is equal to 0. Since all rank correlations in BBN presented in figure 2.2 are algebraically independent, then $r_{4,2|1}$ can take any value in (-1,1). The correlation matrix of the joint normal distribution looks as follows:

$$\Sigma_{4,1,2} = \begin{pmatrix} \rho_{4,4} & \rho_{4,1} & \rho_{4,2} \\ \rho_{4,1} & \rho_{1,1} & \rho_{1,2} \\ \rho_{4,2} & \rho_{1,2} & \rho_{2,2} \end{pmatrix} = \begin{pmatrix} 1 & \rho_{4,1} & \rho_{4,2} \\ \rho_{4,1} & 1 & 0 \\ \rho_{4,2} & 0 & 1 \end{pmatrix}$$

The value of correlation $\rho_{4,1}$ was assessed by expert in previous step. We can determine the relationship between $\rho_{4,2|1}$ and the conditional probability P_2 from useful properties of normal copula. If we know the values of $\rho_{4,1}$, $\rho_{4,2}$ and $\rho_{1,2}$ we can calculate partial correlation $\rho_{4,2;1}$ with the recursive formula [2]:

$$\rho_{4,2;1} = \frac{\rho_{4,2} - \rho_{4,1} \cdot \rho_{1,2}}{\sqrt{1 - \rho_{4,1}^2}\sqrt{1 - \rho_{2,1}^2}} \tag{2.3}$$

As for joint normal distribution partial and conditional correlations are equal, therefore $\rho_{4,2;1} = \rho_{4,2|1}$. Now, using the Pearson transformation ($\rho_{4,2|1} = 2sin(\frac{\pi}{6}r_{4,2|1})$) we can obtain $r_{4,2|1}$.

Relationship between $\rho_{4,2|1}$ and probability obtained from expert is as follows

$$P_2 = \frac{1}{0.5}\frac{1}{0.5}\int_0^{\infty}\int_0^{\infty}\int_0^{\infty}\phi(x_4, x_1, x_2, \rho_{4,1}, \rho_{4,2|1})dx_4dx_1dx_2. \tag{2.4}$$

Figure 2.4 presents the the relationship between conditional probability P_2 and the conditional rank correlation $r_{4,2|1}$, where the previous expert's estimate for P_1 was equal to 0.7. One can notice that the probability of exceedance is constrained by the expert's previous estimate. If $P_1 = 0.7$ then the possible values for probability P_2 are restricted to interval (0.4,1). This can be explained in the following way. While expert provides his estimate for P_1 he/she describes how much of the information of X_4 is explained by X_1. Thus, P_2 can only explain the remaining part of information of X_4 which was not explained by X_1.

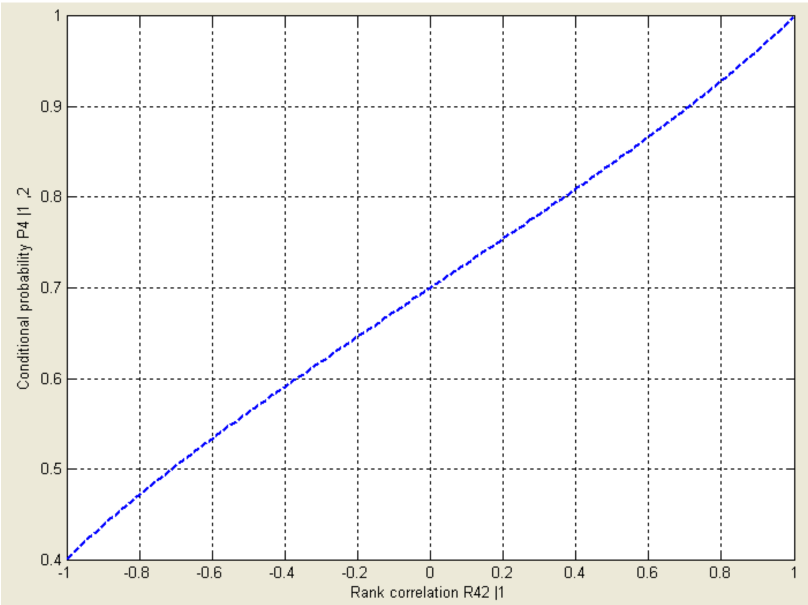


Figure 2.4: Relationship between $P(F_{X_4}(X_4) \geqslant 0.5|F_{X_1}(X_1) \geqslant 0.5, F_{X_2}(X_2) \geqslant 0.5)$ and the rank correlation $r_{4,2|1}$.

For example, if expert provides that $P_2 = 0.6$ then the corresponding conditional rank correlation $r_{4,2|1}$ is equal to -0.37.

A similar procedure would be applied to find other higher order conditional rank correlations in a given BBN.

Different approaches may be used to elicit the unconditional and conditional rank correlations. We present here 2 possible elicitation procedures additional to the one presented so far.

First, for the BBN presented in figure 2.2, the expert is asked:

- *What is the probability that variable 4 is above its q_4^{th} quantile given that 1 is above its q_1^{th} quantile?.*

This information is translated into the rank correlation $r_{4,1}$. As it was already mentioned q_4 and q_1 may be different, however the usual choice is the median for both of them. This question is exactly the same as in procedure 1.

We can ask similar question for the rest of the variables. The second question may be:

- *What is the probability that variable 4 is above its q_4^{th} quantile given that 2 is above its q_2^{th} quantile?* From this exeedence probability one can calculate $r_{4,2|1}$.
- *What is the probability that variable 4 is above its q_4^{th} quantile given that 3 is above its q_3^{th} quantile?* The answer to this question is translated in $r_{4,3|1,2}$.

A third option is to elicit directly a rank correlation as opposed to probabilistic statements. Once the first conditional probability has been assessed and translated to its corresponding rank correlation ($r_{4,1}$ in this case), the analyst may elicit ratios of rank correlations and translate them into conditional rank correlations. The first (unconditional) rank correlation, $r_{4,1}$, follows from eliciting the corresponding probability of exceedence[5]. Later on, the ratio $\frac{r_{4,2}}{r_{4,1}}$ is elicited and translated into the conditional rank correlation $r_{4,2|1}$, followed by the elicitation of $\frac{r_{4,3}}{r_{4,1}}$, which is translated into $r_{4,3|1,2}$.

The answer to each of the subsequent questions is constrained by the answers provided by the previous questions. Therefore at each step of the elicitation, bounds for the rank correlations have to be computed. The experts' assessments will depend on these bounds which are derived from his previous answers.

2.3 Sensitivity Analysis

According to [2], sensitivity analysis is concerned with identifying "important parameters". In general, we would like to know how much input parameters (base variables) influence output data (predicted variables). Sometimes if the analyst considers some dependence to be not strong enough he/she might decide to leave variables outside the model to make the model more simple. Such kind of decisions might be taken on the basis of sensitivity analysis.

To carry out the sensitivity analysis, several statistical and sensitivity measures are calculated using the sensitivity analysis software Unisens which is a satellite program of UniNet. For example, during the sensitivity analysis we can compute

[5]Hence in the three methods the first question would be the same.

the following statistical and sensitivity measures: product moment correlation, rank correlation, regression coefficient, correlation ratio, linearity index, partial correlation, partial regression coefficient, and etc.[6]

However, in our analysis we will only use 3 statistics: product moment correlation, rank correlation and correlation ratio. Below, we present a short description and definitions of these statistics.

Product moment correlation

If random variables X and Y have finite expectation and finite variance, then the product moment correlation is defined as follows:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}, \tag{2.5}$$

where $Cov(X,Y) = E(XY) - E(X)E(Y)$.

The product moment correlation takes values between -1 and 1 and it measures the degree of linear relationship between X and Y:

- $\rho(X,Y) = 1$ if and only if $Y = aX + b$ for some $a > 0$,
- $\rho(X,Y) = -1$ if and only if $Y = aX + b$ for some $a < 0$,
- if X and Y are independent then $\rho(X,Y) = 0$ (although the converse is not always true).

Rank Correlation

For example, when two variables X and Y are functionally related, but in a non-linear way, then the product moment correlation may be small. For this reason the rank correlation is often used as a measure of the degree of monotone relationship (i.e. rank correlation measures the extent to which large values of X occur with the large values of Y and small values of X occur with small values of Y).

The rank correlation of random variables X and Y with cumulative distribution function F_X and F_Y is defined as follows:

$$\rho_r(X,Y) = \rho(F_X(X), F_Y(Y)). \tag{2.6}$$

Correlation ratio

Correlation ratio (CR) is an instrument which can give insight into relations between base variables $X, Y, Z, \dots$ and predicted variable $G(X, Y, Z, \dots)$.

[6] For more information see [2, 19].

Correlation ratio of predicted variable G and base variable X is a squared product moment correlation between G and function $f(X)$ which maximizes this correlation. The function f that maximizes $\rho^2(G, f(X))$ is the conditional expectation of G given X.

$$CR(G, X) = \max_f \ \rho^2(G, f(X)) = \frac{var(E(G|X))}{var(G)}. \quad (2.7)$$

Therefore it is a ratio of the variance of the conditional expectation of G given X and the variance of G. Correlation $\rho^2(X, Y)$ is always less than or equal to correlation ratio $CR(X, Y)$.

Example

The example of the sensitivity analysis is based on Project 4.1 (Investment) from [2][7].

In the sensitivity analysis we will take into consideration the following variables:

- variable "5yrReturn" which can be understood as amount of money which will be return after 5 years when the initial capital equal to 1000,
- variable "start" which is defined as the initial capital (equal to 1000),
- variables "V1, ..., V5" which are representing the yearly interest; each of these variables is uniformly distributed on [0.05,0.15]. V1 is rank correlated 0.7 with V2, V2 is rank correlated 0.7 with V3, V3 is rank correlated 0.7 with V4 and V4 is rank correlated 0.7 with V5. The dependence structure is induced by a dependence tree [2].

$$5yrReturn = start * (1 + V1) * (1 + V2) * (1 + V3) * (1 + V4) * (1 + V5) \quad (2.8)$$

Now, we would like to check if there is any dependence between predicted variable "5yrReturn" and base variables "V1, ..., V5". In this sensitivity analysis we drawn 2000 samples. Figure 2.5 presents the results the results of this analysis.

[7] This example does not have any connection with this master project. It is only used here to explained the sensitivity analysis.

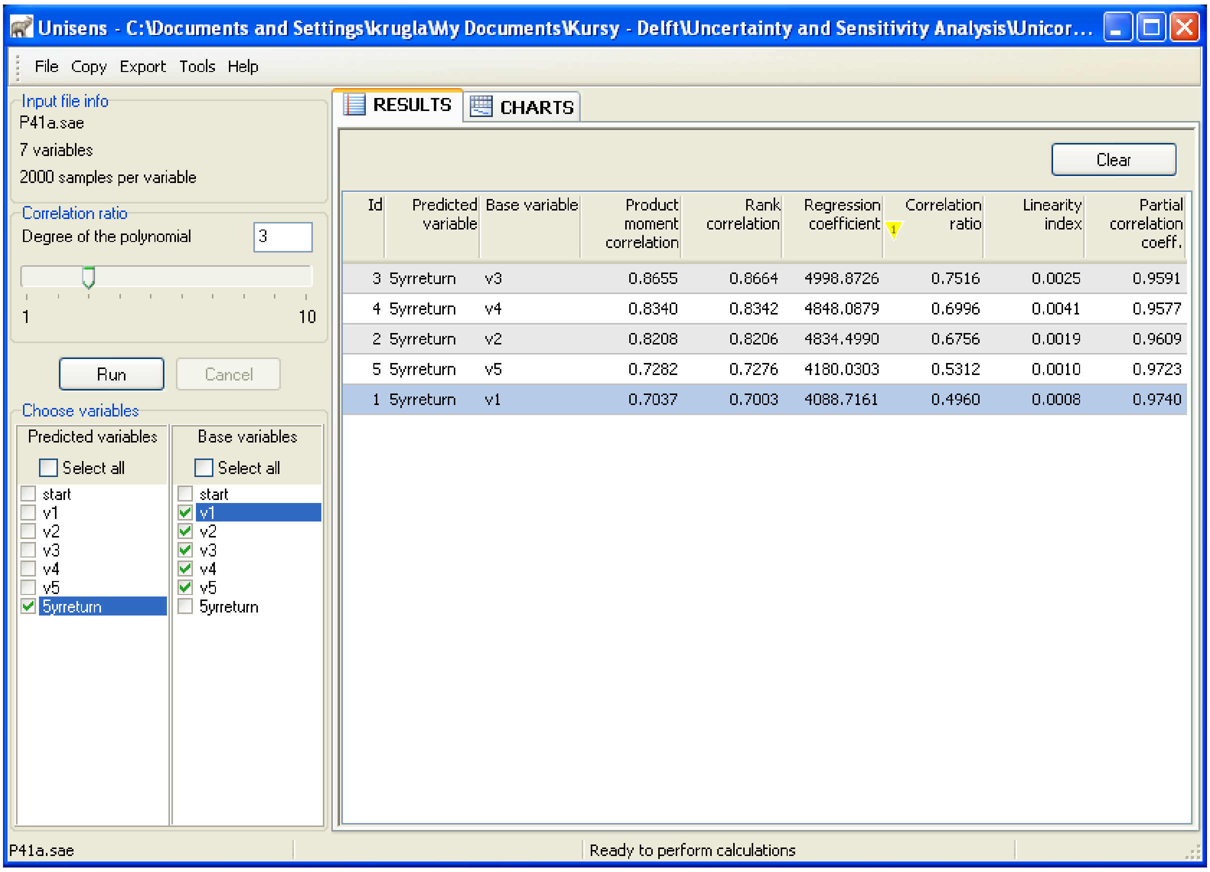


Figure 2.5: Example of the sensitivity analysis.

In figure 2.5 we can see different measures which can be used in the sensitivity analysis, e.g. product moment correlation, rank correlation, regression coefficient, correlation ratio, linearity index and partial correlation coefficient. During the analysis we will focus mostly on the product moment correlation, the rank correlation and the correlation ratio.

As we can observe, the variable V3 (the interest after third year) has the biggest influence on predicted variable “5yrReturn”. The correlation ratio is equal to 0.7516. It means that the variance of the “5yrReturn” is explained by 75.2% of the variance of the conditional expectation of “5yrReturn” given V3. The next variable which has big influence on the “5yrReturn” is the interest after fourth year (V4), whereas the variable V1 has the lowest influence on the predicted variable “5yrReturn”.

The ratio of the highest to lowest correlation ratio is equal to 1.5153. The difference between the product moment correlations and rank correlations do not vary significantly. These measures are between 0.7 and 0.8 approximately indicating a relatively high degree of linear dependence. This can be checked in equation 2.8 by observing that the interaction terms though still present almost vanish. For an extensive account on sensitivity analysis measures the reader is referred to [2].

Chapter 3

Human Reliability Models

According to [13], the "human" factor accounts for 56% of the fatal accidents in worldwide commercial aviation. Therefore the human factor should be properly described in any risk model that illustrates the causes of aviation accidents. The CATS model includes the human factors through human performance models, which describe the human errors that can occur in different flight phases. Two human performance models have been previously developed. These were Flight Crew (FC) and Air Traffic Control (ATC) performance models. The goal of our work is to develop a third human reliability model, the Maintenance Technician (MT) model.

As it was mentioned already, generic human performance models can be used to represent the roles of flight crew, air traffic control crew and maintenance crew in different situations. Some of the requirements of the human performance models were:

- The human performance models should focus on those human operations that influence the accident scenarios the most and be as simple as possible.
- It is not required to include every task or job to the model.
- These factors should have a simple and clear definition and operational meaning.

This last remark is essential as the quantification of the model is based either on available data or on expert opinion. The human performance is modeled with continuous-discrete non parametric Bayesian Belief Nets, which allows the representation of "probabilistic influences rather than deterministic cause-effect relationships" ([5]).

In this chapter first, we briefly present the Flight Crew (FC) and Air Traffic Control (ATC) models and describe the development of the Maintenance Technician (MT) model. Next an analysis of the model will be performed.

3.1 Flight Crew Performance Model

3.1.1 Description of the model

The Flight Crew Performance Model was the first of the generic models that have been developed to describe the influence of humans in aviation operations. Figure 3.1 presents the BBN representing the Flight Crew Performance Model. Variables taken into account in the model are described below. The names of most variables in the text appear abbreviated in the BBN. There are only two exceptions, intra-cockpit communication, which is described as language difference (LangDif) in the BBN, and man machine interface, which is described in BBN as aircraft generation (AirGen).

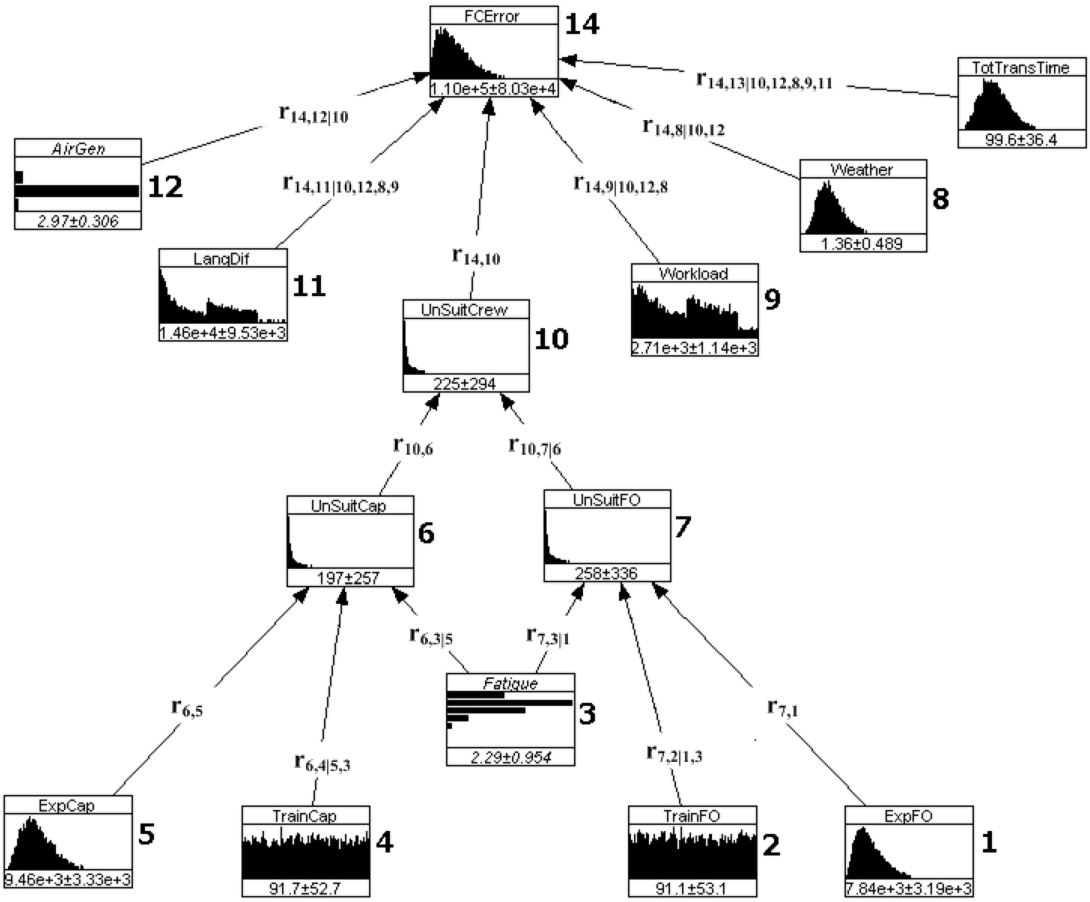


Figure 3.1: BBN representing the Flight Crew Performance Model.

The distributions of the variables are obtained from data or experts as indicated in the description. However, the distribution of flight crew error is obtained from the associated Fault Tree[1].

[1]The description and explanation of FT will be presented in more details in Chapter 4.

1. FIRST OFFICER EXPERIENCE (Data): Total number of hours flown (all type aircrafts) for the First Officer.
2. FIRST OFFICER TRAINING (Data): Number of days since the last type recurrent training for the First Officer.
3. FATIGUE (Data): Stanford Sleepiness Scale, where: 1-Completely awake and 7-sleep onset soon .
4. CAPTAIN TRAINING (Data): Number of days since the last type recurrent training for the Captain.
5. CAPTAIN EXPERIENCE (Data): Total number of hours flown (all type aircrafts) for the Captain.
6. CAPTAIN UNSUITABILITY (Structured Expert Judgment): Likelihood that the Captain fails a proficiency check.
7. FIRST OFFICER UNSUITABILITY (Structured Expert Judgment): Likelihood that the First Officer fails a proficiency check
8. WEATHER (Data): Rainfall rate in mm/hr translated to cockpit radar.
9. WORKLOAD (Structured Expert Judgment): Likelihood that the flight crew needs to follow a procedure of the “abnormal/emergency procedures” section of the Aircraft’s Operating Manual.
10. CREW UNSUITABILITY (Structured Expert Judgment): Likelihood the Captain or the First Officer fail a proficiency check.
11. INTRA-COCKPIT COMMUNICATION (Structured Expert Judgment): Difference in mother tongue between Captain and First officer; yes or no
12. MAN MACHINE INTERFACE (Data): Aircraft generation; 1, 2, 3, or 4.
13. TOTAL TRANSMISSION TIME (Expert Judgment): The total duration (in seconds) of the air/ground communications, per aircraft, for the approach and landing flight phase.
14. FLIGHT CREW ERROR (From the associated Fault Tree): Likelihood that the flight crew makes an unrecovered error that is potentially hazardous for the safety of the flight.

The nodes of the BBN in figure 3.1 show the marginal distributions of the variables listed above. The mean and the standard deviation of the distribution of each variable is shown at the bottom of each node. The marginal distributions that were not obtained from data were elicited from experts [7]. The (conditional) rank correlations were elicited according to the second method described in section 2.2.2. Five experts were interviewed and asked 23 questions in total. 11 of these questions were used to obtain the dependence information, 4 to obtain the marginal distributions and 8 of them to elicit the calibration variables. Experts answers were combined in the standard way, as described in [7]. Table 1 presents the resulting (conditional) rank correlations. For details about the FC performance model, we refer to [5].

(Un)Conditional Rank Correlations			
$r_{7,1}$	-0.95	$r_{14,10}$	0.30
$r_{7,3\|1}$	0.86	$r_{14,12\|10}$	-0.32
$r_{7,2\|1,3}$	0.24	$r_{14,8\|10,12}$	0.46
$r_{6,5}$	-0.95	$r_{14,9\|10,12,8}$	0.18
$r_{6,3\|5}$	0.86	$r_{14,11\|10,12,8,9}$	0.19
$r_{6,4\|5,3}$	0.24	$r_{14,13\|10,12,8,9,11}$ [a]	0.16
$r_{10,6}$	0.71		
$r_{10,7\|6}$	1.00		

Table 3.1: Dependence Information in the Flight Crew Performance Model.

[a] $r_{14,13|10,11,8,12,9}$ was elicited later from a single expert who is not a pilot but a risk analyst at National Aerospace Laboratory (NLR).

For instance, $r_{7,1}$ represents the dependence between "first officer unsuitability" and "first officer experience" and they are highly negatively correlated meaning that the first officer with low experience is more likely to make errors.

Next, we present the analysis and the sensitivity analysis of the FC performance model.

3.1.2 Analysis of FC performance model

Once the model is quantified we can further use it for prediction. We can conditionalize on certain values of the bottom variables (ancestors of maintenance error) and observe how this changes propagate to the top node "FC Error". Let us assume that the crew flies in "good" weather (figure 3.2). That would be translate in 0.8 mm/hour rainfall rate. We can observe that the average number of errors when flight crew is flying in good weather conditions will be reduced from 11e+4 to 6,83e+4.

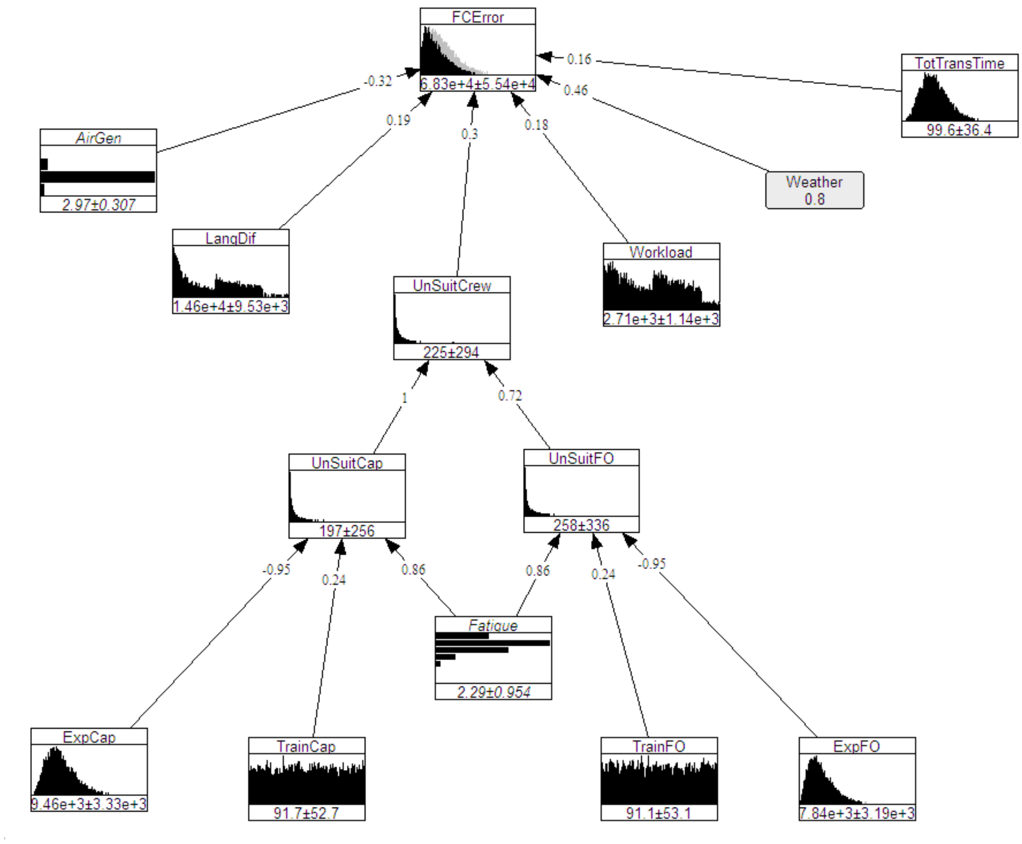


Figure 3.2: Conditional distribution of FCError|Weather=0.8.

If additionally we consider flights that not only take place in good weather, but also "high" workload (describes the number of times the crew encounters an abnormal procedure per 100,000 flights) then we can notice that the mean value for the FC error has increased to 7.67e+4 (figure 3.3). This means that high workload makes flight crew errors more probable but does not neutralize the benefit of good weather. It is still lower than the unconditional mean 11e+4.

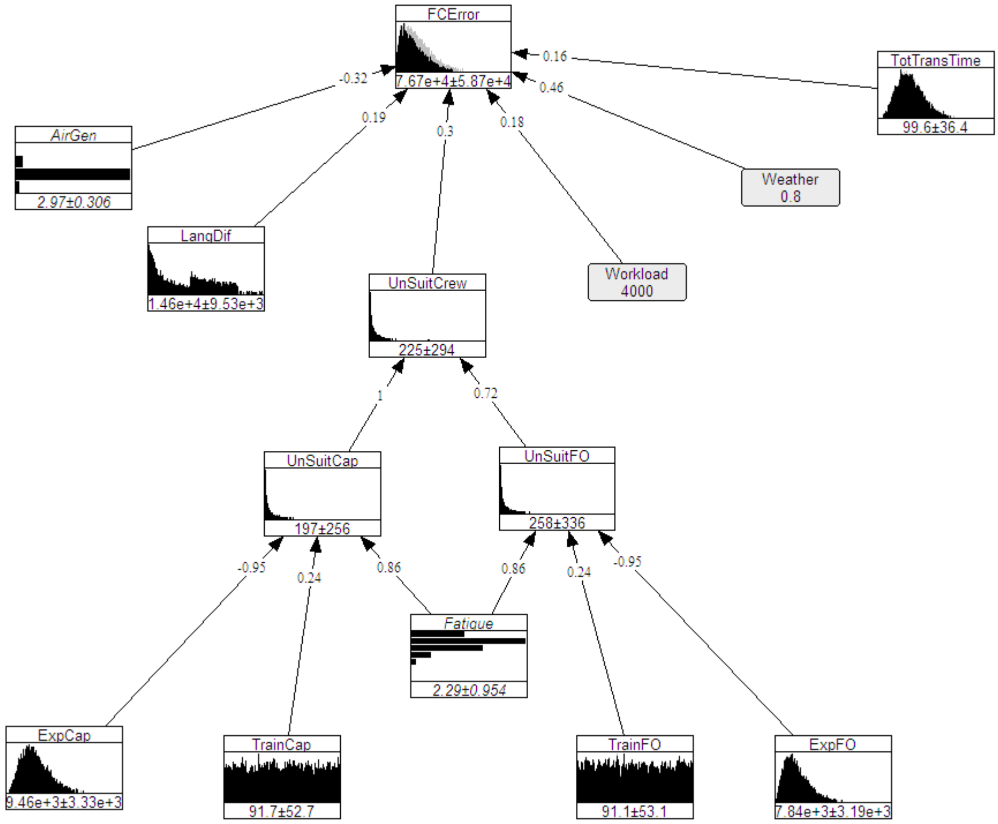


Figure 3.3: Conditional distribution of FCError|Weather=0.8, Workload=4000.

Finally, we conditionalize on the aircraft generation. Namely, we know that the aircraft was designed in the 1950s, hence the oldest design, generation 1. All previous conditioning has been kept. As we can notice, the mean value for the flight crew error has increased to 15.7e+4 (figure 3.4). This value is higher than the unconditional mean of FC error, meaning that "good" weather cannot compensate for "high" workload and an old aircraft.

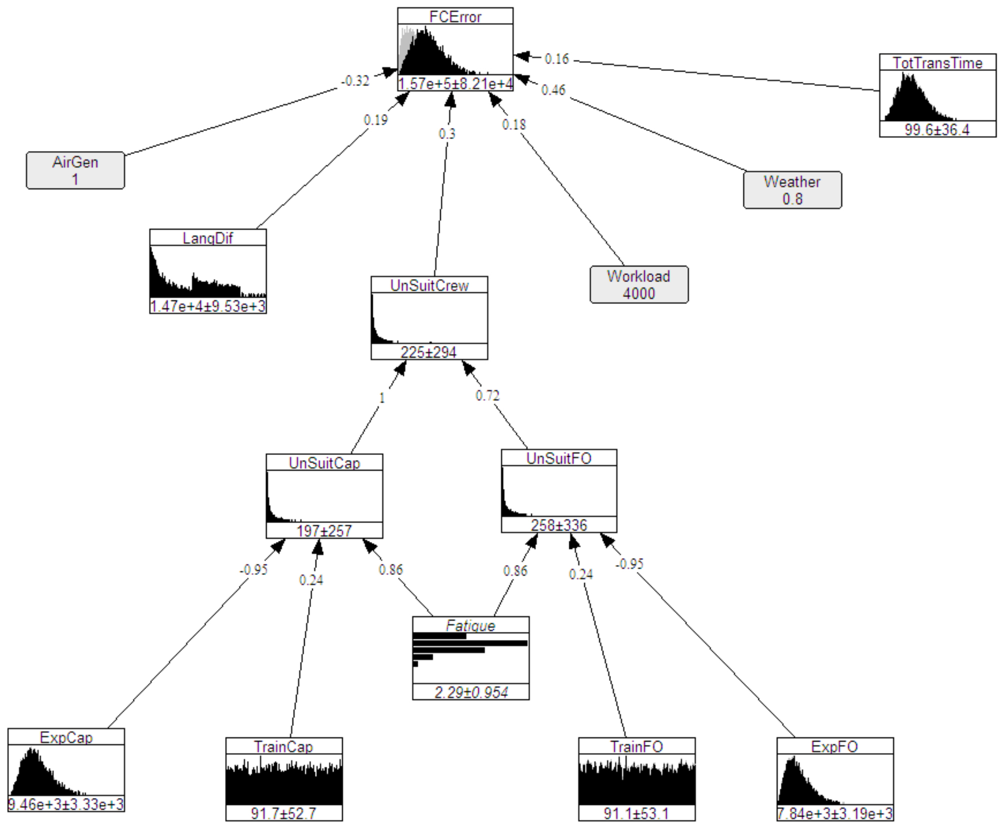


Figure 3.4: Conditional distribution of FCError|Weather=0.8, Workload=4000, AirGen=1.

This type of analysis can be done quite efficiently with non parametric continuous BBNs and with UniNet software.

To see which variable influences the FC error the most or in other words to try to explain the variability of which variable contributes the most to the variability of FC error, sensitivity analysis can be performed.

The sensitivity analysis is based on 50000 samples drawn from the BBN created in UniNet.

We would like to check how the "predicted variable", i.e. the FC error, depends on the "base variables": aircraft generation, weather, man-machine interface, workload, total transmission time, etc. As it was mentioned in Chapter 2, in our analysis we only used 3 statistics: product moment correlation, rank correlation and correlation ratio.

In figure 3.5 we present the sensitivity analysis for FC error related to each of the base variables based on the three measures described before.

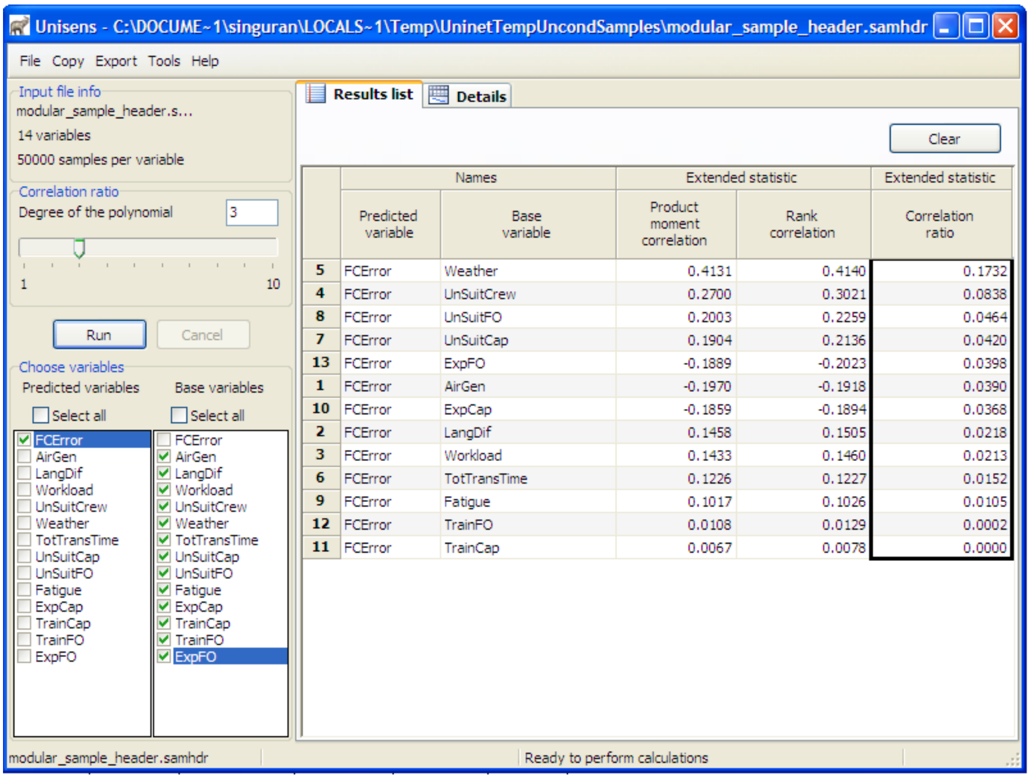


Figure 3.5: Sensitivity indices for the predicted variable FC Error and given base variables.

Each row in figure 3.5 shows the sensitivity indices for a given base variable, with respect to the predicted variable FC error. We can notice that the smallest correlation ratio of FC error is obtained with captain training and first officer training, while the highest with weather and it is equal to 0.1732. This means[2] that the variance of the FC error is explained by 17.3% of the variance of the conditional expectation of FC error given weather. The ratio of the largest to smallest rank correlation (different than zero) is 895. The product moment correlation and rank correlation do not vary significantly. The largest rank correlation is equal to 0.414 and it is corresponds to weather. Weather is followed by crew unsuitability with the rank correlation equal to 0.3021. The smallest rank correlations corresponds to first officer training and captain training and they are equal 0.0129 and 0.0078 respectively.

Next, we will discuss the Air Traffic Controller performance model.

[2]According to the formula 2.7.

3.2 Air Traffic Control Performance Model

3.2.1 Description of the model

The second of the generic models is the Air Traffic Control Performance Model. The BBN representing the Air Traffic Control Performance Model is shown in figure 3.6. Variables taken into account are briefly described below. The names of all variables in the text appear abbreviated in the BBN.

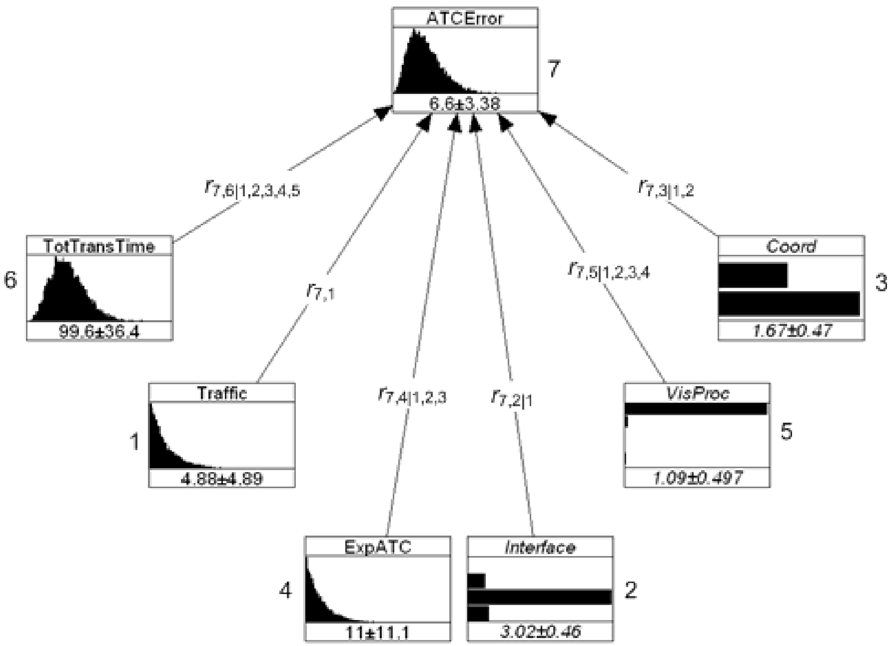


Figure 3.6: BBN representing the Air Traffic Control Performance Model.

The distributions of the variables are obtained form the data as indicated in the description. However, the distribution of ATCo error comes from the associated Fault Tree[3].

1. TRAFFIC (Data): Number of aircraft (any type) simultaneously under control.

2. MAN-MACHINE INTERFACE (Data): Four states variable. From 1- using radio only to 4-using radio, primary and secondary radar and additional tools.

3. COMMUNICATION - COORDINATION (Data): 1 - The communication with other ATCos takes place in the same room 2 - The communication with other ATCos does not take place in the same room.

4. ATCo EXPERIENCE (Data): Number of years working as an ATCo in the same position.

[3]The description and explanation of FT will be presented in more details in Chapter 4.

5. VISIBILITY PROCEDURE (Data): Five states variable. From 1 - normal operations to 5 - operations below 200 meters visibility.

6. TOTAL TRANSMISSION TIME (Expert Judgment): The total duration (in seconds) of the air/ground communications, per aircraft, for the approach and landing flight phase.

7. ATCo ERROR (Form the associated FT): Likelihood that the ATC control will make an error of a given kind.

The (conditional) rank correlations are obtained from experts using the third elicitation method described in section 2.2.2. 6 experts were interviewed and asked 19 questions. 5 of these questions were used to assign the marginal distributions and 12 to elicit calibration variables. Unfortunately, the estimates of one expert could not be used because the ratios which he/she gave were inconsistent[4]. The (un)conditional rank correlations obtained with the performance based combination of experts' assessments are presented in table 2. The ATC performance model is explained in more details in [6].

(Un)Conditional Rank Correlations	
$r_{7,1}$	-0.179
$r_{7,2\|1}$	-0.21
$r_{7,3\|1,2}$	0.18
$r_{7,4\|1,2,3}$	-0.06
$r_{7,5\|1,2,3,4}$	0.02
$r_{7,6\|1,2,3,4,5}$ [a]	0.18

Table 3.2: Dependence Information in the Air Traffic Control Performance Model Model.

[a] $r_{7,6|1,2,3,4,5}$ was elicited later from a single expert who is a risk analyst at NLR.

The rank correlations in table 2.2 are indexed by the variables labelling in figure 3.6. For instance $r_{7,1}$ represents the dependence between "ATCo Error" and "traffic" and $r_{7,2|1}$ represents the dependence between "ATCo Error" and "traffic" given "man-machine interface".

3.2.2 Analysis of ATC Performance Model

We can analyze the ATC performance model by conditioning on different values of some variables.

[4]The ratios of rank correlations are constraint by expert's previous estimates. In this case some estimates were outside the range available.

First, we conditionalize on the Man-Machine Interface (figure 3.7). We have selected a high value for this discrete variable, meaning that the air traffic controllers have the most advanced surveillance technology, i.e. radio, primary and secondary radar and additional tools (4). We can observe that the average number of errors when air traffic controllers are using the most advanced technology will be reduced from 6.6 to 5.56.

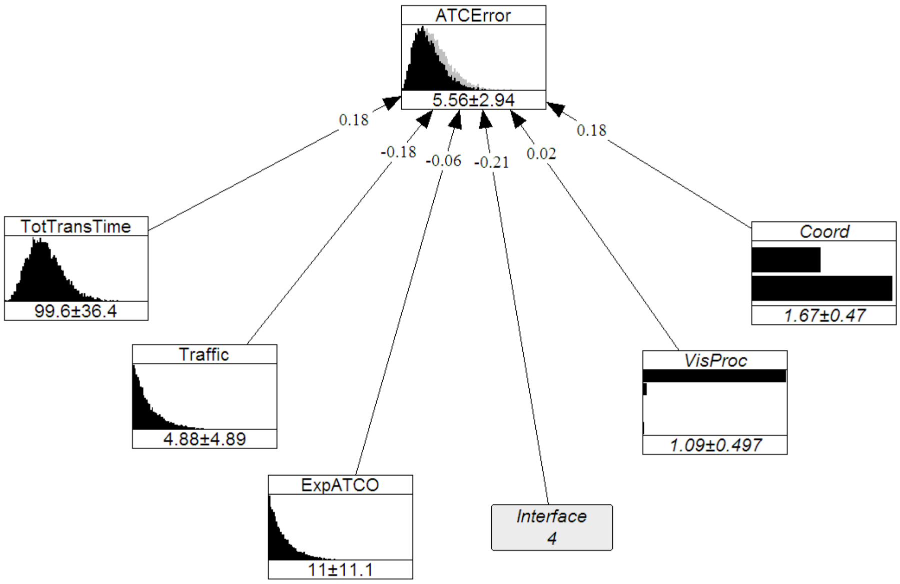


Figure 3.7: Conditional distribution of ATCError|Interface=4.

If in addition to the controllers using the most advanced technology we consider that the communication with other air traffic controllers takes place in different rooms (Coord=2) we notice a small change in the mean value of the ATC error. It decreases from 5.56 to 5.54 (figure 3.8). It means that given that the ATC works with proper technology (Man-Machine interface = 4) obstruction of the communication with other controllers does not have much influence on the number of ATC errors.

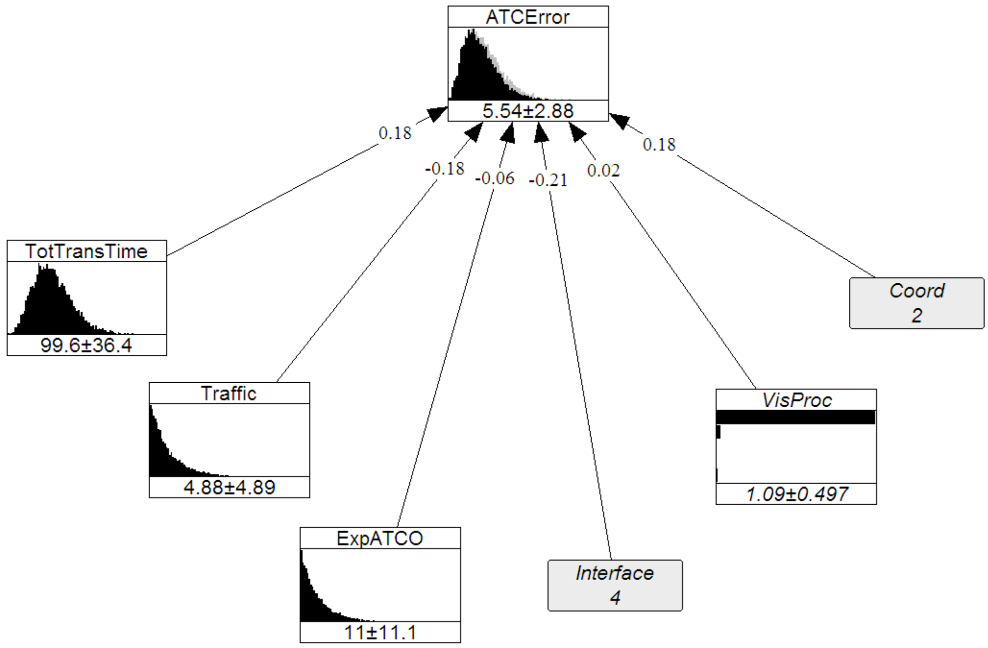


Figure 3.8: Conditional distribution of ATCError|Interface=4, Coord=2.

Finally, we condition on the continuous variable that is traffic. We assume that the air traffic controller has to direct and monitor 15 aircraft simultaneously. The mean value for the ATC error decreases to 4.66 (figure 3.9). It can be interpreted as high values of traffic influence ATC controller in a such way that he/she makes less errors[5].

[5]According to the experts working in high traffic situations kept them alert and less likely to commit errors.

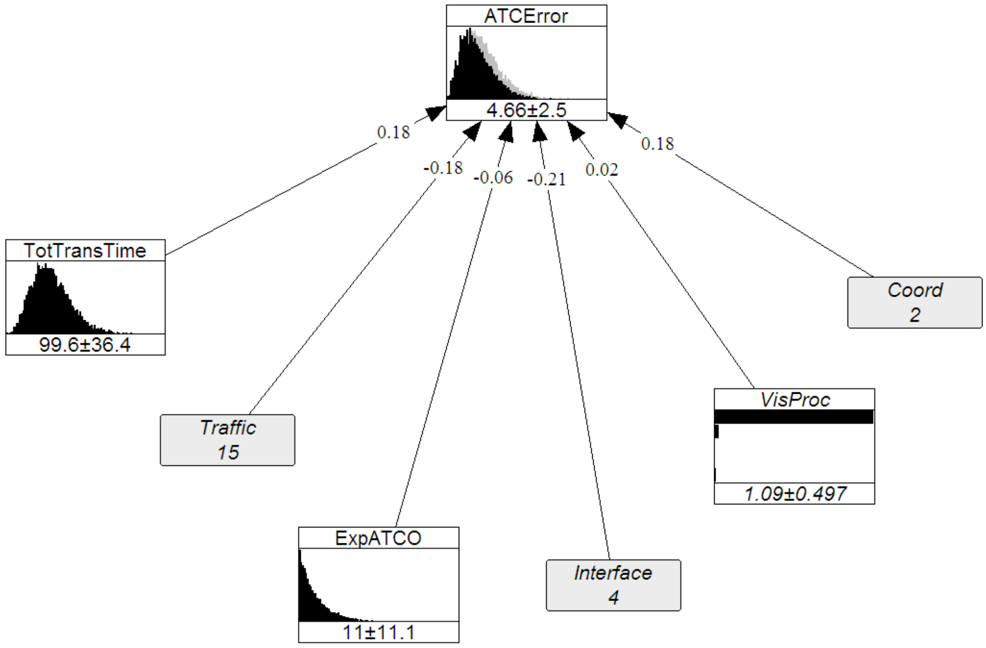


Figure 3.9: Conditional distribution of ATCError|Interface=4, Coord=2, Traffic=15.

Next, we present sensitivity analysis of ATC performance model.

We would like to check how the predicted variable, i.e. ATC error, depends on the base variables: traffic, man-machine interface, communication-coordination, ATC experience, visibility procedure and total transmission time. In the analysis we use the same statistics as for FC performance model. As in the case of the FC performance model, the sensitivity analysis is based on 50000 samples drown from the BBN created in UniNet.

In figure 3.10, we present the sensitivity analysis for ATC error related to each of the base variables.

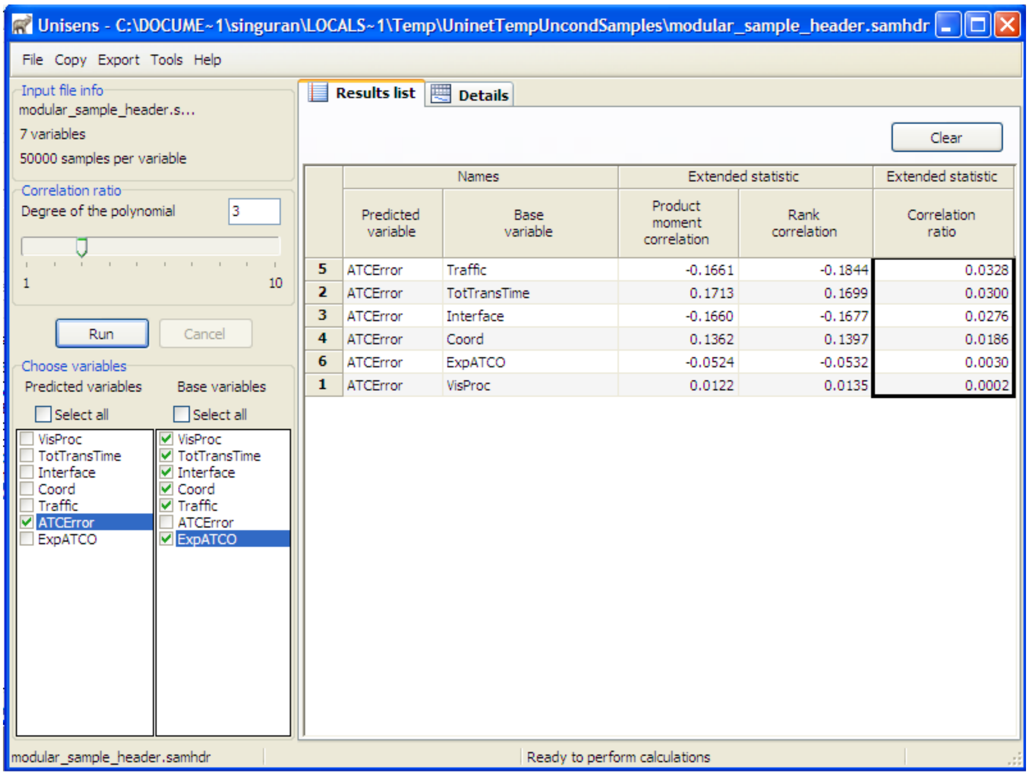


Figure 3.10: Sensitivity indices for the predicted variable ATC Error (AL19B8122 in this case) and given base variables.

Each row in figure 3.10 shows the sensitivity indices for a given base variable, with respect to the predicted variable ATC error. We can notice that visibility procedure has the smallest correlation ratio to ATCo error; traffic on the other hand has the highest value for correlation ratio of the model variables equal to 0.0328. In general it may be observed that the pattern of "importance" shown by the model variables is consistent across the three measures of dependence shown in figure 3.10. However according to the combination of experts' dependence estimates all variables are only slightly rank correlated to the ATC error the largest one (in absolute value) being traffic explaining only about 3.3% of the ATC error variance. However, the ratio of the largest to smallest correlation ratio is equal to 164. The values of the product moment correlation and rank correlation are similar. The largest rank correlation (in absolute value) corresponds to traffic (0.1844) while the smallest one to visibility procedure (0.0135).

Next, we discuss the Total Transmission Time node and its role in the model.

3.3 The Total Transmission Time node

In Chapter 4 we describe in more details how human performance models are connected to the rest of the CATS model. It is enough to mention here that different copies of these models are incorporated in different flight phases represented by the model. The structure of these models does not differ from one flight phase to the other. Some nodes, e.g. aircraft generation, experience, intra-cockpit communication, working condition, shift overlap time, etc. do not change through flight phases. Nodes like weather, workload, traffic, man-machine interface, communication-coordination, etc. may change during flight.

For en route and approach and landing phases an extra node connecting ATC and FC models is added. Interaction between flight crew and air traffic controllers should be modeled. The added node is called the total transmission time and its formal definition is as follows: the total duration (in seconds) of the air/ground communications per aircraft, for the approach and landing flight phase. Figure 3.11 represents the structure of the model for approach and landing which includes the Flight Crew performance model and the Air Traffic Control performance model.

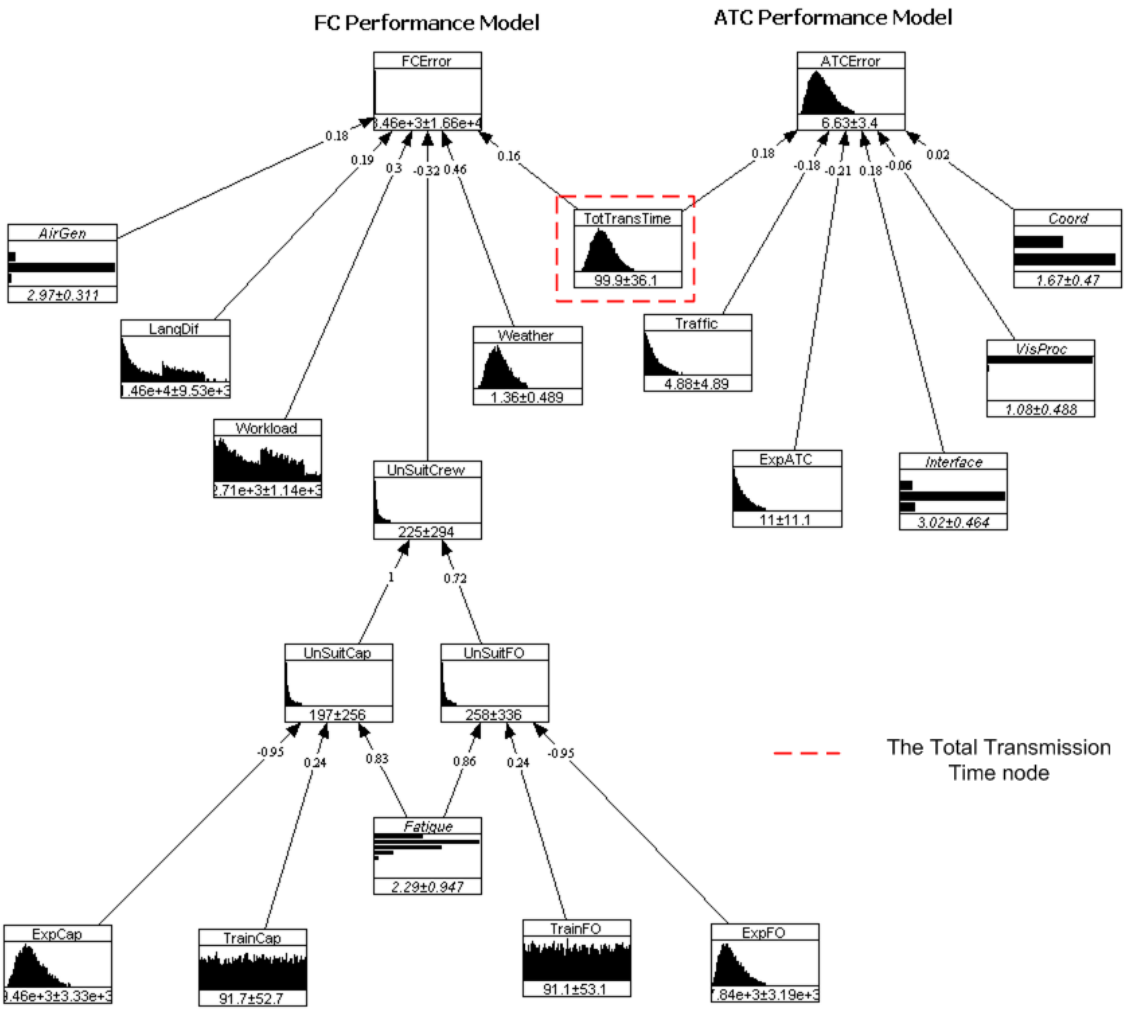


Figure 3.11: Two Human Reliability Models for Approach and landing flight phase. From left: Flight Crew performance model and Air Traffic Control performance model.

The conditional rank correlations for that node were elicited later from a single expert who is a risk analyst at NLR. In table 3.3, we present the conditional rank correlations of this node with the Flight Crew error and the Air Traffic Controller error.

Conditional Rank Correlations	
FC Error	ATC Error
0.16	0.18

Table 3.3: Dependence Information for the Total Transmission Time in the Flight Crew Error and the Air Traffic Control Error.

Next, we present the last of human reliability models: the Maintenance Technician performance model. The development of the Maintenance Technician model as well as an analysis of the model will be performed.

3.4 Maintenance Technician Performance Model

Maintenance refers to "all activities necessary to keep the aircraft in, or restore it to, a specified condition". Aircraft maintenance is a complicated and costly process. It is characterized by large amounts of regulations, procedures and documentation [9].

In the next sections we will present the Maintenance Technician Performance Model. We start by describing the influencing factors of the Maintenance Technician Error that have been considered the most important. We will briefly describe the variables and propose a way to quantify them.

3.4.1 Building the model

Since human behavior is quite difficult to predict, we looked for factors which would influence the human error. An overview of these influencing factors and the motivation for our choice will be presented in the following section. BBNs have been used to model the influences of those factors on the error probability. In order to specify the BBN a marginal distribution for each factor considered has to be assessed. Only for two variables (experience and aircraft generation), information was available from data. For the other variables (working condition, fatigue, shift overlap time and workload), expert judgment was used. To complete the Maintenance Technician Performance Model, we have quantified the dependencies between those factors and the maintenance technician error with help of expert opinion.

Figure 3.12 displays the structure of the Maintenance Technician Performance Model. It is a simplified representation of how the number of maintenance technician errors can be influenced. Five factors considered as the most important have been chosen: working condition, fatigue, experience, shift overlap time, aircraft generation and workload.

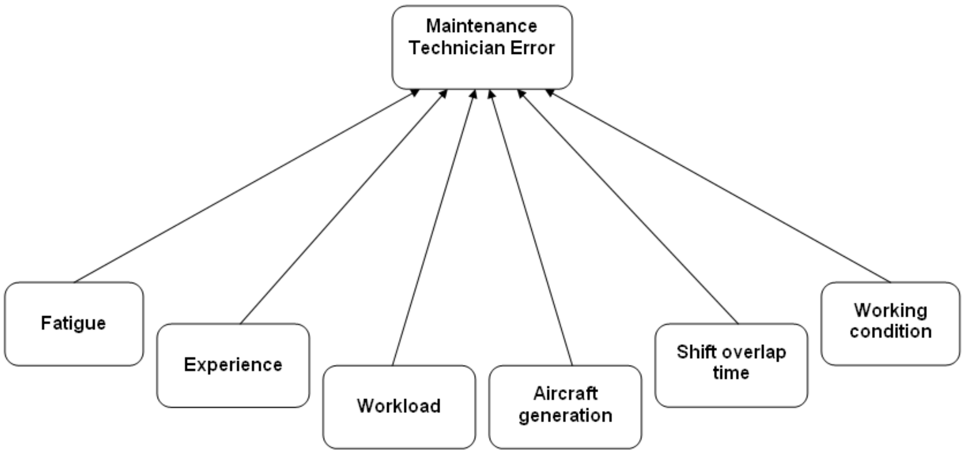


Figure 3.12: Structure of Maintenance Technician Performance Model.

The final graphical structure of the model in figure 3.12 was decided together with analysts of NLR and TU Delft upon the considerations stated at the beginning of this chapter. As stated previously one of the main requirements was to keep the model as simple as possible. Some of the dependencies that could have appeared in the model are not considered due to the operational definition of the nodes in the BBN (see next section). For instance, the time available to transfer a job was considered to be independent of wether the job is performed outside the ramp or inside the ramp because it depends on the shift duration rather than on the place where the job is being carried out.

Another rank correlation that was not considered in the model is that between fatigue and the estimated delay in the release of the aircraft. The delay in the release of the aircraft depends most of times on the availability of parts to be replaced and the availability of the mechanic to release the aircraft from a particular job. With these kind of considerations in mind and the fact that the model required to focus only on those factors that have the largest direct influence on errors, it was the decision of the research team to concentrate on the joint distribution as represented in figure 3.12.

Next, we present a short description of the variables considered in the Maintenance Technician Performance Model.

3.4.2 Model variables

The Maintenance Technician Performance Model contains the following variables:

Working Condition

Depending on the job to be performed, the maintenance technicians can work in the hangar or outside of the hangar. In general, the maintenance working conditions are preferable in the hangar rather than at the ramp. The main reason to do a job outside, e.g at the ramp, is lack of space or time. For instance, when only "small" work have to be performed on the aircraft, there is no need to take the aircraft inside the hangar. Taking the aircraft inside the hangar and then bringing it back takes on average 20 minutes. This imposes difficult time restrictions on the maintenance crew when the aircraft needs to be quickly released.

For the purpose of the model, the working condition is defined as a variable with two possible states: at the ramp (outside - 1) and in the hangar (inside - 2). The marginal distribution for this variable, describing the number of operations which are performed outside or inside the hangar, is obtained via expert judgment.

Fatigue

Fatigue is often mentioned as one of the most important human performance shaping factors in aircraft maintenance. One of the reasons is that much of the maintenance work is carried out during the night. Since maintenance teams work in shifts, some of the workers might have difficulties to adjust to the schedule. A shift system has, as well, influence on their social life. The human performance is influenced by the time of day. In general, people perform worst during the early hours in the morning [9]. According to a BASI study, the biggest relative frequency of incidents occurred during the night shift between 2 am and 4 am [11].

In our model Fatigue is represented by the Stanford sleepiness scale, where:

- 1 is feeling active, vital, alert, or wide awake;
- 2 is functioning at high levels, but not at peak; able to concentrate;
- 3 is awake, but relaxed; responsive but not fully alert;
- 4 is somewhat foggy, let down;
- 5 is foggy; losing interest in remaining awake; slowed down;
- 6 is sleepy, woozy, fighting sleep; prefer to lie down;
- 7 is no longer fighting sleep, sleep onset soon; having dream like thoughts.

The marginal distribution for this variable, describing percentage of the maintenance technicians been on one of 7 states during the time of performing the job, is obtained using expert judgment.

Experience

In our model Experience will be described as a number of years in the current position. We use data from the Bureau of Labor Statistics. In table 3.4 the data is presented.

Experience in years	% of the maintenance crew
3 or less years	22.8%
4-9 years	28.5%
10-19 years	16.2%
more than 20 years	32.5%
Median	9.4 years

Table 3.4: Summary of experience data.

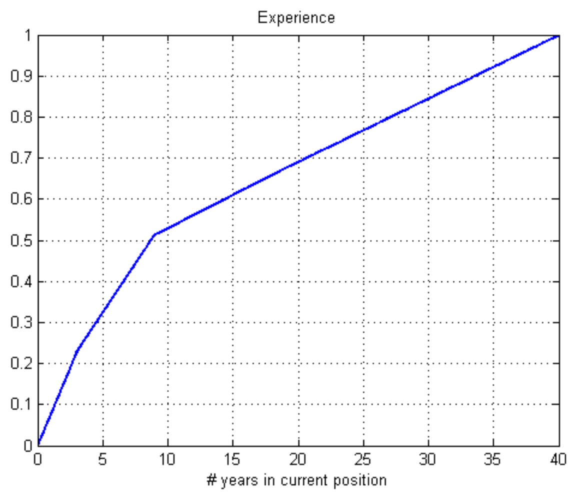


Figure 3.13: Data about experience, where the "solid" curve represents the minimum information solution assuming no MT has more that 40 years experience.

As we can see in the table 3.4, 67.5% of the whole maintenance crew has less than 20 years of experience. In the same table, we notice that 32.5% of the maintenance crew has more than 20 years of experience. The number of years working in the same position is considered as a continuous variable. Therefore we built the histogram from 0 to 40 years and after integration we obtain the cumulative distribution function in figure 3.13.

Shift overlap time

Most of the time, maintenance technicians have to deal with the situation when the tasks span more than one shift. This requires information to be passed from one shift to another. In these cases, the transfer of information might be a source of errors. In the model, shift overlap time will be described as a time available to transfer a task.

Expert judgment has been used to quantify the marginal distribution of this variable describing the available time (in minutes) to transfer a job.

Aircraft generation

In general, the ease with which aircrafts are maintained might change according to their design (older or newer designs); although some of maintenance technicians may still have personal preference for older types of aircraft. When thinking about the effect of technological advances on safety of air transport we should consider four different generations of aircraft.

First generation of aircraft is typically designed in the 1950s. Most of the aircraft were certified before 1965, according to BCAR's (British Civil Airworthiness Requirements) or other certification bases. Jet engines were still very new and the aircraft had very limited cockpit automation, simple navigational aids and limited approach equipment. Examples of first generation of aircraft are: the DH Comet, Fokker F-27 and Boeing 707.

Second generation of aircraft is designed in the 1960s and 1970s and have more reliable engines. The aircraft were certified between 1965 and 1980, but not yet based on common JAR-25/FAR-25 rules. Cockpit equipment is more advanced, with better auto pilots, auto throttles, flight directors and better navigational aids. Examples of second generation aircraft are: Fokker F-28, Boeing 737-200 and Airbus A-300.

Third generation of aircraft is designed in the 1980s and 1990s, typically show considerations for human factor aspects in the cockpit. Electronic Flight Instrument Systems (EFIS) and improved auto pilots are being used. Furthermore, the aircraft are equipped with ACMS data systems and high-by-pass engines designed according to higher certification standards. Examples of third generation aircraft are: Fokker 50 and Boeing 737-700.

Fourth generation of aircraft, like the Airbus A 320 and Boeing 777, have fully glass cockpits and digital fly-by-wire systems. The four aircraft generations provide a convenient classification for the human factors aspects of the man-machine interface and the associated maintenance procedures. Even though the aircraft operator has some freedom in developing its own maintenance schedule and associated procedures, this will have to be based upon the documentation developed by the aircraft manufacturer,which is subject to the aircraft type certification process [9].

In the table 3.5, we present a summary of aircraft generation data.

Type of aircraft	Participation
1	0.08%
2	6.14%
3	90.78%
4	3%

Table 3.5: Summary of aircraft generation data.

The data which we consider for aircraft generation is obtained from "Schiphol Statistical Annual Review 2000-2002". The data was defined by a scale from 1 to 4, where 4 is the most recent generation of aircrafts and 1 is the oldest generation of aircraft used nowadays.

Workload

High workload exists when task requirement is close to the operator's maximum capacity, while workload is low when the task requirement is much below the operator's capacity. Hence, workload is not only sensitive to multiple characteristics of tasks, i.e. task requirement, but as well to the operator's capacity. When thinking about the task requirement we should have in mind that it is determined by:

- number of actions,
- sequence of actions,
- time required for action to be completed,
- and type of action.

The number of tasks which have to be completed by a single maintenance technician depends of the number of available technicians per aircraft. Furthermore, it depends on certain characteristics, such as visibility of crew, average age of the crew and the distribution of the tasks between technicians [9].

In the Maintenance Technician Performance Model, workload is defined as estimated delay in release of the aircraft. The delay in the release of the aircraft is normally influenced by the four statements above and additionally by the availability of the human or material resources requires for actions to be computed. Since we do not have available data about this variable, we use expert judgment to obtain the marginal distribution.

Next, we will present Maintenance Technician Error variable.

Maintenance Technician Error

The formal definition of this variable is the number of unrecovered errors that the maintenance technician makes per number of jobs that are potentially hazardous for the safety of the flight. Since the duration of a job may vary from place to place, from time to time, or from position to position we are mostly interested in number of maintenance technician errors in the context of specific jobs, like component replacement, wheel replacement or fuel unit replacement, etc., rather than complete checking (which might even take 1,5 week).

In table 3.6, below we present a summary of the variables in the Maintenance Technician Performance Model, their formal definitions and the data sources for the marginal distributions.

Node	Definition	Unit	#	Source for marginal distribution
Maintenance Technician Error	Number of unrecovered errors that the maintenance technician makes per # of jobs that are potentially hazardous for the safety of the flight	Number of errors per job	7	DNV FT
Working Condition	Whether the work is performed at the ramp (outside - 1) or in the hangar (inside - 2)	1-2	1	Expert Judgment
Fatigue	Stanford sleepiness scale, where: 1 - Feeling active, vital, alert, or wide awake; 2 - Functioning at high levels, but not at peak; able to concentrate; 3 - Awake, but relaxed; responsive but not fully alert; 4 - Somewhat foggy, let down; 5 - Foggy; losing interest in remaining awake; slowed down; 6 - Sleepy, woozy, fighting sleep; prefer to lie down; 7 - No longer fighting sleep, sleep onset soon; having dream-like thoughts.	1-7	2	Expert Judgment
Experience	Number of years in current position	Number of years	3	Data
Shift Overlap Time	Time available to transfer a job	Minutes	4	Expert Judgment
Aircraft Generation	Four generations of aircraft, where 4 is the most recent generation of aircrafts	1-4	5	Data
Workload	Estimated delay in release of the aircraft	Hours	6	Expert Judgment

Table 3.6: Variables used in the Maintenance Technician Performance Model.

After we defined the model variables the influences on the maintenance error will be considered as in figure 3.14. Each of the variable presented in table 3.6 is a node in the structure of the model.

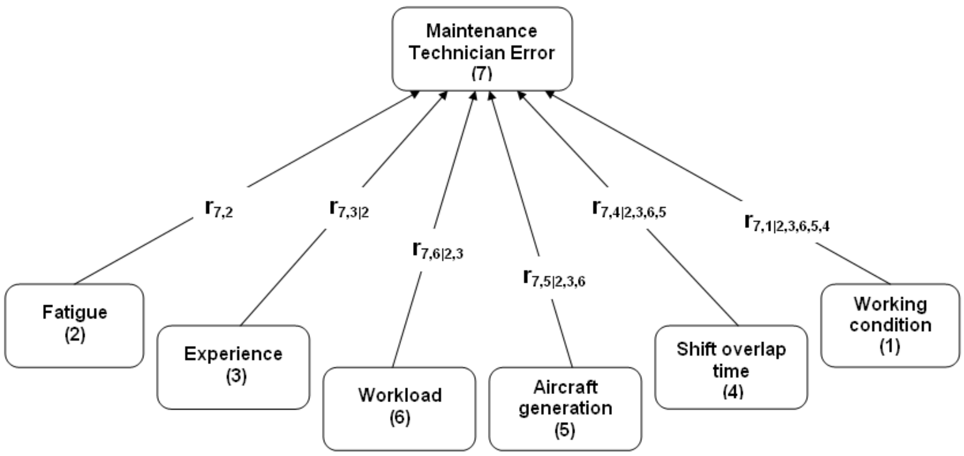


Figure 3.14: Maintenance Technician Performance Model.

In the next section we will present the quantification of the Maintenance Technician Performance Model.

3.4.3 Quantification of the model

In table 3.6 we can see that only for two variables, Experience and Aircraft Generation, the data for marginal distribution is available. For the other variables: Working Condition, Fatigue, Shift Overlap Time, Workload and Maintenance Technician Error as well as for dependence information, the expert judgment procedure needs to be performed.

The meeting with the single expert, who is a maintenance engineer for NLR (National Aerospace Laboratory), took place at NLR, on 28^{th} May 2008. One of the reasons that we did elicitation only with one expert was the lack of time. The CATS model should have been built and completed at the end of June and it should have contained the Maintenance Technician model. The other reason was that it was difficult to find a good expert who is working as a maintenance engineer and who was available that time.

The expert was asked 21 questions in total, where 5 questions were used to assign marginal distributions and 13 of them were used to elicit calibration variables. As stated previously the quantification presented here is not the final one for the CATS model. In principle the opinions of more experts will be obtained at later stages and combined using the classical method of expert judgment ([7]). All questions in the elicitation protocol refer to a population of maintenance technicians in the Western world (e.g. Europe, North America, Australia) and Western-built large aircraft ($> 5,700$kg Maximum Take-off Weight) currently flying in worldwide commercial operations.

As it has been already mentioned, the expert was asked 5 questions concerning marginal distribution. The expert was asked to specify his $5\%, 50\%$ and 95% quantile of uncertainty distribution for each of the variable of interest.

As an example, we will present one of the questions that have been asked the expert for the elicitation of marginal distribution. The rest of the questions and the protocol for elicitation can be found in the appendix.

- **Q4 (Shift Overlap Time):** What are your estimates for the 5^{th}, 50^{th} and 95^{th} percentiles of the distribution of the time available to transfer a job (minutes)?

He gave us the following answers to that question:

5%	5	50%	10	95%	20

In figure 3.15 we present a distribution of shift overlap time which was obtained from expert's answers.

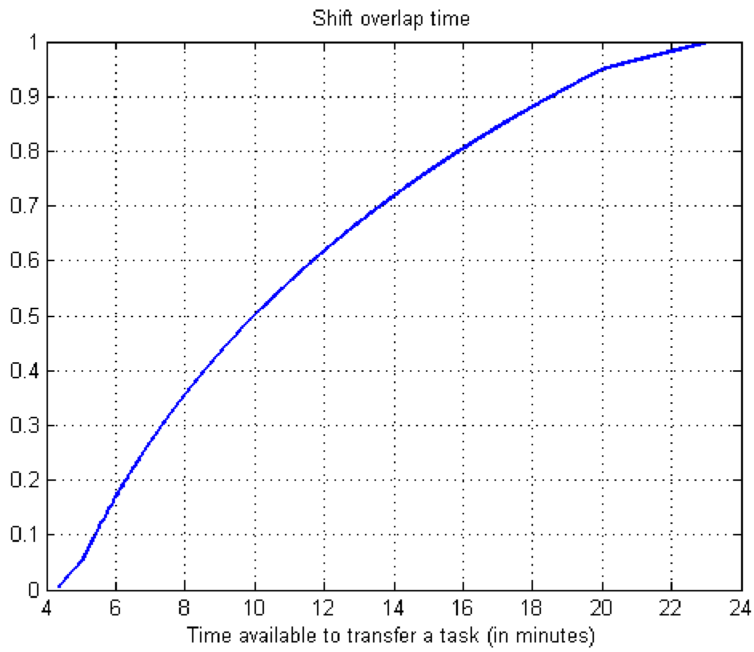


Figure 3.15: Distribution of shift overlap time obtained from expert's answers.

In order to gain insight about the relationship between the number of maintenance technician errors and the variables: working condition, fatigue, experience, shift overlap time, aircraft generation and workload, the expert was asked to rank variables according to the one which the considered most highly rank correlated (in absolute value) to the number of the maintenance technician errors. The results are presented in the table 3.7.

Variable	Rank
Working Conditions	4
Fatigue	1
Experience	1
Shift overlap time	3
Aircraft generation	3
Workload	2

Table 3.7: Ranking of variables according to the expert.

The expert chose Fatigue and Experience as variables witch have the strongest influence on the number of Maintenance technician error. The first rank correlation elicited was that between Fatigue and Maintenance technician Error and the order of variables were elicited according to the expert's ordering shown in table 3.7 as follows: 2,3,6,4,1 (e.g. fatigue, experience, workload, aircraft generation, shift overlap time and working condition). To obtain the rank correlation between fatigue and number of maintenance technician errors the usual probability of exceedence was used. For Fatigue the expert was asked the following question (**Q7**):

Suppose that 20,000 ***Maintenance technician*** *are randomly chosen from our total population. Out of those, 10,000 are selected for which the chosen variable (namely:* ***Fatigue****) has values above its median value (or above a certain percentile).*

What portion of these 10,000 ***Maintenance technician*** *will commit more than the median number of errors per median (*50% *quantile in question* ***Q1****) jobs. Observe that the median number of Maintenance technician errors was specified in question* ***Q2****).*

where

- **Q1 (Number of Maintenance technician jobs):** What are your estimates for the 5^{th}, 50^{th} and 95^{th} percentiles of the distribution of number of jobs that each Maintenance technician makes during a day?
- **Q2 (Number of Maintenance technician errors):** What are your estimates for the 5^{th}, 50^{th} and 95^{th} percentiles of the distribution of number of errors that each Maintenance technician makes per **median** (50% quantile in previous question) number of jobs ***median** average duration in hrs of a job?

Observe that the average duration of hours per job was asked previously. The expert was asked to express the influence of the remaining variables (other than Fatigue) as a portion of the influence of the variable ranked the highest $(0 - 100\%)$. He was asked also to specify the dominant direction (positive or negative) of the influence. This information was needed to compute the unconditional and conditional rank correlations required by the Maintenance Technician performance model. The results of expert's answers are shown in figure 3.16. This figure presents the panel form UniExp[6] software which accompanied elicitation process. The (conditional) rank correlations are obtained using the third elicitation method briefly described in section 2.2.2.

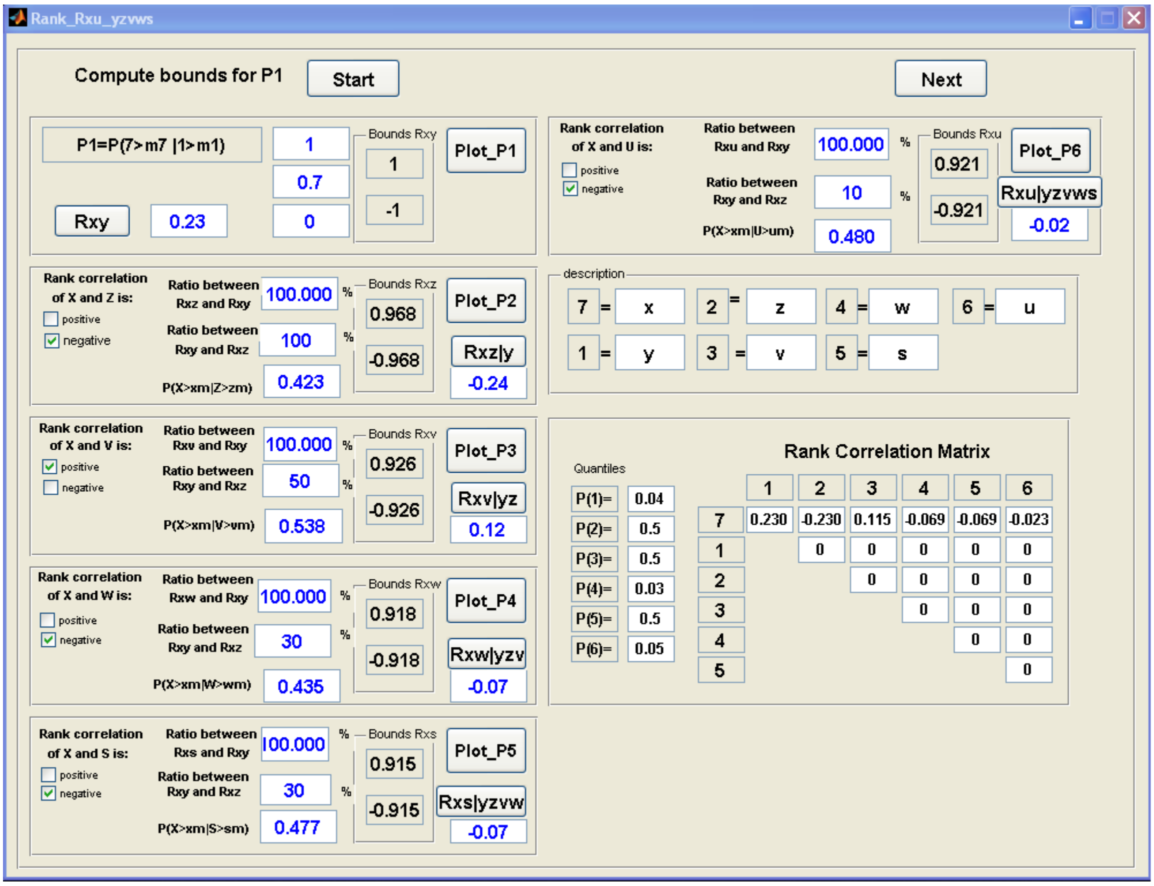


Figure 3.16: The panel from UniExp, where the numbers 1-7 corresponds to the following variables: 1 - Fatigue, 2 - Experience, 3 - Workload, 4 - Aircraft generation, 5 - Shift overlap time, 6 - Working condition and 7 - Maintenance technician error.

The "Description" window contains an explanation of the node of interest and all of its parents. Number 7 is assigned to Maintenance Technician Error (x), whereas numbers from 1-6 are assigned to contributed factors in the following order: 1 - Fatigue (y), 2 - Experience (z), 3 - Workload (v), 4 - Aircraft generation (w), 5 - Shift overlap time (s), 6 - Working condition (u).

Bounds for the conditional probability $P1$ (**Q7**) are filled in by the program. As mentioned in the Chapter 2, $P1$ is not restricted in any way and the expert can give us any number from $[0, 1]$.

[6]Brief description of that software is presented in [10, 15]

The conditional probability $P1$ corresponds to the question **Q7**, i.e. the expert was asked to give us the conditional probability that MT error is above median given that fatigue is above 4^{th} class. The answer which he gave was 0.7 and from that answer we obtain that MT error and Fatigue are positively correlated, i.e. when fatigue takes low values then MT error will be low as well. Having that probability we can compute the rank correlation between "maintenance technician error" and "fatigue" which is equal to 0.23. Figure 3.17 presents the plot of the relationship between probability $P1$ and rank correlation $r_{7,1}$.

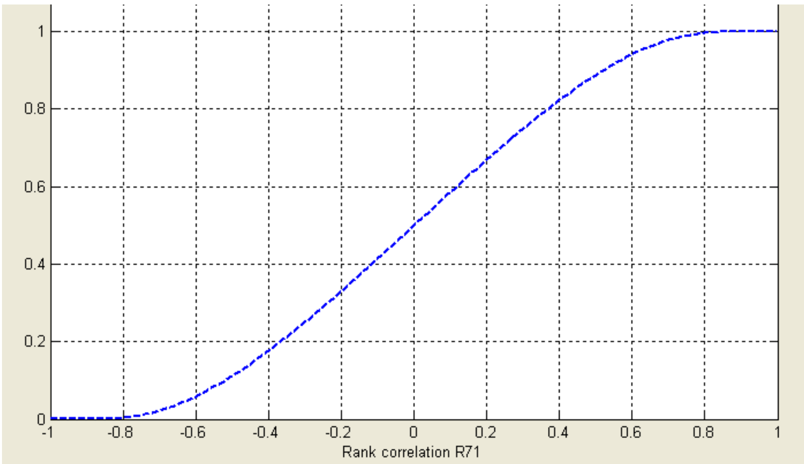


Figure 3.17: Relationship between $P(MTError \geqslant median|Fatigue > 4)$ and rank correlation $r_{7,1}$.

UniExp program fills the correlation obtained in the rank correlation matrix which is placed in the right hand side in the main window.

The expert was asked for the direction of the rank correlation (positive or negative) between maintenance technician error and experience, workload , aircraft generation, shift overlap time and working condition. As we can see, the direction of rank correlation ($r_{7,1}$) between "maintenance technician error" and "experience" is negative (it means that when experience takes high values then the MT error takes low values). Next we can see, the maximum value that the ratio between $r_{7,2}$ and $r_{7,1}$ may take according to the expert's previous answer. This ratio is computed by the program while the second ratio is the actual expert's belief. The procedure to compute the maximum and minimum values for the expert's next assessment follows the same idea presented in chapter 2 for computing the limits of the exceedence probabilities (see [12]).

Observe that in this case the ratio given by the expert has to be smaller than the ratio computed by the program and each ratio given by the expert cannot be bigger than the previous one[7]. Having specified direction of the rank correlations between

[7]Because the expert was asked to rank variables according to the absolute value of the rank

"maintenance technician error" and "experience" and the ratio between $r_{7,2}$ and $r_{7,1}$ we can compute the conditional rank correlation $r_{7,2|1}$. $r_{7,2|1}$ represents the rank correlation between "maintenance technician error" and "experience" given "fatigue" and it is equal to -0.24. Observe that the unconditional rank correlation in the correlation matrix is equal to the previous one as this was the expert's constraint.

Immediately after $r_{7,2|1}$ is computed, we obtain the conditional probability of exceedence that the expert would have stated if we would have been asked the probability that MTError is above median given that Experience is above the median. Figure 3.18 presents the plot of relationship between conditional probability and conditional rank correlation $r_{7,2|1}$.

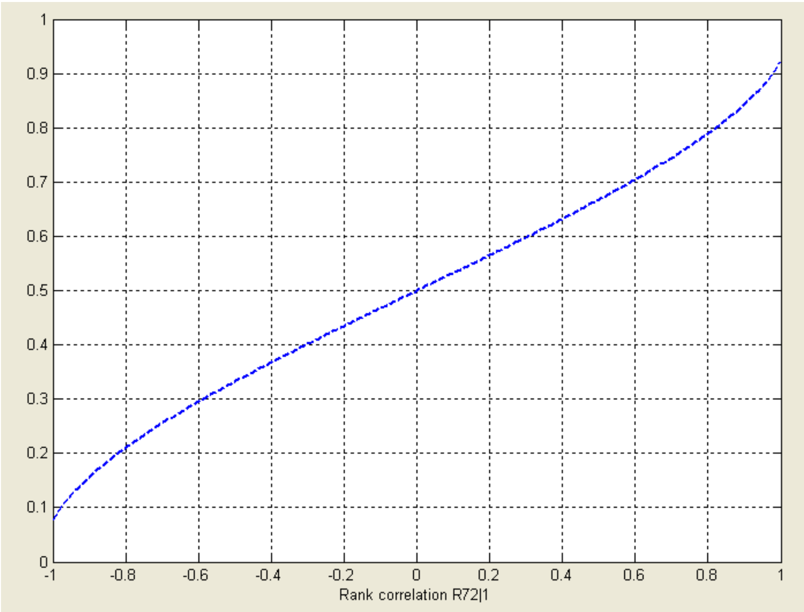


Figure 3.18: Relationship between $P(MTError \geqslant median|Experience \geqslant median)$ and conditional rank correlation $r_{7,2|1}$.

Observe that in this case the conditional probability of exceedence is $\in (0.08, 0.91)$ because part of the dependence has been explained by fatigue. The same procedure is used to obtain other conditional rank correlations.

The summary of the results from the elicitation of the dependence information is presented in table 3.8.

correlation and the dependence is asked in the order specified by each expert.

Probability[a]		(Un)Conditional Rank Correlations	
P_1	0.7	$r_{7,2}$	0.23
P_2	0.423	$r_{7,3\|2}$	-0.24
P_3	0.538	$r_{7,6\|2,3}$	0.12
P_4	0.435	$r_{7,5\|2,3,6}$	-0.07
P_5	0.477	$r_{7,4\|2,3,6,5}$	-0.07
P_6	0.48	$r_{7,1\|2,3,6,5,4}$	-0.02

Table 3.8: Dependence Information in the Maintenance Technician Performance Model.

[a] $P_1 = P(MT\ Error \geqslant median|Fatigue > 4)$, $P_2 = P(MT\ Error \geqslant median|Experience \geqslant median)$, $P_3 = P(MT\ Error \geqslant median|Workoad \geqslant median)$, $P_4 = P(MT\ Error \geqslant median|Aircraft\ generation = 4)$, $P_5 = P(MT\ Error \geqslant median|Shift\ overlap\ time \geqslant median)$, $P_6 = P(MT\ Error \geqslant median|Working\ condition = 1)$

The Maintenance Technician Performance Model has been quantified. Figure 3.19 presents the (un)conditional rank correlations assigned to each arc of the model, together with marginal distribution to each of the nodes.

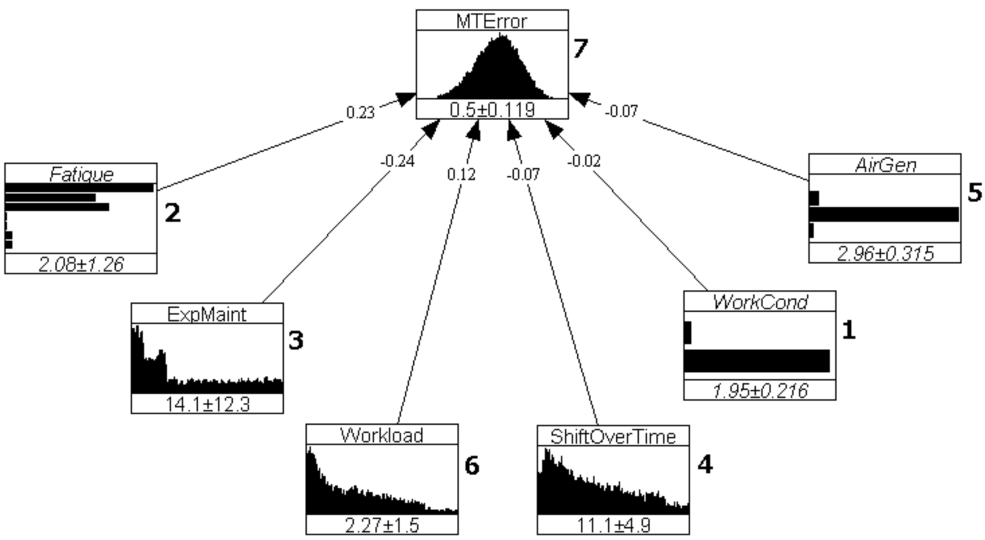


Figure 3.19: Quantification of the Maintenance Technician Performance Model.

In the next section, we will present the analysis of the Maintenance Technician Performance Model.

3.4.4 Analysis of the Maintenance Technician Performance Model

In the following section the goal is to update our belief in the Maintenance Technician Performance Model given some observations. For the simulations we use UniNet.

In the table 3.9 we present unconditional rank correlations between MT error and fatigue, experience, workload, aircraft generation, shift overlap time and working condition which we calculated and have seen in the rank correlation matrix presented in figure 3.16:

Unconditional Rank Correlations						
	Fatigue	Exp	Workload	AirGen	ShiftOverTime	WorkCond
MTError	0.23	-0.23	0.115	-0.069	-0.069	-0.023

Table 3.9: Rank correlations for MT model.

From table 3.9, we notice that 4 variables are negatively correlated with variable MT error. Therefore, when we conditionalize on high values (above median) of experience, aircraft generation, shift overlap time, working condition, we should observe a decrease in the estimated number of MT Errors. Fatigue and Workload are positively correlated with MT error. Therefore, when we conditionalize on low values (below median) of Fatigue, Workload we should observe an increase in the estimated number of MT errors.

Figure 3.19 presents the Maintenance Technician model with assigned conditional rank correlations and marginal distributions. The probability of MT Error is equal to 0.50 in this particular case.

From table 3.9 we can notice that the highest rank correlations in absolute value are between MT error and the 2 variables, Fatigue and Experience, they are equal to 0.23 and -0.23 respectively.

In figure 3.20 we can see how low value of the Experience, i.e. maintenance technician has 3 years of experience, influences the total number of mistakes by maintenance technician. Such a small experience can "negatively" influence the behavior of maintenance technician and the number of possible mistakes he/she can make. The expected number of MT Errors increased in this situation from 0.5 to 0.52.

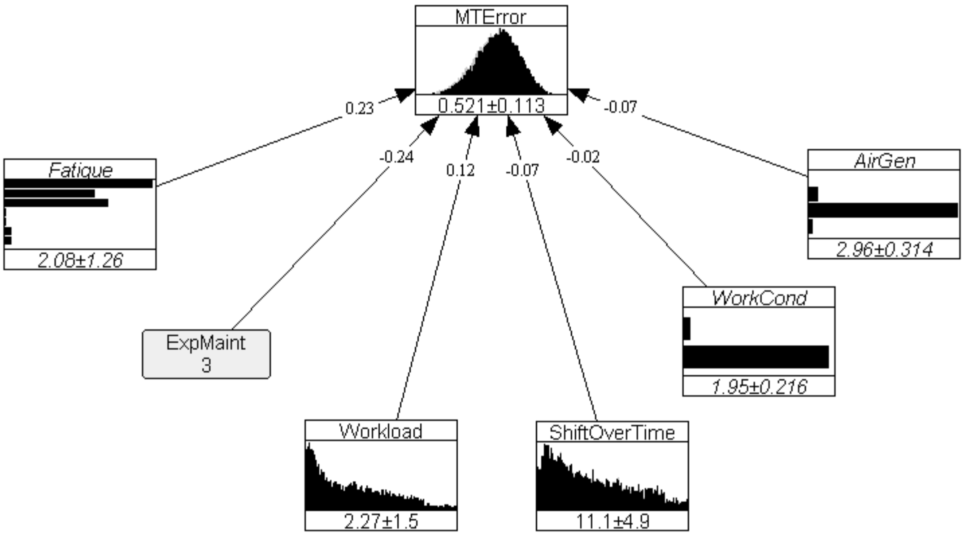


Figure 3.20: Conditional distribution of X_7|Exp=3, where X_7 refers to MTError.

In figure 3.21 we can observe how the previously obtained expected value of MT error changes if evidence additional to experience is available. We also consider that the maintenance technicians are feeling active, vital, alert or wide awake during their work (1^{st} class). We can notice that the mean MT errors increases back to 0.515, which means that inexperienced but not fatigued maintenance technician does not commit too many mistakes or are nit vary different than the unconditional distribution as shown in figure 3.19.

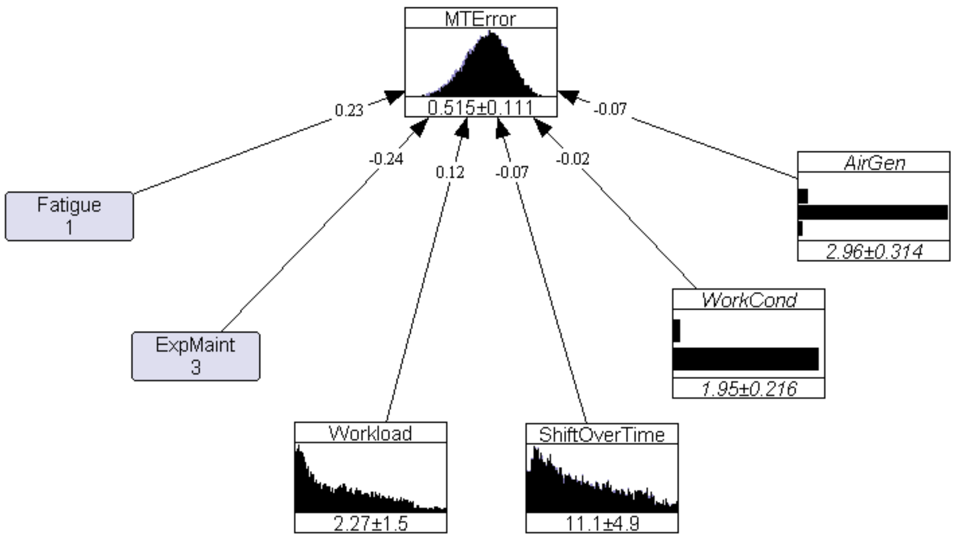


Figure 3.21: Conditional distribution of X_7|Exp=3, Fatigue=1, where X_7 refers to MTError.

Finally, we conditionalize on another two variables, workload and shift overlap time. We took into consideration the estimated delay in release of aircraft 5.5 hours (above median) and the time available to transfer a job is 6 minutes (below median) (3.22). We noticed that the mean of the MT error increased form 0.515 to 0.548, which means that if we additionally consider high delay in release of the aircraft and low time to transfer a task between workers then the maintenance technician can commit more mistakes than with unconditional case.

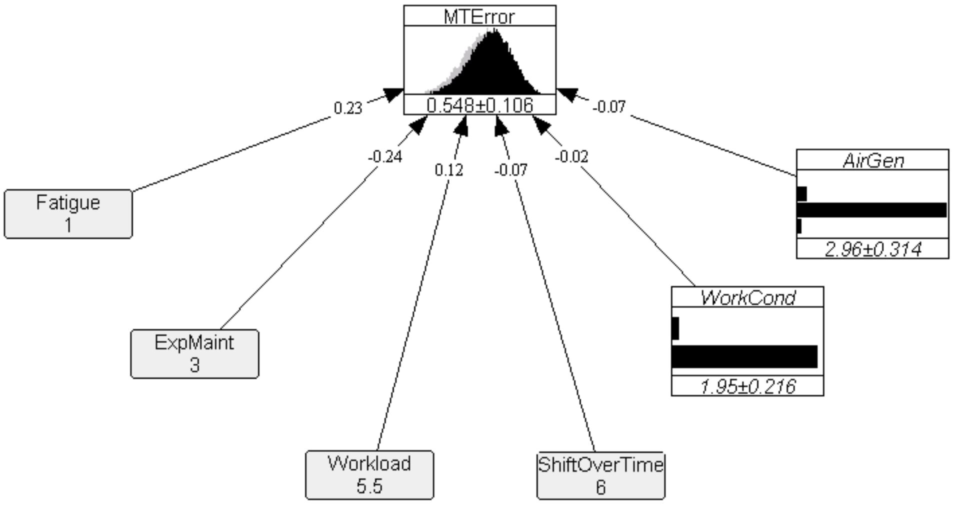


Figure 3.22: Conditional distribution of X_7|Exp=3, Fatigue=1, Workload=5.5, ShiftOverTime=6, where X_7 refers to MTError.

In the next section, we will present the analysis of the Maintenance Technician Performance Model.

3.4.5 Sensitivity Analysis of Maintenance the Technician Performance Model

As with the ATC model and the FC performance model, to carry out the sensitivity analysis several statistical and sensitivity measures are obtained by using UniNet. We drawn 50000 samples from the BBN created in UniNet and analyzed in Unisens.

The objective is to check how the predicted variable, i.e. Maintenance technician error, depends on the base variables: fatigue, experience, workload, aircraft generation, shift overlap time and working condition.

In figure 3.23, we present the sensitivity analysis for maintenance technician error related to each of the base variables.

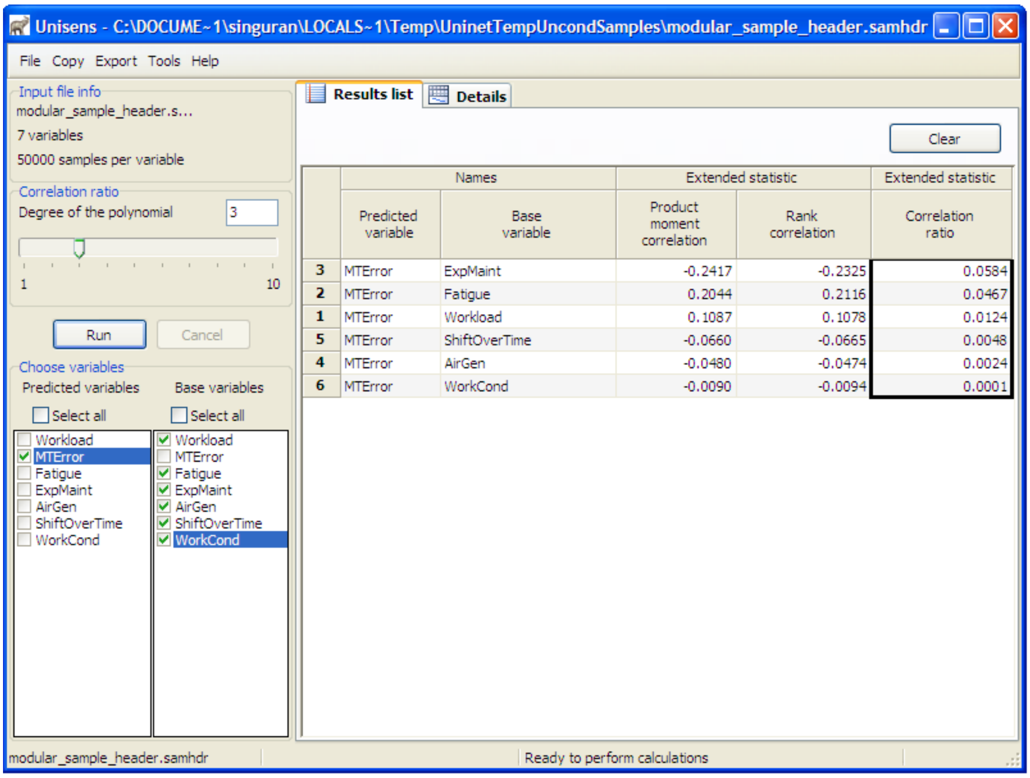


Figure 3.23: Sensitivity indices for the predicted variable Maintenance technician error (AL25B21 in this case) and a given base variable.

We can notice that the smallest correlation ratio is between the MT error and the working condition; while the highest, equal to 0.0584, is between experience and MT error. It means that[8] the variance of the MT error is explained by 5.8% of the variance of conditional expectation of MT error given experience. The ratio of the highest to lowest rank correlation is equal to 584. The differences in absolute value of the product moment correlation and the rank correlation are not significant. We can notice that the highest rank correlation is equal to 0.2325 in absolute value and is obtained for experience. Experience is followed by fatigue with the rank correlation ratio equal to the 0.2116. The rank correlation of fatigue and MT error is slightly lower that the theoretical one because fatigue is a discrete variable. As expected, the sensitivity analysis gives almost the same results given by the expert.

3.5 Final Remarks

In table 3.10, we present the result of scoring the expert that participated in the elicitation of the marginal distribution and unconditional and conditional rank correlations in the MT performance model.

[8] According to the formula 2.7.

Id	Calibration	Mean relative realization
Expert1	9.215E-005	0.383

Table 3.10: Expert's performance.

In table above, we present the calibration and information score for the expert used in the study of MT performance model. The first column gives expert id and the second column gives the calibration score. As we can observe, the expert is poorly calibrated with calibration score equal to $9.215E - 005$. If robustness analysis[9] is performed itemwise the calibration score would rise to 0.0006036 by removing any of CQ3, CQ4, CQ5, CQ7, CQ12 or CQ13 in Appendix A.

The information score for calibration items are shown in the third column. We can notice that this score is equal to 0.383 and it is relatively small.

For comparison purposes, during the elicitation procedure in ATC performance model the best calibrated expert was expert A with the correlation score equal to 0.1012 and in the same time he/she had the lowest information score for the calibration variables, which is a recurring pattern. Whereas, two of the experts were less calibrated than the Expert1. In case of the elicitation procedure in FC performance model, the expert with the best calibration score (expert B) has also one of the lowest information scores for the calibration variables, which is recurring pattern. His/her calibration score is equal to 0.6638. For more information about expert's performance in FC performance model and ATC performance model, check [5, 6].

In this chapter we saw that the sensitivity analysis performed, showed the most influential factors on human errors represented by the FC performance model, ATC performance model and MT performance model. We noticed that the largest product moment correlation and correlation ratio is in the FC performance model. The smallest product moment correlation, rank correlation and correlation ratio is in the ATC performance model. This may reflect the expert's belief that the air traffic controllers are aided by technological developments more than pilots or maintenance crew.

Weather has the biggest influence on the FC error. The correlation ratio between them is equal to 0.1732. In the ATC model, traffic has the strongest influence on ATC error with the correlation ratio between them equal to 0.0328. Whereas, the MT error is influenced the most by experience. The correlation ratio between these variables is equal to 0.0584. We can notice that correlation ratios between ATC error and traffic and between MT error and experience are small. It may be inter-

[9]Removing one item at the time and recomputing the scores in table 3.10 to see wether removing some item would give very different scores from those presented in table 3.10.

preted that only 3.28% of the variance of ATC error is explained by the variance of traffic, while the variance of MT error is explained by 5.84% of the variance in the maintenance technician experience. As compared with ATC error and MT error, the FC error variance is explained by 17.32% of the variance in weather.

It is worth mentioning that in the ATC performance model traffic and ATC error are negatively correlated. One may think that these variables should be positively correlated and maybe it is reasonable to think that high traffic may cause more errors. On the other hand, the experts told us that when ATC workers have more aircraft to control they are more careful, paying more attention and are more aware of the errors. That is the reason why these variables are negatively correlated.

One may notice that some of the (conditional) rank correlations in the ATC model, are rather small, for example, $r_{7,5|1,2,3,4}$ and $r_{7,4|1,2,3}$. This means that experts considered these two variables as the variables with the lowest influence on the ATC error.

Chapter 4

Causal Model for Air Transport Safety

As it was already mentioned, the Dutch Ministry of Transport and Water Management has initiated a research to investigate risks in aviation safety and develop a causal model for aviation safety. In this chapter we will present CATS model in more details. First, we present introduction to CATS model. We will give a brief description and example of Event Sequence Diagrams (ESDs), Fault Trees (FTs) and how human reliability models are attached in this model. Next, we show the development of the CATS model. We finish this chapter with procedure to obtain the underlying error distribution.

4.1 Introduction to CATS model

The Causal Model for Air Transport Safety (CATS) combines Event Sequence Diagrams (ESDs), Fault Trees (FTs) and Bayesian Belief Nets (BBNs) into a single structure, i.e. to continuous-discrete non parametric BBN. A schematic representation of the CATS model is presented in figure 4.1

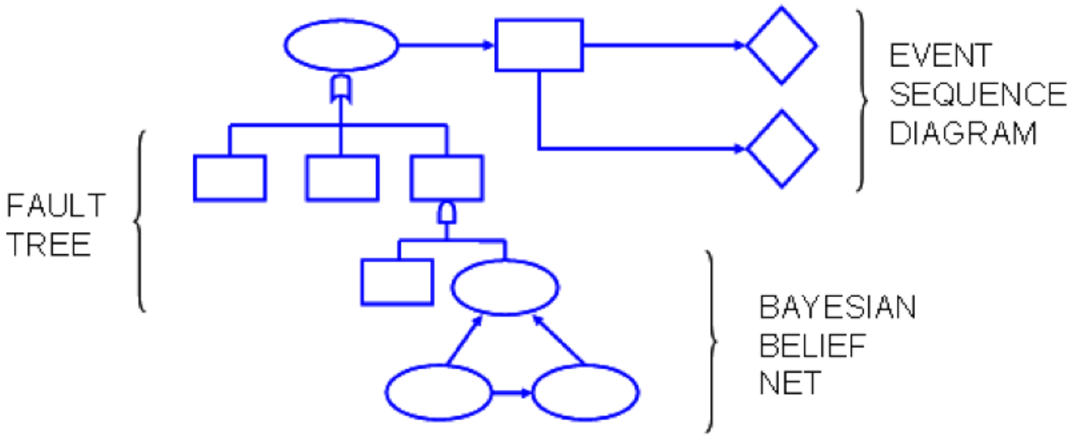


Figure 4.1: Schematic representation of the CATS model with ESDs, FTs and BBNs.

Figure 4.1 presents a conceptual structure of CATS model. ESDs are used to represent the top part of the model. They start with the initiating event which connects ESDs with FTs. The FT contains base events which are connected to BBNs. These represent one of the three human reliability models (FC performance model, ATC performance model and MT performance model) presented in previous chapter. It is visible that the simple idea presented in figure 4.1 realizes a very complicated graphical structure integrated into a single BBN.

More details about ESDs and FTs will be presented in the next section.

4.1.1 Event Sequence Diagram (ESD)

In figure 4.2[1] we can see an Event Sequence Diagram (ESD) which is a flow chart of paths leading to different end states. Each path through the flow chart is an accident scenario. Different accident scenarios are identified: abrupt manoeuvre, cabin environment, uncontrolled collision with ground, forced landing, controlled flight into terrain, mid-air collision, collision on ground, structure overload, and fire/explosion. In this way, 31 generic accident scenarios have been developed. Along each path, pivotal events are identified as either occurring or not occurring. The event sequence starts with an initiating event that requires some kind of response from operators or pilots or one or more systems ([3]). Operators, pilots or systems can solve the problem caused by the initiating event then aircraft continuous flying (green state). They may however fail to react properly which leads to accident (red state) or to incident that may influence later flight phases (orange state). Figure 4.2 presents the schematic representation of ESD with green, red and orange paths[2].

[1] Figure is taken from [3]

[2] For more information see [20].

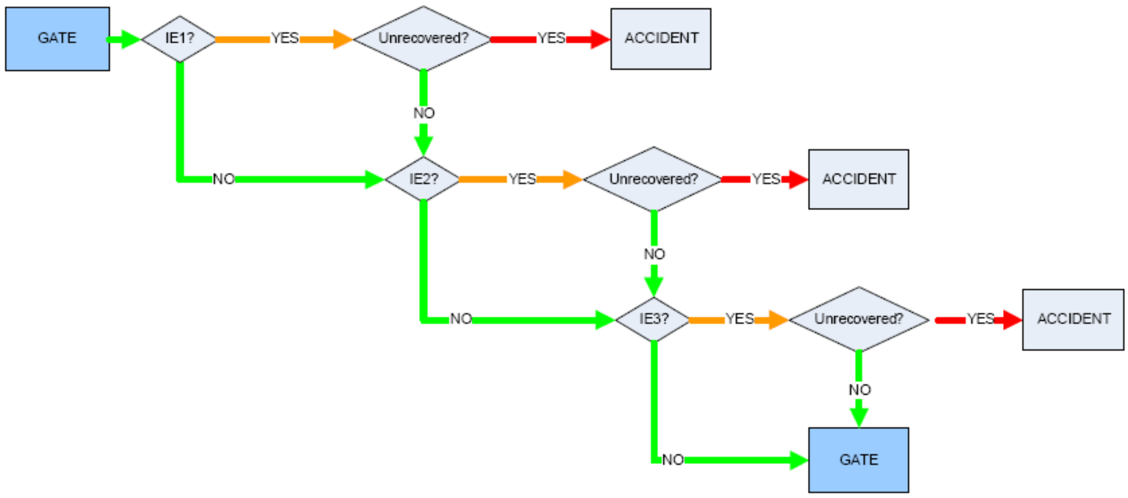


Figure 4.2: Event Sequence Diagram with green, red and orange paths.

The quantification of Event Sequence Diagram will be discussed next.

Quantification

In [3] three different ways to quantify ESDs are presented. In this section we only focus on one idea which was used to quantify ESDs. For more details see [3].

In the study, all probabilities are expressed as probability of occurrence conditional to the preceding pivotal event. The numbers at the pivotal events represent conditional probabilities, while the numbers at the initiating event and the end states are absolute probabilities. The sum of the probabilities of the end states is equal to the probability of the initiating event. At each pivotal event, the conditional probabilities add up to 1.

For the quantification of ESDs the data from NLR Air Safety Database were used. The NLR Air Safety Database includes the following databases[3]

- FAA Service Difficulty Reports (SDRs)
- Air Safety Reports (ASRs)
- Airclaims database
- ICAO ADREP
- NTSB aviation accident database
- FAA AIDS database

[3]For more details see [3]

- Aviation Safety Reporting System

Next, we will present one of ESDs as an example. We will explain in short initiating event, pivotal events and end sates. We will present also quantification of that ESD.

Example

As an example we present ESD 33, which is described as "Cracks in aircraft pressure cabin". Figure 4.3[4] represents the structure of that ESD, with an accident type described as "structure overload" and initiating event as "cracks in aircraft pressure boundary" in the take-off, initial climb, en route and approach and landing flight phases.

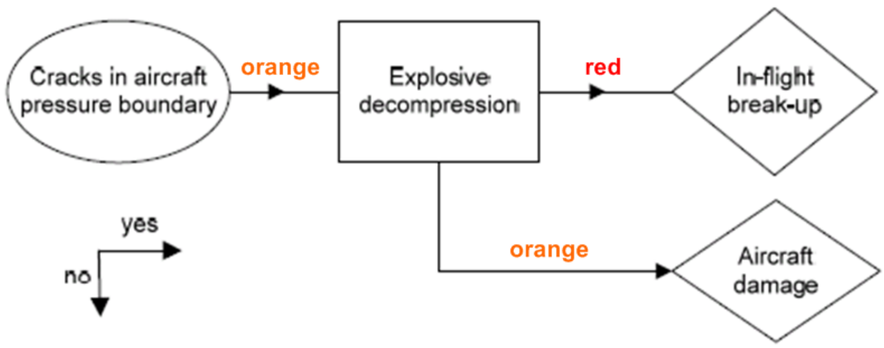


Figure 4.3: Structure of ESD 33 - Cracks in aircraft pressure cabin.

The initiating event, pivotal events and end states, in ESD 33, are defined as follows.

Cracks in aircraft pressure boundary (initiating event)

This event covers a crack in an aircraft pressure boundary. This crack can be different in location and size and it is developed over time. In this ESD, the focus is on those cracks that should have been detected during maintenance or line checks.

Explosive decompression (pivotal event)

In an explosive decompression, the aircraft cabin quickly decompresses resulting in major structural failure to the aircraft fuselage. Although there have been cases where aircraft landed "safely" following an explosive decompression (e.g. Aloha Airlines flight 243 on 28 April 1988).

[4]Figure is taken form [3]

In-flight break-up (end state)

The end state covers a severe damage to the aircraft caused by an explosive decompression. This could be a crash of the aircraft but also damage without crashing.

Aircraft damage / aircraft continues flight (end state)

This is the outcome of a crack in the aircraft pressure boundary which did not cause an explosive decompression. This could mean that there has been decompression of the pressure cabin but not of an explosive nature and it did not result in an accident, or nothing happened at all and the aircraft could safely continue the flight.

Quantification of ESD 33

Figure 4.4 presents the schematic representation of quantified ESD 33 with associated set of the probabilities representing initiating event, pivotal events and end states.

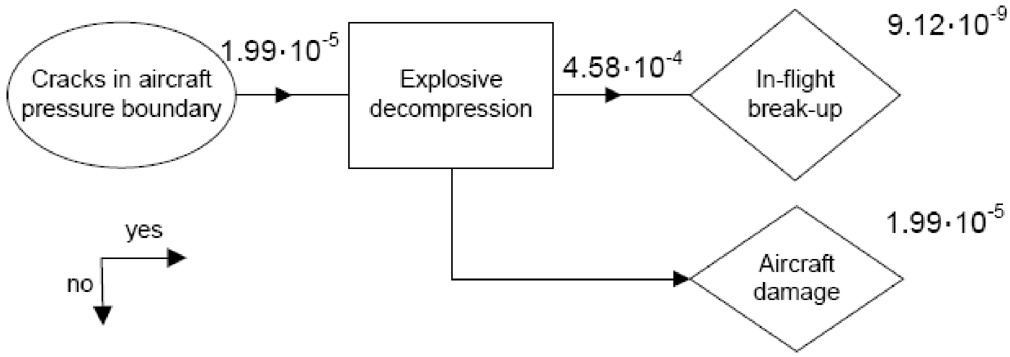


Figure 4.4: Quantified ESD 33 - Cracks in aircraft pressure cabin.

For example, an estimate of the frequency of occurrence of the initiating event "cracks in the pressure boundary" is based on service difficulty reports[5]. The dataset of service difficulty reports contains descriptions of all occurrences where "reportable" cracks have been detected during maintenance or line checks.The dataset contains at total of 3050 occurrences of cracks in aircraft pressure boundary and covers a total of 153.5 million flights. The associated frequency is equal to 1.9910^{-5} per flight. The probability of end state, "in-flight break-up", is computed as follows: $1.99 \cdot 10^{-5} * 4.58 \cdot 10^{-4} = 9.12 \cdot 10^{-9}$, whereas the probability of end event "aircraft damage" is calculated as follows: $1.99 \cdot 10^{-5} * (1 - 4.58 \cdot 10^{-4}) = 1.99 \cdot 10^{-5}$

[5]For details see Appendix A in [3].

For more details about quantification of the ESD 33 see [3].

In the CATS model, Event Sequence Diagrams (ESDs) are combined with Fault Trees (FTs). In the next section we present an overview of FTs.

4.1.2 Fault Tree (FT)

In practice, Event Sequence Diagrams are typically used to describe progress of events over time, while Fault Trees best represent the logic corresponding to failure of complex systems. Fault Trees are used to model initial and pivotal events in Event Sequence Diagrams in sufficient detail. The initiating and pivotal events in the Event Sequence Diagram are the top events in the Fault Trees.

FT's are used to describe the occurrence of an event in terms of the occurrence of another events or causes. They are logical trees and the sate of each block can only be "true" or "false". "True" usually can be interpreted as the failed state in the block described by the process, while the "false" describes the opposite state of that process.

When the FT is quantified, a probability is attached to each of the blocks, i.e. probability at any given time that the condition in the block is true (system is failed). Usually, the quantified probabilities are assumed to be independent (e.g. the probability of rapid decompression following fuselage failure is independent of the probability of rapid decompression following window failure). Other probabilities are assessed under the condition that other blocks in FT are true and then these probabilities are conditional probabilities.

For example, in figure 4.5, the probability of the right hand side branch under AND gate is conditional on the left hand block being true. In figure 4.5 the probability of "Failure propagates catastrophically" is conditioned on the left event, "Rapid Decompression". For each event, the FT provides the probability per demand. Mostly, for the events on the left, the regarding demand is a flight. As well, the contribution of each event to the failure is provided, as we can see in figure 4.5.

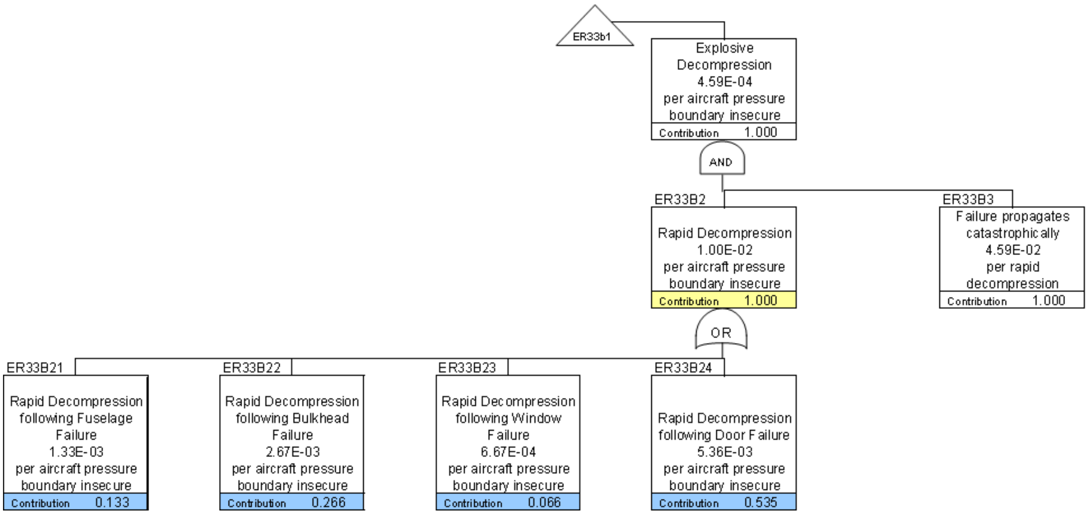


Figure 4.5: Part of FT associated with ESD 33.

To each FT a color coding is assigned to indicate the pedigree of the probability and contribution for each event. Four categories are distinguish:

- probabilities based on probability data (green),
- contributions based on the causal factors distribution of the accident data (blue),
- probabilities and contributions deducted from other values (white),
- probabilities based on pure expert judgment in the absence of data (yellow).

In the FT, the probabilities of events are generated by the AND or OR gates. They are quantified from the probabilities of the two inputs, which are assumed to be independent, as follows:

$$P(A \ \text{ AND } \ B) = P(A)P(B)$$

and

$$P(A \ \text{ OR } \ B) = 1 - (1 - P(A))(1 - P(B)).$$

Extensions to more that two events follow the same line. The FTs, in CATS model, were quantified using a top-down approach, which protected that the overall probabilities are consistent with the actual accident data. Some FTs were quantified independently (by DNV and NLR) form the ESDs, these independent quantification gave the consortium an opportunity to compare them. The results showed a remarkable resemblance in both quantifications, hence later quantification of FTs was based in NLR's ESDs. The FTs provide the best estimates of the average probabilities of events among commercial flights worldwide. The development of the fault trees is documented and presented in details in the DNV reports [16].

The base events are the most detailed layer of the FTs. Some of the base events represent human errors, detailed in human reliability models. This will be discussed in the next section.

4.1.3 Human reliability models in CATS model

CATS model describes human error, along with its contributing factors through 3 human reliability models: Flight Crew performance model, Air Traffic Control performance model and Maintenance Technician model. These generic models are presented in previous chapter.

Maintenance Technician model is defined for all flight phases, whereas Flight Crew performance model and Air Traffic Control performance model are specific to each of the three flight phases (take-off, en route and approach and landing).

As an example, we present the Flight Crew performance model for the approach and landing flight phase (figure 4.6).

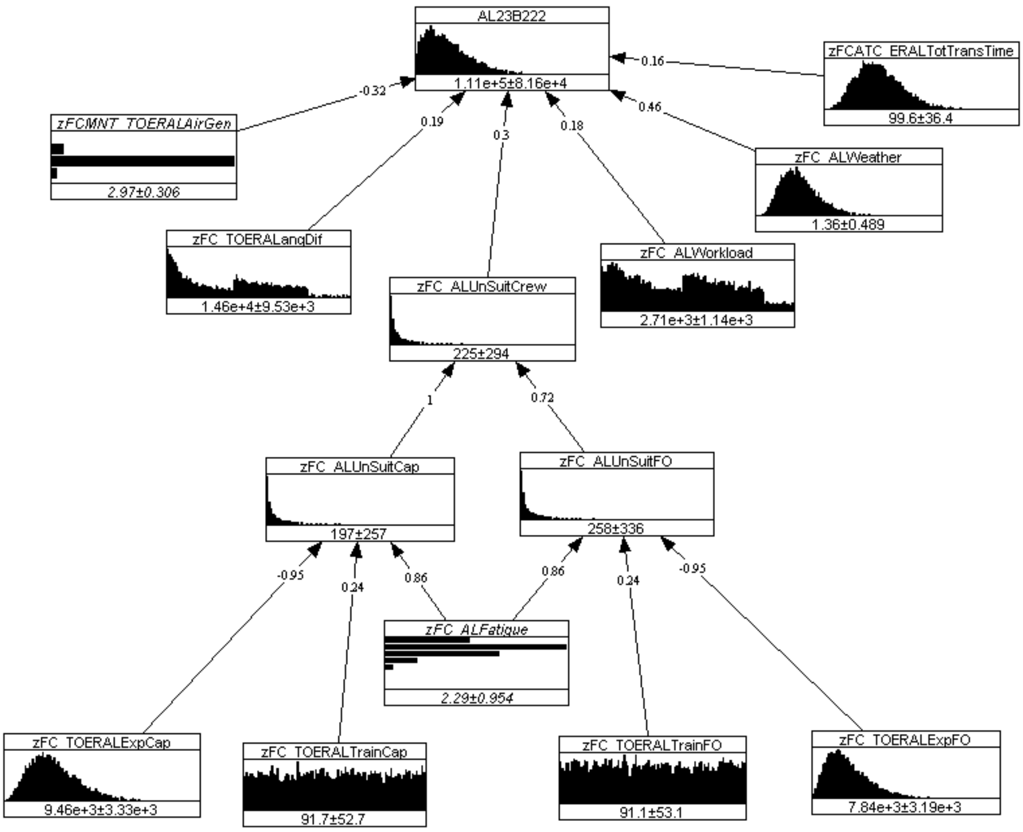


Figure 4.6: Information about human errors.

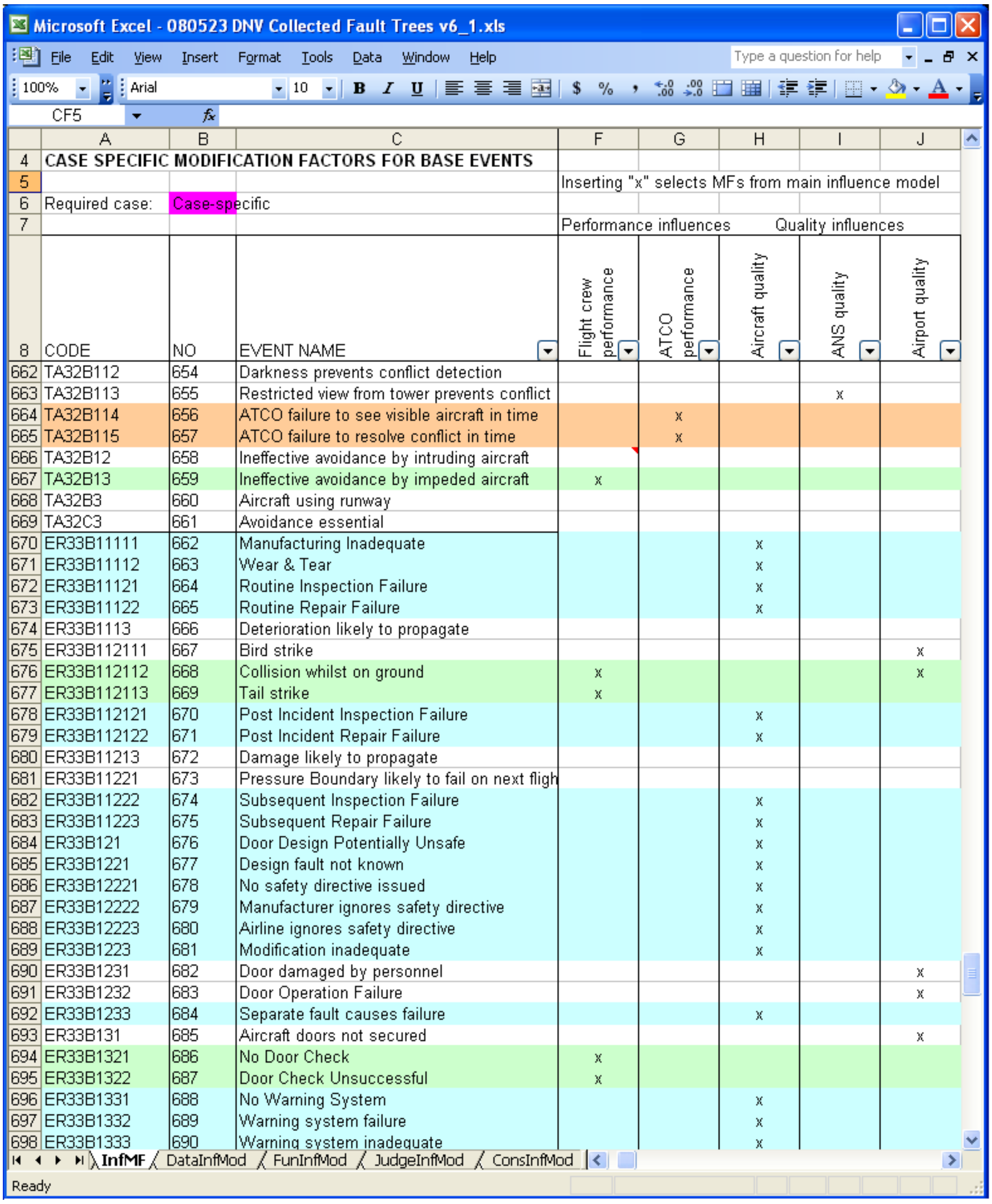
As we can see in figure 4.6, the FC error is replaced by the base event "AL23B222". AL23B222 refers to the flight for which the base event is defined as "pilot fails to execute a successful windshear escape manoeuvre (WEM)" and its formal definition is "the flight crew fails to execute and complete a successful windshear escape manoeuvre (WEM)". The procedure to obtain the underlying error distribution will be presented in section 4.3 in this chapter. For example, variable "zFC_TOERALAirGen", i.e. aircraft generation, is defined for the whole flight, whereas variable "zFC_ALWeather" is defined for each flight phase separately.

The same applies for the Air Traffic performance model.

4.2 Development of CATS model

In this section we present an overview of building the Causal Model for Air Transport Safety (CATS) in UniNet.

The Excel file (**080523 DNV Collected Fault Trees v6_1.xls**) provided by DNV, specifies among other things, to which base event is a specific human reliability model attached to. Figure 4.7 presents the classification of the base events according to human reliability models, FC performance model, ATC performance model and MT performance model. For example, as presented in figure 4.7, the event "Door Check Unsuccessful" (ER33B1322) belongs to the FC performance model. The event TA32B114, on the other hand corresponds to an ATC performance model outcome.



	A	B	C	F	G	H	I	J
4	CASE SPECIFIC MODIFICATION FACTORS FOR BASE EVENTS							
5				Inserting "x" selects MFs from main influence model				
6	Required case:	Case-specific						
7				Performance influences		Quality influences		
8	CODE	NO	EVENT NAME	Flight crew performance	ATCO performance	Aircraft quality	ANS quality	Airport quality
662	TA32B112	654	Darkness prevents conflict detection					
663	TA32B113	655	Restricted view from tower prevents conflict				x	
664	TA32B114	656	ATCO failure to see visible aircraft in time		x			
665	TA32B115	657	ATCO failure to resolve conflict in time		x			
666	TA32B12	658	Ineffective avoidance by intruding aircraft					
667	TA32B13	659	Ineffective avoidance by impeded aircraft	x				
668	TA32B3	660	Aircraft using runway					
669	TA32C3	661	Avoidance essential					
670	ER33B11111	662	Manufacturing Inadequate			x		
671	ER33B11112	663	Wear & Tear			x		
672	ER33B11121	664	Routine Inspection Failure			x		
673	ER33B11122	665	Routine Repair Failure			x		
674	ER33B1113	666	Deterioration likely to propagate					
675	ER33B112111	667	Bird strike					x
676	ER33B112112	668	Collision whilst on ground	x				x
677	ER33B112113	669	Tail strike	x				
678	ER33B112121	670	Post Incident Inspection Failure			x		
679	ER33B112122	671	Post Incident Repair Failure			x		
680	ER33B11213	672	Damage likely to propagate					
681	ER33B11221	673	Pressure Boundary likely to fail on next fligh					
682	ER33B11222	674	Subsequent Inspection Failure			x		
683	ER33B11223	675	Subsequent Repair Failure			x		
684	ER33B121	676	Door Design Potentially Unsafe			x		
685	ER33B1221	677	Design fault not known			x		
686	ER33B12221	678	No safety directive issued			x		
687	ER33B12222	679	Manufacturer ignores safety directive			x		
688	ER33B12223	680	Airline ignores safety directive			x		
689	ER33B1223	681	Modification inadequate			x		
690	ER33B1231	682	Door damaged by personnel					x
691	ER33B1232	683	Door Operation Failure					x
692	ER33B1233	684	Separate fault causes failure			x		
693	ER33B131	685	Aircraft doors not secured					x
694	ER33B1321	686	No Door Check	x				
695	ER33B1322	687	Door Check Unsuccessful	x				
696	ER33B1331	688	No Warning System			x		
697	ER33B1332	689	Warning system failure			x		
698	ER33B1333	690	Warning system inadequate			x		

Figure 4.7: Information about human errors.

Each of the human reliability models is represented by a different color. For example, the FC performance model is represented by green, the ATC model by orange/red and the MT model by blue. We may notice that some base events do not represent one of mentioned human errors. This reflects the fact that the base event corresponds to a non-human error. These events are represented by the white color. For example, in figure 4.7 when we are looking for the FC error and ATCo error we should look on the performance influences column in the Excel file. For the MT error, we should look at the aircraft quality in the quality influences columns in Excel file provided by DNV.

In building the CATS model the aforementioned Excel file has been consulted to be able to identify all base events corresponding to one of the three considered human errors. In the CATS model there are 280 nodes representing the FC error, 57 nodes representing the ATC error, 197 nodes representing the MT error and 240 nodes representing non-human error. It is worth to mention that the number of FC errors outruns both ATC and MT errors, which means that the FC performance has bigger influence on the accident probability than ATC or MT perfroamnce.

Next, we present a brief description of the structure of one ESD, ESD 33 - cracks in aircraft pressure cabin. This ESD already quantified corresponds to figure 4.3. Figure 4.8 presents the initiating event, the pivotal event and end events in the ESD 33. The initiating event is described as “Cracks in aircraft pressure boundary”, whereas the pivotal event is described as “Explosive Decompression” and it is either occurring or not occurring. The end states are described as follow: “In-flight break up” and “Aircraft damage”.

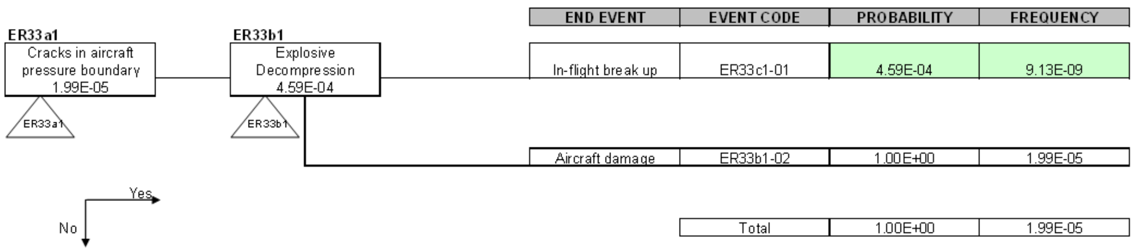


Figure 4.8: Initiating event, pivotal event and end events in ESD 33.

As previously mentioned, where necessary, the initiating and pivotal events are detailed into breakdown causes using fault trees. The initiating and pivotal events are the top events of the FTs. For ESD 33, both initiating and pivotal events have been detailed using FTs.

Figure 4.9 presents the structure of ESD 33 and its corresponding FTs, as it has been build in UniNet. The whole ESD 33 along with its corresponding FTs, contains 56 nodes. 2 of them are end states and they are the top nodes in figure 4.9. Below these two end state nodes, 1 initiating event and 1 pivotal event are observed. 17 nodes are intermediate events and 35 of them represent base events. Out of the 35 base events, 23 are for MT error, 5 for FC error, 0 for ATC error and 7 for non-human error. Whenever the base events represent a human error they are connected to the correspond human reliability models as illustrated in figure 4.9.

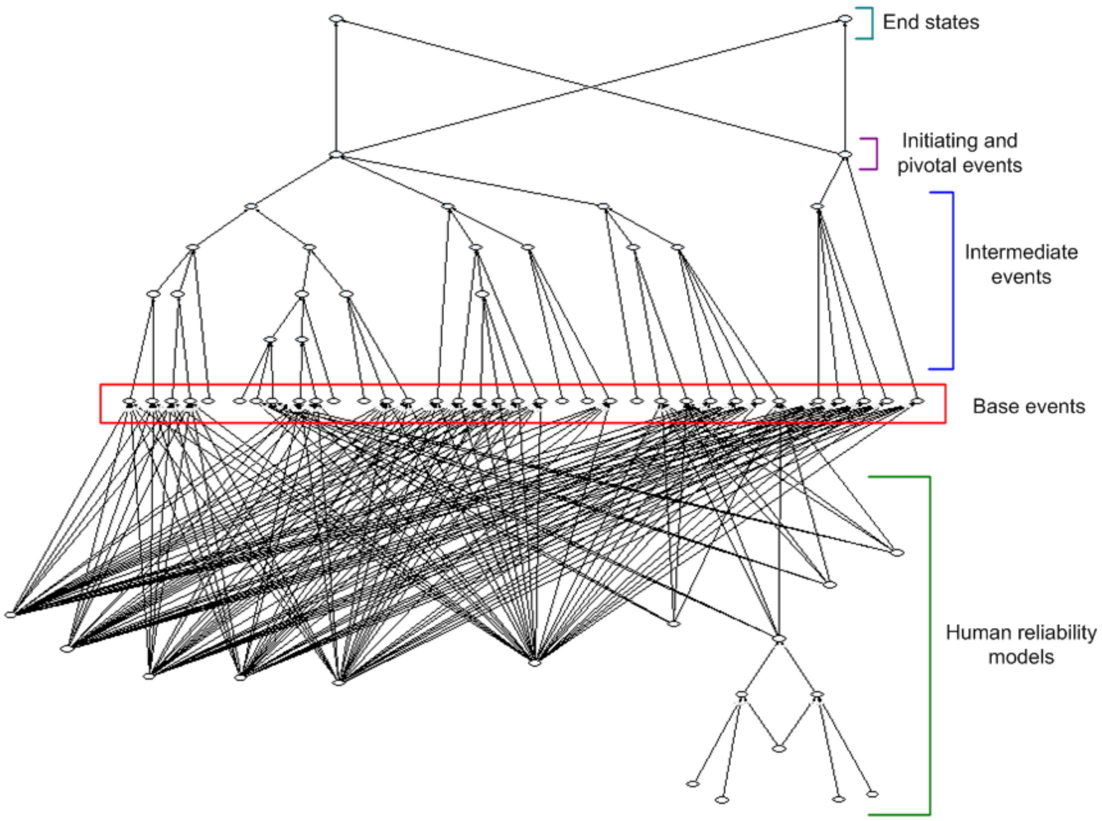


Figure 4.9: The representation of ESD 33 with human reliability models (from left: MT performance model and FC performance model).

We will start our description from the top of BBN. As we can see in figure 4.9, there are 2 nodes representing end states. The first one from the left is "in-flight break up" and the next one on the right is "aircraft continues flight damaged". These 2 end events are connected with the initiating event, described as cracks in aircraft pressure boundary (node on the left side), and the pivotal event, described as explosive decompression. The pivotal event is identified as either occurring or not occurring. The initiating event and the pivotal event represent the top events in the corresponding FTs. Therefore, they are connected to the 17 intermediate events. The end events, the initiating event, the pivotal event and the intermediate events are functional nodes. Finally, the intermediate events are connected to the base events. Some base events are representing one of human errors, i.e. FT error (36.2%), ATC error (7.4%) and MT error (25.4%). The other base events represent non-human errors (31%)[6].

Figure 4.10 presents the CATS model in UniNet as it stood before the work described in this thesis began. The model contained at the time 1479 arcs and 649 nodes. From these nodes, 332 were functional nodes representing ESDs and

[6]As stated before, the procedure to obtain the underlying error distribution will be presented in section 4.3 in this chapter.

FTs as boolean functions and 317 were probabilistic nodes. The 20 ESDs (out of 31 generic accidents scenarios) presented in the figure 4.10 are listed below. Three phases are represented in the CATS model: Take-off (TO), En-route (ER) and Approach-landing (AL). As we can notice not all of end event(s) are connected to the Accident/Incident node (the node on the top of BBN).

A similar situation is observed on the bottom part of the BBN, where we have FC performance model for: take-off, en route and approach and landing and as well as ATC performance model for: take-off and approach and landing. The first 20 ESDs shown in figure 4.10 are listed next.

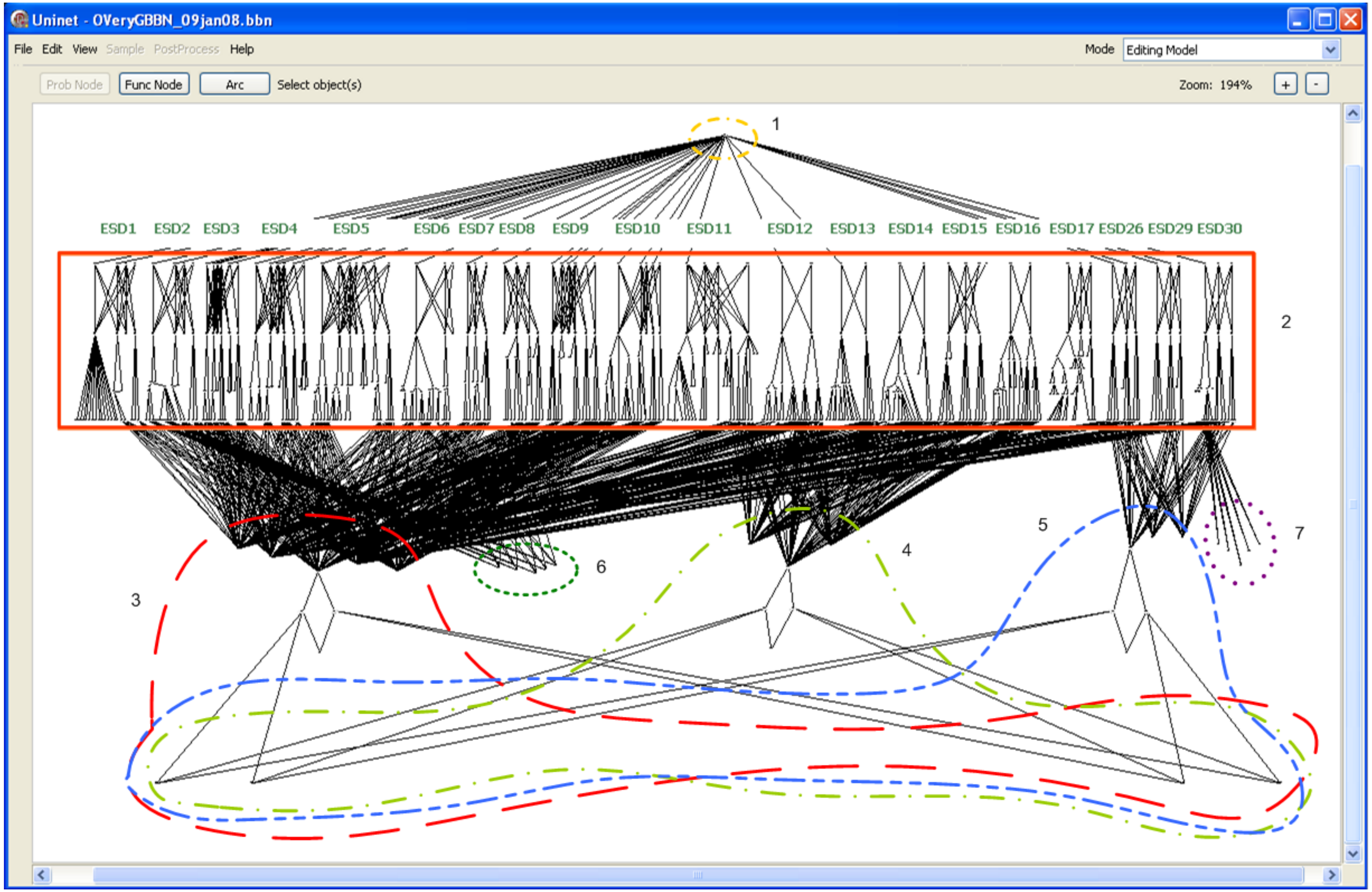


Figure 4.10: 1 - An Accident/Inicident node, 2 - Boolean functions representing ESDs and FTs, 3 - FC performance model for take-off, 4 - FC performance model for en route, 5 - FC performance model for approach-landing, 6 ATC performance model for take-off, 7 - ATC performance model for approach and landing.

- ESD 1 - Aircraft system failure,
- ESD 2 - ATC event,
- ESD 3 - Aircraft handling by flight crew inappropriate,
- ESD 4 - Aircraft directional control related system failure,
- ESD 5 - Incorrect configuration,
- ESD 6 - Aircraft takes off with contaminated wing,
- ESD 7 - Aircraft weight and balance outside limits during take-off,
- ESD 8 - Aircraft encounters a performance decreasing wind shear after rotation,
- ESD 9 - Single engine failure during take-off,
- ESD 10 - Pitch control problems,
- ESD 11 - Fire onboard aircraft,
- ESD 12 - Flight crew spatially disoriented,
- ESD 13 - Flight control system failure,
- ESD 14 - Flight crew incapacitation,
- ESD 15 - Anti-ice/de-ice system not operating,
- ESD 16 - Flight instrument failure,
- ESD 17 - Aircraft encounters adverse weather,
- ESD 26 - Aircraft handling by flight crew during landing roll inappropriate,
- ESD 29 - Thrust reverser failure,
- ESD 30 - Aircraft encounters unexpected wind.

The first aim was to add all missing ESDs to the model. The 11 ESDs (out of 31 generic accidents scenarios) that were added to the model are:

- ESD 18 - Single engine failure in flight,
- ESD 19 - Unstable approach,
- ESD 21 - Aircraft weight and balance outside limits during approach,
- ESD 23 - Aircraft encounters wind shear during approach,
- ESD 25 - Aircraft handling by flight crew during flare inappropriate,

- ESD 27 - Aircraft directional control related system failure during landing,
- ESD 28 - Single engine failure during landing,
- ESD 31 - Aircraft are positioned on collision course,
- ESD 32 - Incorrect presence on runway in use,
- ESD 33 - Cracks in aircraft pressure cabin,
- ESD 35 - Flight crew decision error/operation of equipment error (CFIT).

After we add the missing ESDs to the model, the number of arcs increased by 731, the number of functional nodes by 337 and the number of probabilistic nodes by 300.

The next aim is to add all arcs which connect the end event(s) with Accidents/Incidents node on top of the BBN. We specified as well the Accident/Incident node for flight phases: take-off, en route and approach and landing and connect them to end event(s). In the bottom part we have connected the probabilistic nodes to FC performance model for: take-off, en route and approach and landing as it can be seen in figure 4.11. 759 arcs were added to the model. After these implementations the model consisted of 2969 arcs and 1288 nodes. From these nodes, 671 are functional nodes representing ESDs and FTs as boolean functions, and 617 are probabilistic nodes.

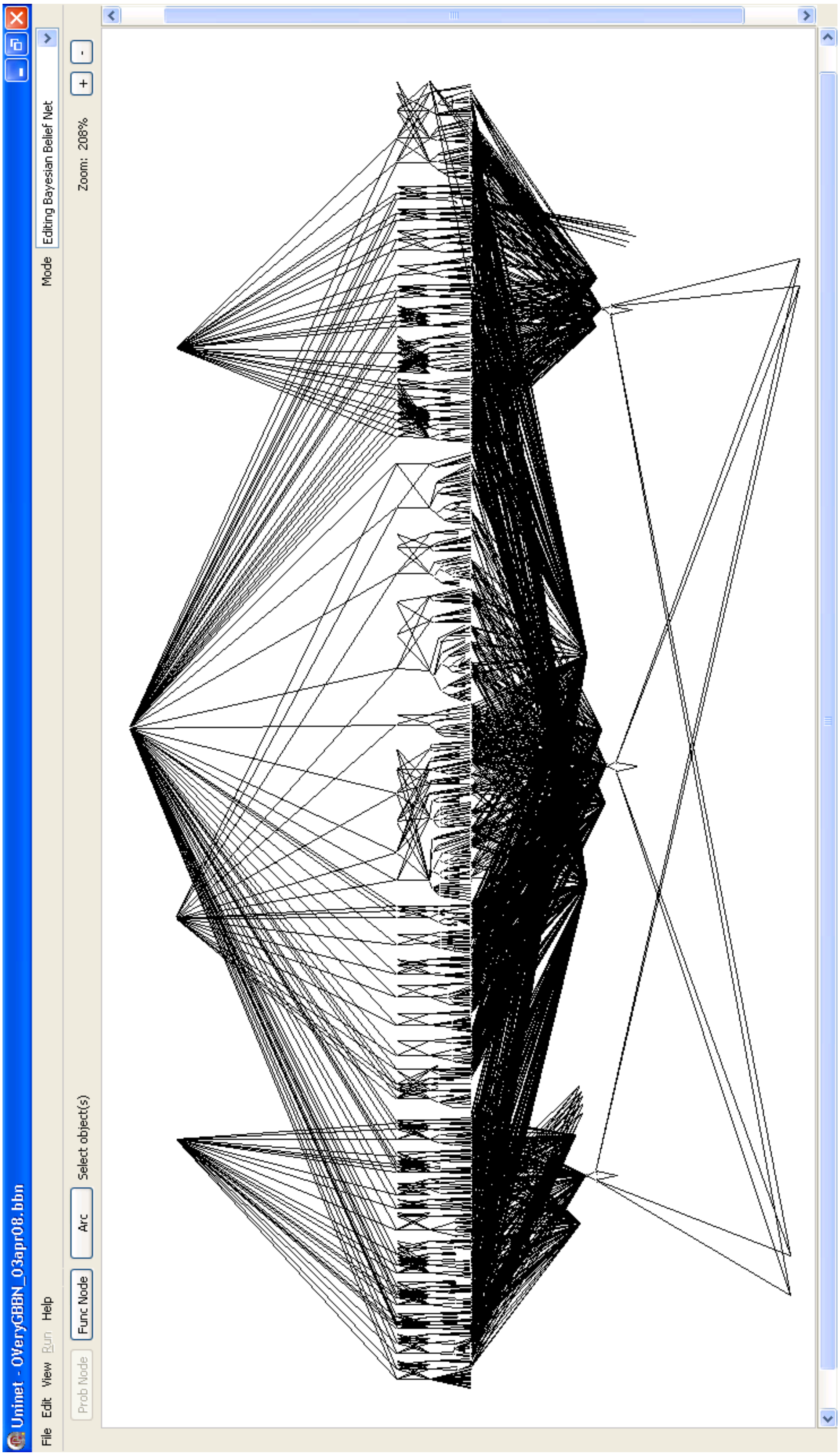


Figure 4.11: The CATS model.

We grouped all ESDs which belong to take-off, en route and approach and landing to make the structure of this BBN more readable. The breakdown groups are listed below.

- Take-off: ESD 1, ESD 2, ESD 3, ESD 4, ESD 5, ESD 6, ESD 7, ESD 8, ESD 9, ESD 10,
- En route: ESD 11, ESD 12, ESD 13, ESD 14, ESD 15, ESD 16, ESD 17, ESD 18, ESD 21, ESD 31, ESD 32, ESD 33,
- Approach and landing: ESD 19, ESD 23, ESD 23, ESD 26, ESD 27, ESD 28, ESD 29, ESD 30, ESD 35.

At this stage, several implementations and changes were done in the model. For example, the accident/incident node representing en route flight phase has been delated from the model. Besides that, some specific accident nodes, such as: runway overrun, runway veer-off, collision with ground, collision in mid-air, collision on runway, aircraft damaged and etc., have been added to the model. Some of those nodes are representing all flight phases, e.g. collision with ground. Some of them are specific to each of the flight phase, e.g. fire in flight in en route, collision on runway in en route and etc. Some of accidents are represented in two flight phases, e.g. runway overrun in take-off and approach and landing, aircraft damaged in take-off and approach and landing and etc. These changes were done to gain a better understanding of different types of accidents in CATS model.

Throughout the construction of the model, DNV's Excel file was updated (for example, more quantiles were added to that file). As the reason of this update, the structure of BBN suffered minor changes. For example, ESD 32 (Incorrect presence of aircraft/vehicle on runway in use) changed and was split into two flight phases - take-off and approach and landing (previously it has been linked to the en route flight phase exclusively). Some of the functional nodes were changed to probabilistic nodes.

In the bottom part of the BBN the last of human reliability models, namely the Maintenance Technician model has been added representing all phases during the flight. Moreover, the Total Transition Time node has been added to the BBN. This node is connecting to models: ATC performance model and FC performance model in approach and landing and en-route flight phases.

After these changes, the model consists of 4745 arcs and 1366 nodes. From these nodes, 532 are functional nodes representing ESDs and FTs as boolean functions and 834 are probabilistic nodes. Figure 4.12 presents the latest version of the model from 27 June 2008. We have to be aware of the fact that the presented model is not a final one, it is still under development and in the future may vary from this one.

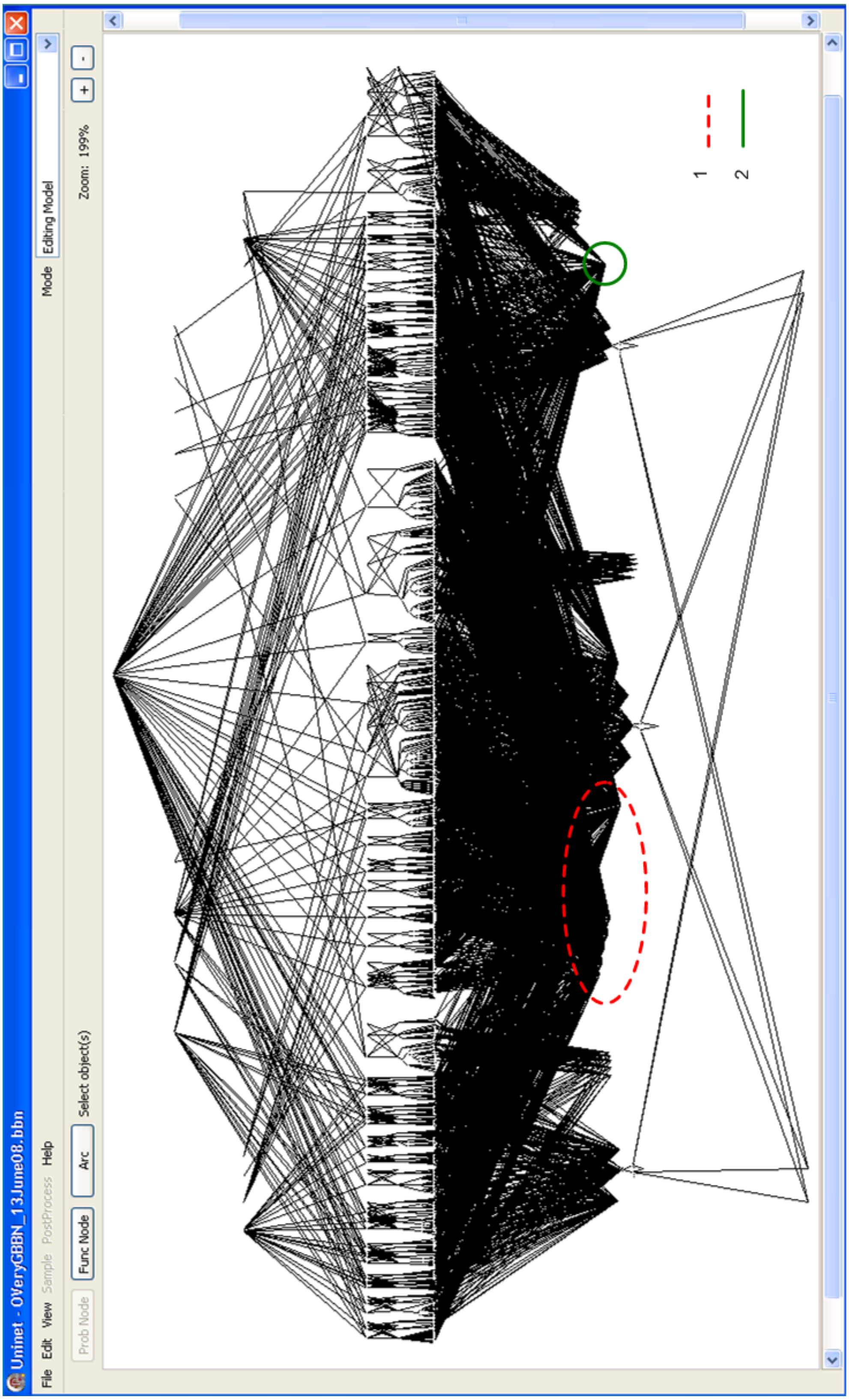


Figure 4.12: The CATS model, where: 1 - Maintenance Technician Performance Model and 2 - The Total Transmission Time.

Next, we present some summary figures describing the progress in the construction of the CATS model.

The complete model in figures

In the table below we present summary of arcs and nodes which have been added to the model in different stages.

	# of functional nodes	# of probabilistic nodes	# of arcs
1	332	317	1479
2	669	617	2210
3	671	617	2969
4	532	834	4745

Table 4.1: Number of nodes and arcs in individual stages.

In table 4.1, 1 corresponds to the period when 20 ESDs were added to the model, 2 corresponds to the period when all ESDs were added to the model, 3 corresponds to the time when the bottom part of BBN was connected with the FC performance model and the top part was connected with accident node in flight phases and general accident node, and 4 correspond to 27 June 2008 when the latest model was updated. As stated above the model presented in figure 4.12 is not final. Two notorious changes which the model will undergo in short are the addition of 2 missing ESDs: ESD 36 and ESD 37.

4.3 Procedure to obtain the underlying error distribution

Once the generic models for human reliability were available the marginal distribution of each basic event (top node in each of the three models previously described) had to be obtained. A central estimate (average) of the probability of occurrence of a human error has been computed using Fault Trees by DNV. Estimates of the "variability" of the central estimates have also been obtained by them. Firstly, our main source of information about the distribution of errors (representing base events in the FTs) was in the form of values for 5%, 50% and 95% - quantiles. This information was included in the Excel file (**080214 DNV Collected Fault Trees v5_1.xls**) and provided by DNV. Since the data about distribution of errors have changed, our main source of information about these distributions have been enhanced by the values for 10%, 25%, 75%, 90%, 99% - quantiles and minimum and maximum value of the base events. Finally, in the procedure of finding error distribution we used the Excel file (**080523 DNV Collected Fault Trees v6_1.xls**) containing the form of values for 5%, 10%, 25%, 50%, 75%, 90%, 95%, 99% - quantiles and minimum and maximum value of base events. In figure 4.13, we present the percentiles of the FTs base event error distribution from DNV.

Microsoft Excel - 080523 DNV Collected Fault Trees v6_1.xls

CAUSAL MODEL FOR AIR TRANSPORT SAFETY
Det Norske Veritas for CATS Consortium and Ministerie van Verkeer en W
STORE OF MODIFICATION FACTORS AND VARIABILITY OF BASE EV
MF defined as Case-specific base event prob/Generic base event prob
Distribution of variability in cols DP to DY

CALCULATED COMBINED UNCERTAINTY AND VARIABILITY (assuming all influences are independent)

	CODE	NO	EVENT NAME	Minimum	5%ile	10%ile	25%ile	50%ile	75%ile	90%ile	95%ile	99%ile	Maximum
670	ER33B11111	662	Manufacturing Inadequate	0.071	0.175	0.270	0.454	0.817	1.867	3.491	14.328	85.134	91.534
671	ER33B11112	663	Wear & Tear	0.071	0.175	0.270	0.454	0.817	1.867	3.491	14.328	85.134	91.534
672	ER33B11121	664	Routine Inspection Failure	0.129	0.284	0.325	0.509	0.781	1.521	2.418	2.418	2.418	2.418
673	ER33B11122	665	Routine Repair Failure	0.018	0.079	0.151	0.398	0.898	1.791	4.652	8.255	8.255	8.255
674	ER33B1113	666	Deterioration likely to propagate	0.932	0.932	0.945	0.971	1.000	1.001	1.001	1.001	1.001	1.001
675	ER33B112111	667	Bird strike	0.310	0.380	0.533	0.775	1.192	1.938	2.979	3.729	7.614	9.321
676	ER33B112112	668	Collision whilst on ground	0.075	0.255	0.340	0.638	1.116	2.279	4.653	50.216	256.309	313.798
677	ER33B112113	669	Tail strike	0.087	0.297	0.398	0.742	1.160	1.855	3.616	24.485	124.977	153.009
678	ER33B112121	670	Post Incident Inspection Failure	0.020	0.088	0.166	0.403	0.910	1.907	4.452	7.792	60.000	60.000
679	ER33B112122	671	Post Incident Repair Failure	0.088	0.203	0.310	0.455	0.863	1.651	3.173	17.754	29.500	29.500
680	ER33B11213	672	Damage likely to propagate	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
681	ER33B11221	673	Pressure Boundary likely to fail on next flight	0.075	0.075	0.152	0.397	0.871	2.064	2.356	2.356	2.356	2.356
682	ER33B11222	674	Subsequent Inspection Failure	0.086	0.198	0.304	0.446	0.853	1.682	3.250	17.200	30.426	30.426
683	ER33B11223	675	Subsequent Repair Failure	0.000	0.000	0.000	0.000	0.114	1.259	6.620	10.204	161.449	331.247
684	ER33B121	676	Door Design Potentially Unsafe	0.263	0.269	0.269	0.413	0.606	2.170	3.345	52.931	52.931	52.931
685	ER33B1221	677	Design fault not known	0.020	0.088	0.166	0.403	0.910	1.907	4.452	7.792	27.129	27.129
686	ER33B12221	678	No safety directive issued	0.071	0.246	0.315	0.495	0.728	1.393	2.605	14.369	35.790	35.790
687	ER33B12222	679	Manufacturer ignores safety directive	0.000	0.000	0.000	0.001	0.158	1.543	7.660	12.503	197.820	349.383
688	ER33B12223	680	Airline ignores safety directive	0.000	0.000	0.000	0.001	0.158	1.543	7.660	12.503	197.820	349.383
689	ER33B1223	681	Modification inadequate	0.020	0.088	0.166	0.403	0.910	1.907	4.452	7.792	30.000	30.000
690	ER33B1231	682	Door damaged by personnel	0.000	0.000	0.000	0.000	0.015	2.602	5.338	5.338	11.879	12.834
691	ER33B1232	683	Door Operation Failure	0.277	0.393	0.627	0.898	1.116	1.277	1.900	2.483	2.973	3.071
692	ER33B1233	684	Separate fault causes failure	0.020	0.087	0.164	0.402	0.908	1.892	4.481	7.881	109.274	124.683
693	ER33B131	685	Aircraft doors not secured	0.853	0.853	0.853	0.853	0.853	1.311	1.898	2.051	2.051	2.051
694	ER33B1321	686	No Door Check	0.000	0.000	0.000	0.000	0.150	1.105	5.120	7.062	176.941	372.813
695	ER33B1322	687	Door Check Unsuccessful	0.182	0.396	0.398	0.603	0.772	1.473	1.937	10.180	10.180	10.180
696	ER33B1331	688	No Warning System	0.000	0.000	0.000	0.000	0.001	0.522	3.870	7.405	102.810	337.771
697	ER33B1332	689	Warning system failure	0.025	0.109	0.196	0.409	0.913	2.113	4.036	6.661	18.341	18.341
698	ER33B1333	690	Warning system inadequate	0.025	0.109	0.196	0.409	0.913	2.113	4.036	6.661	18.341	18.341
699	ER33B1334	691	Crew ignore warning system	0.000	0.000	0.000	0.000	0.177	1.238	5.735	7.540	183.412	183.412

FunInfMod / JudgeInfMod / ConsInfMod / InSum / MFStore / RRStore / BaseEventOut / FreqOut

Figure 4.13: Percentiles of the FT's base event error distribution from DNV.

To obtain the minimum value that the base event probability that event ER33B11111 can take, we do: $0.071 * E(\text{ER33B11111})$[7] where E(ER33B11111) denotes the expectation of ER33B11111. In the same way the maximum value of ER33B11111 should be equal to $91.534 * E(\text{ER33B11111})$. Other quantiles maybe obtained in the same way.This information is used to fit a parametric distribution to the data to represent the distribution over the base events probability. In total there are 756 base events representing 31 ESDs. With the percentiles provided by DNV in the document **080523 DNV Collected Fault Trees v6_1.xls** a minimally informative distribution (with respect to uniform measure) can be found. This distribution will be compatible with the percentiles provided by DNV. However, a parametric distribution could be fit to these data as well. The reasons to use a parametric distribution are as follow:

[7] The expectation is provided as the probability of occurrence of each event in the FTs. See, for example figure 4.8.

- It is easier to maintain model. The minimally informative distribution requires to store the whole distribution, while a parametric distribution only requires to store a number of parameters which describe the distribution.

- In general, the minimally informative distribution fitted to the quantiles provided by DNV will not preserve the expectation provided by DNV[8].

- We wanted to have the model with the functionality that by specifying a different mean than the one computed by DNV and keeping the variance constant, we could obtain a new distribution.

The procedure used to obtain the underlying error distribution is described below:

1. Obtain the minimally informative distribution (in log uniform scale to avoid negative values) that fits provided information. As it was mentioned already, this distribution is a distribution over the probability of error. Observe that the minimally informative distribution will always capture the percentiles provided by DNV, however, the minimally informative solution might give sometimes inconsistent results[9]. To meet these goals the following steps are taken:

2. Find the parameters of a Weibull or Gamma distribution that minimize the sum of squared difference between the minimally informative solution found in step 1 and the parametric distribution such that,

   a) The expectation of the parametric distribution is equal to the estimate provided by DNV and,

   b) The distribution lies in the interval (0,1),

   c) The 0.999999999% - quantile of distribution is less or equal than the maximum value provided by DNV.

3. This Parametric distribution will be used as an estimate of the error distribution in the CATS model.

In general, it was observed that Weibull and Gamma better characterized the data provided by DNV. After all, to improve the speed of finding parametric distribution, the procedure presented above was applied only to search Weibull and Gamma parameters to the 756 base events provided by DNV.

[8]This is because the expectation of the minimally informative distribution is determined by quantiles specified by DNV. While, the expectation provided by DNV is not determined by these quantiles, [16]

[9]In previous Excel files provided by DNV, for example part of the distribution might be outside the (0,1) interval, because the minimum and maximum of teh distribution were not specified. In general, this distribution would not have the expectation provided by DNV (as the probability of base events in FTs).

In the situation, where no convergence in step 2 was possible, we focus on searching the parameters for Beta and Log-normal distribution. Since the mean and maximum value of DNVs data entered as constrains in the optimization procedure, were better captured by the fitted distribution. In some cases, some fits may be far from DNVs data. Next, we present two examples of "good" and "bad" fit only for illustration purposes.

In figure 4.14 we present two examples showing different levels of convergence. One could say that the parametric distributions obtained with the procedure described above shown in the left present "better" convergence than those in the figure in the bottom.

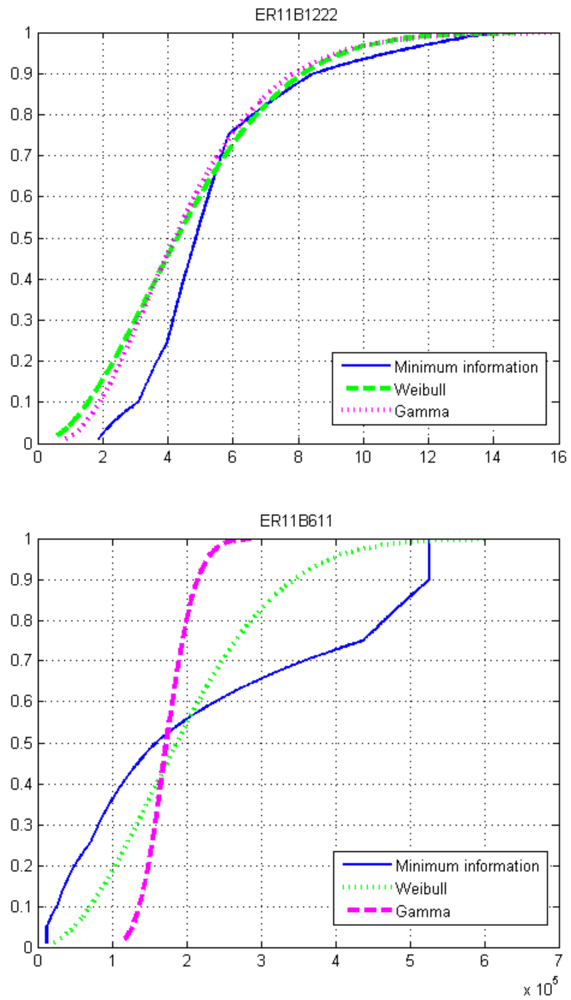


Figure 4.14: From the left: "better" example of convergence and "worst" example of convergence.

In each of the two plots in figure 4.14 we see 3 curves. The "solid" curve represents the minimum information solution, the "dashed" one a Weibull distribution and the "dotted" one represents a Gamma distribution. As we can see on the first figure (the top one), the Gamma and Weibull distributions are relatively close to the quantiles specified by DNV (comparing to the bottom figure). They capture the mean (which is equal to 4.63) and the variability of the minimum information solution relatively well. In this case the Gamma distribution converge in 26 iterations and the value of the sum of squared difference is 3.73871495e+002. The Weibull distribution converged in 22 iterations wit a value of the sum of squared difference is 4.53437414e+002.

On the second figure we see that the curves "dashed" and "dotted" do not capture well the variability of the minimum information solution as well as mean for Gamma distribution. Only the mean is well captured for Weibull distribution (which is equal to 2e5) which converged in 46 iterations . We can observe big difference in the values of the sum of the squared differences between the Gamma and Weibull distribution. The sum of squared differences for Weibull distribution is equal to 4.51882540e+012, whereas for Gamma distribution is equal to 14.7308225e+012. The example was taken from ESD 11. Once the parametric distribution corresponding to each base event in the FTs is obtained, one instance of the corresponding human reliability model is introduced in the CATS model. In the next chapter analysis of the full model as now stands will be performed.

Chapter 5

Analysis of the CATS BBN

In this chapter we present the analysis of the CATS BBN. During the analysis we use the version of the model from 27 June 2008. First, we present the analysis of the accident in take-off, en route and approach and landing with respect to the MT performance model. The main reason of this choice is that accident node is of prime importance at this stage of the analysis and MT performance model has been elaborated in the context of this thesis. In general, we would like to check if there is any influence between accident probability and contribution factors representing MT model and how strong this relation is. First, we will conditionalize on the contribution factors and check the differences in the mean of accident probability. Later on, we perform the sensitivity analysis to point out the most influential contribution factors on accident probability.

In the next sections, we will present the sensitivity analysis. First, the sensitivity analysis of FC performance model, ATC performance model and MT performance model with respect to the overall accident will be presented. Next, the same analysis will be presented with respect to the accident in take-off. Also, the sensitivity analysis of base event representing human errors on accident node in take-off, en route and approach and landing flight phases and other different accident scenarios will be performed. The base events used in the sensitivity analysis are those from the ESD 30 - “Aircraft encounters unexpected wind”.

5.1 Conditional and Unconditional distribution of accident probability

In this section, we will present and use in the analysis two ways to perform the conditioning on the contribution factors. Firstly, we can perform sampling-based conditioning, which can be applied to both probabilistic and functional nodes. Secondly, we can perform analytical conditioning (as we did in Chapter 3), which can be applied only to probabilistic nodes. One of the differences between them is that using analytical conditioning we can conditionalize on single values of the conditioning nodes, whereas sample-based conditioning allows us to conditionalize on intervals. Mostly, we are interested in conditioning on intervals and that is why we chose sampling-based conditioning. The interest is in checking the differences between the means of the conditional and unconditional distributions. For example, we would like to check how the mean of the unconditional distribution of the output variable will change when we conditionalize on the group of the maintenance technicians who just start their job as the maintenance worker and maintenance technicians who already worked in current position for some years.

In table 5.1 we present results of sampling based-conditioning. To perform sampling based-conditioning we used 100000 samples. The mean of the accident probability node[1] representing take-off, en route and approach and landing after sampling is equal to 3.2319e-6.

Condition	**Conditional mean**	**Conditional standard deviation**	**Ratio**[a]
Fatigue=[1,4]	3.1709e-6	2.6370e-5	0.9811
Fatigue=[5,7]	4.5367e-6	2.4659e-5	1.4037
Fatigue=[1,2]	3.0358e-6	2.7163e-5	0.9393
Fatigue=[6,7]	4.6513e-6	2.5615e-5	1.4392
AirGen=[1,2]	1.4974e-5	7.0090e-5	4.6332
AirGen=[3,4]	2.4370e-6	1.9896e-5	0.7540
AirGen=[2,3]	3.2852e-6	2.6216e-5	1.0165
Exp=[1,5]	5.7895e-6	2.4455e-6	1.7914
Exp=[8,14]	3.2620e-6	2.2542e-5	1.0093
Exp=[18,25]	2.8989e-6	2.9033e-5	0.8970
Workload=[0.1666,2]	3.0847e-6	2.5768e-5	0.9545
Workload=[3,7]	3.4161e-6	2.8480e-5	1.0570
Workload=[9,16]	3.8919e-6	2.9406e-5	1.2045
ShiftOverTime=[0.5,5]	3.3042e-6	3.0293e-5	1.0224
ShiftOverTime=[7,13]	3.1172e-6	2.5944e-5	0.9645
ShiftOverTime=[17,22]	2.6913e-6	1.3347e-5	0.8327
# of samples = 100000			
Mean and standard deviation of the accident node after sampling			
mean = 3.2319e-6			
stdDev = 2.6298e-5			

Table 5.1: Sample based-conditioning on contribution factors with respect to accident node.

[a] Ratio=$\frac{\text{conditional mean}}{\text{unconditional mean}}$

[1] See figure 4.10

The largest difference in the conditional distribution of accident as measured by the ratio of conditional mean to unconditional mean is observed in aircraft generation. When we conditionalize on AirGen=[1,2] (1 and 2 are the oldest generations of the aircraft), the conditional mean increased from 3.2319e-6 to 1.4974e-5. In chapter 3 it was observed that experience was the most influential factor in maintenance error. However its importance is not as notorious in the overall accident probability as aircraft generation is. Indeed, when we conditionalize on Exp=[1,5], we see that the mean of accident node increased to 2.5576e-6. This is due to the fact that the aircraft generation is also a node in the FC performance model. Hence its influence in the overall accident probability is not only through MT performance but also through the FC error.

It is worth mentioning that the lowest conditional mean is obtained when we conditionalize on the most recent generation of the aircraft (3 and 4). The ratio between conditional mean and unconditional mean is equal to 0.7540 and it is relatively small comparing to other ratios of conditional and unconditional means[2]. It means that the accident probability decreases when maintenance technicians have to maintain the most recent generation of aircraft. This point out a possible performance of maintenance technicians to work on newer aircraft.

After we obtained these results perform also analytical conditioning and conditionalize on single values of the contribution factors. In table 5.2 we present the analytical conditioning for different contribution factors.

[2]For comparison the ratio of the smallest percentile to the mean in the unconditional distribution of accident probability is equal to 0.0326, and the ratio of the the largest percentile to the mean of accident probability is 550.8874

Condition	Conditional mean	Conditional standard deviation	Ratio
Fatigue=1	2.9729e-6	2.7093e-5	0.9360
Fatigue=6	4.6383e-6	3.2694e-5	1.4604
AirGen=1	5.1411e-5	1.9272e-4	6.1779
AirGen=2	8.9990e-6	4.8716e-5	2.8333
AirGen=3	5.1091e-7	2.2270e-6	0.1609
AirGen=4	2.4494e-7	5.5711e-7	0.0771
Exp=1	9.9652e-6	4.9652e-6	3.1376
Exp=3	5.9813e-6	3.6873e-5	1.8832
Exp=30	2.5569e-6	2.5379e-5	0.8050
Workload=10min	2.9128e-6	2.6701e-5	0.9171
Workload=30min	3.0238e-6	2.7108e-5	0.9520
Workload=22hrs	4.0257e-6	3.0403e-5	1.2675
ShiftOverTime=1min	3.1656e-6	2.7570e-5	0.9967
ShiftOverTime=22min	2.9290e-6	2.6731e-5	0.9222
WorkCond=1	3.2370e-6	2.7796e-5	1.0192
WorkCond=2	3.0135e-6	2.7022e-5	0.9488

# of samples = 100000
Mean and standard deviation of the accident node after analytical conditioning
mean = 3.1761e-6
stdDev = 2.7587e-5

Table 5.2: Analytical conditioning of the contribution factors with respect to accident node.

When we conditionalized on single values of aircraft generation we may see how big influence the type of the aircraft has on the accident node. The first generation of the aircraft represents the class of the aircraft designed in the 1950s. From table 5.2 we can see that when we conditionalize on this class of aircraft generation the conditional mean increases from 3.1761e-6 to 5.1411e-5. On the other hand, when we conditionalize on the fourth generation of the aircraft, the most recent, we see that the ratio between conditional and unconditional mean is equal to 0.0771. Observe that the third generation of aircraft contains most of the aircraft (90.78%). The ratio between the conditional mean and unconditional mean of the accident node is equal to 0.1609. This node reflects the belief in the industry that technological advances have made the industry more safe over the years. We may also speculate that experts consider the interaction between men and technology fundamental in risk mitigation.

Next, we will present a graphical interpretation of the accident node before and after conditioning on first and third generation of the aircraft. Figure 5.1 presents the empirical cumulative distribution function of the accident node (the left one is normal scale and the right one is zoom in to see better the difference between the empirical cumulative function before and after analytical conditioning on AirGen=1).

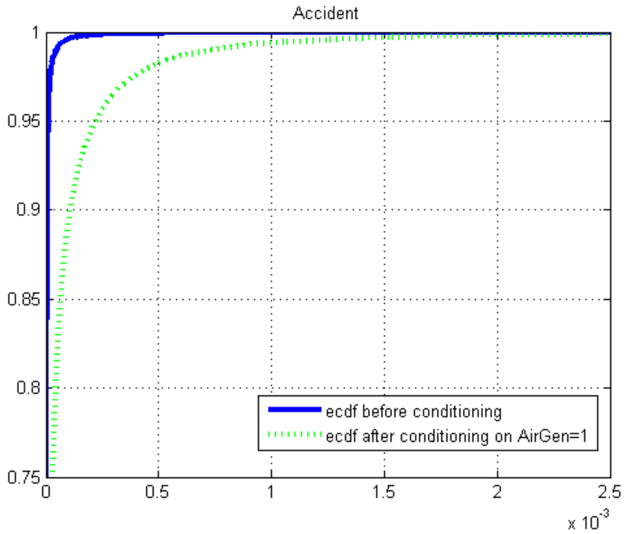


Figure 5.1: The empirical cumulative distribution function of accident node before conditioning (solid line) and after conditioning on AirGen=1 (dotted line).

In figure 5.1, we see two plots. The “solid” line represents the empirical cumulative distribution function of the accident node before performing the conditioning and the “dotted” line represents the empirical cumulative function of the accident node after performing the conditioning on the AirGen=1. We can notice the differences between these two curves. The conditional probability of the accident is smaller than the probability when no conditioning is performed. The biggest differences are observed after the 85th percentile of the distribution.

In figure 5.2 we present the relationship between the empirical cumulative distribution function of the accident node when no conditioning is performed and the empirical cumulative distribution function after conditioning on AirGen=3. Similar explanation of the curves is provided here.

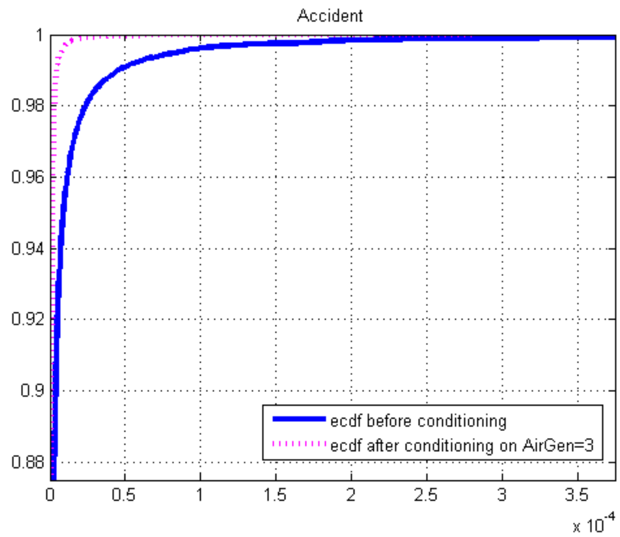


Figure 5.2: The empirical cumulative distribution function of accident node before conditioning (solid line) and after conditioning on AirGen=3 (dotted line).

As we can see in figure above, the conditional probability of the accident node is bigger than the probability of accident node without conditioning. The biggest differences are observed after the 96th percentile of the distribution.

Now, we would like to check how the probability of accident representing all flight phases will change when we conditionalize not only on the oldest class of aircraft (AirGen=1) but also on low values of maintenance technician experience (Exp=1 year). Figure 5.3 presents the relationship between the empirical cumulative distribution function of the accident node without conditioning ("solid" curve) and the empirical cumulative function of the accident node after the conditioning on AirGen=1 and Exp=1 was performed ("dotted" curve).

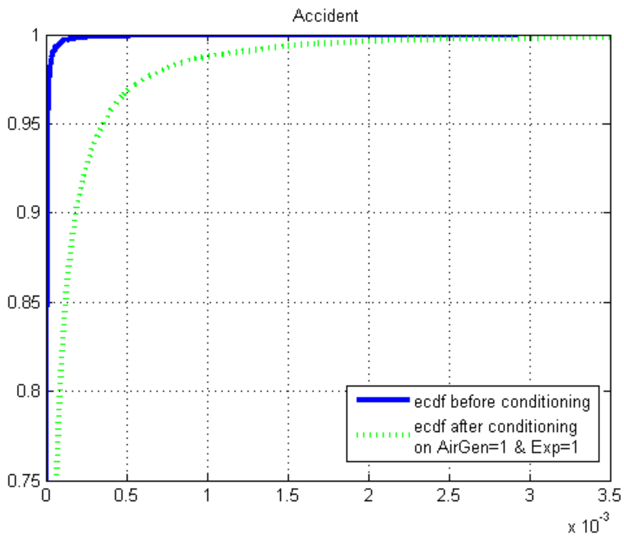


Figure 5.3: The empirical cumulative distribution function of accident node before conditioning (solid line) and after conditioning on AirGen=1 and Exp=1 year (dotted line).

As we can see in figure above, the conditional probability of the accident node is bigger than the probability of accident node without conditioning. The biggest differences are observed after the 85th percentile of the distribution. We want to mention that the mean of accident node increased from 3.1761e-6 to 9.0749e-5.

From the data we know that 0.08% of the aircraft represent the oldest generation of aircraft. To make sure that the aircraft generation is influencing the accident node the most, we will do the sensitivity analysis and point out the most influential contribution factors.

5.2 Sensitivity analysis of MT performance on accident probability

Figure 5.4 presents the sensitivity analysis of the predicted variable accident in take-off, en route and approach and landing (OUT_TOERALAccident)[3] with respect to the base variables representing the MT performance model. In the analysis we drawn 100000 samples.

The sensitivity analysis confirms that the aircraft generation has the strongest influence on the accident node from the MT performance model. The next, big influence on the accident node is that of experience. The difference between the correlation ratio of the aircraft generation and other contribution factors is signif-

[3]See figure 4.10.

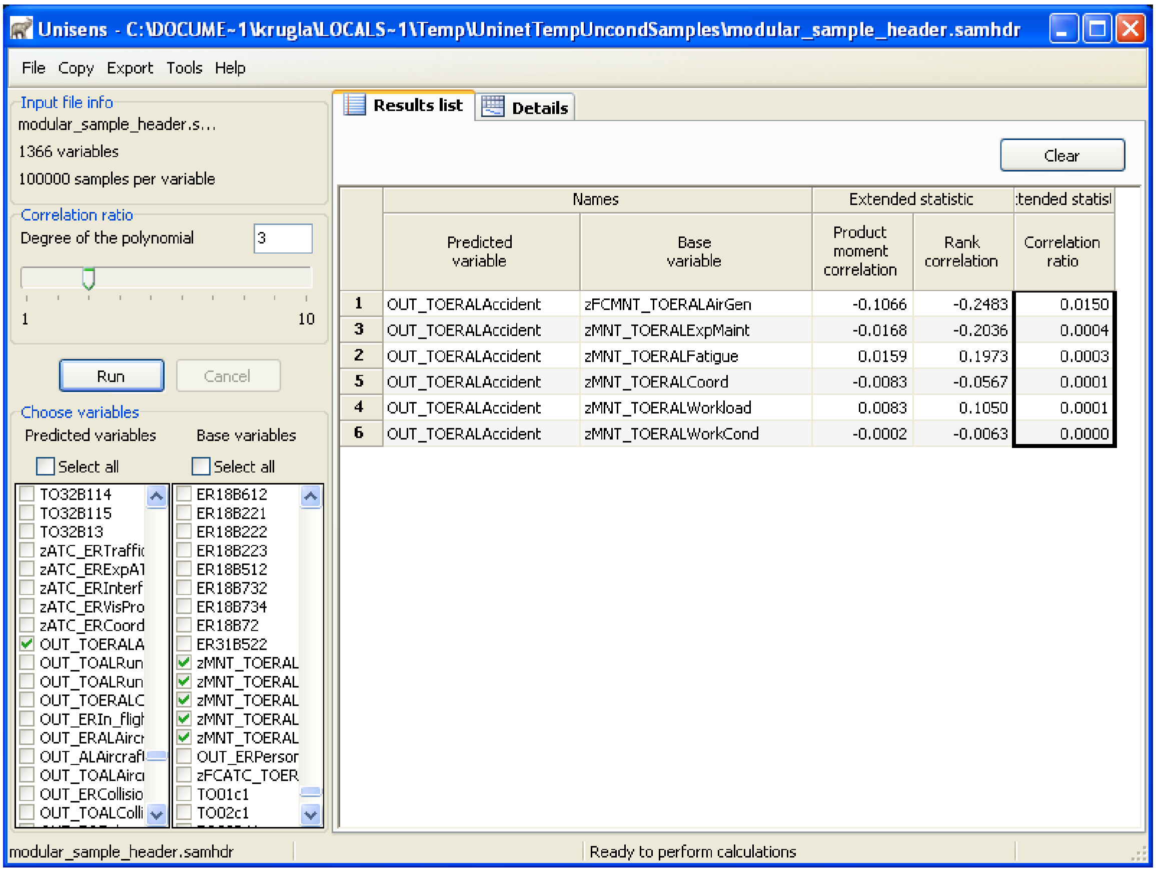


Figure 5.4: Sensitivity analysis of predicted variable OUT_TOERALAccident with respect to base variables from MT performance model.

icant. For example, correlation ratio of the aircraft generation is equal to 0.0150 while the correlation ratio of the experience is equal to 0.0004. This is explained as mentioned earlier because the aircraft generation is an influencing factor in both the MT and FC performance models[4].

Now, we would like to check which one of the contribution factors representing the human reliability models has the strongest influence on the accident node. To check this we will perform sensitivity analysis of the predicted variable, accident in take-off, en-route and approach and landing, with respect to the base variables representing the contribution factors from FC model, ATC model and MT model. Also, in this sensitivity analysis 100000 samples were used. Figure 5.8 presents result of these analysis.

[4]Sensitivity analysis was also done using 50000 samples. The results were similar and the order of the most influential base variables on the predicted variable, OUT_TOERALAccident is the same as when using 100000 samples. That is, aircraft generation has the strongest influence on the accident node, then experience, etc. We noticed only the differences in the absolute value of the correlation ratio, for example, aircraft generation has correlation ratio equal to 0.0365 while the correlation ratio of the experience is equal to 0.007. As we can see the difference between these values are about a factor two.

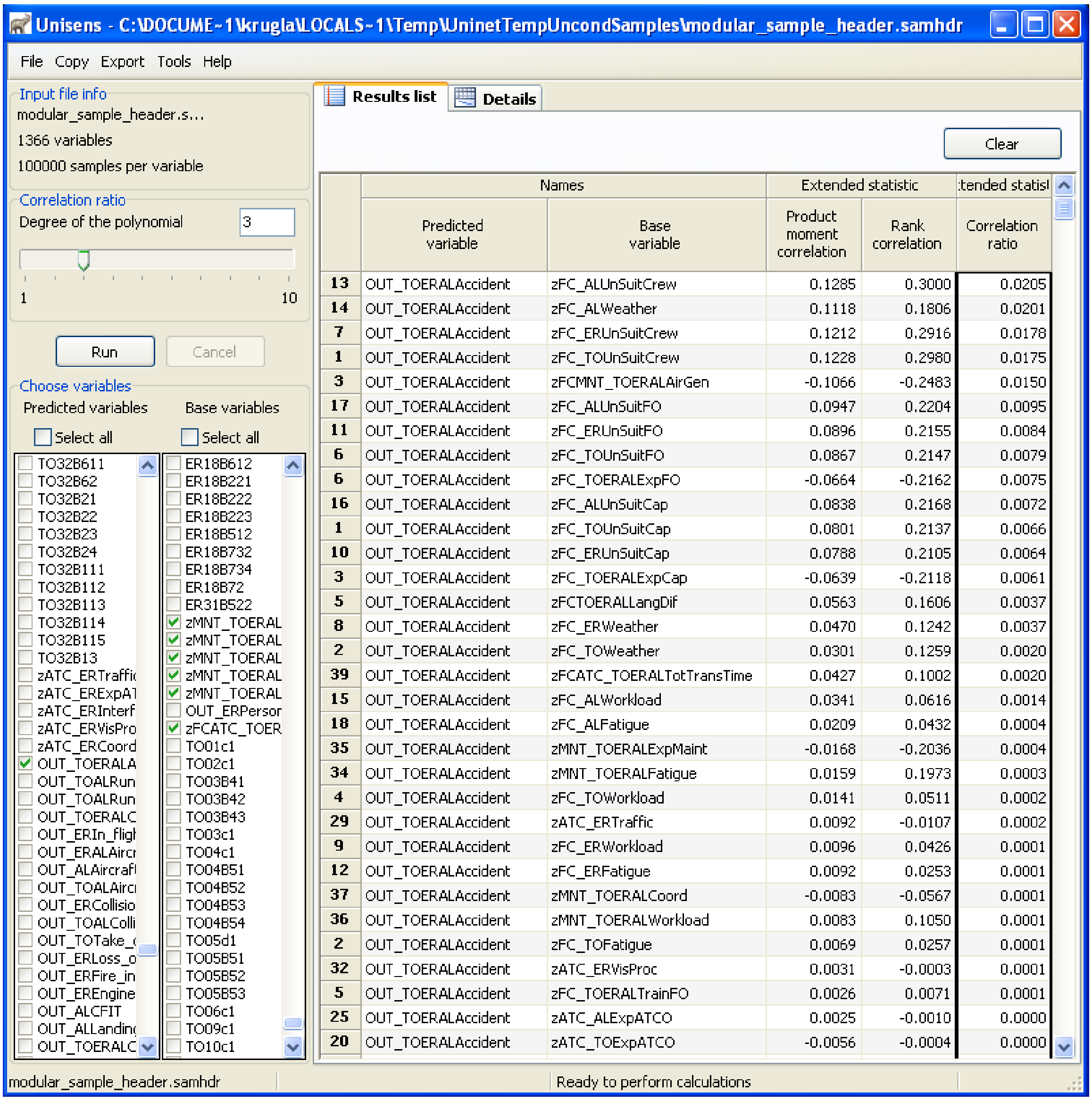


Figure 5.5: Sensitivity analysis of predicted variable OUT_TOERALAccident with respect to base variables representing the human reliability models.

As we can observe, the crew unsuitability at approach and landing flight phase has the strongest influence on the accident node. However, the correlation ratio is small and equal to 0.0205. We can also notice the fact that the most influential factors on the predicted variable are those representing the FC performance model. The next variable which has the biggest influence on the accident node is a weather during the approach and landing, a crew unsuitability representing en route and take-off flight phase, then is the aircraft generation, first officer unsuitability for approach and landing, en route and take-off flight phases, and etc.

In general the correlation ratio values for all variables are small. For example the ratio of highest to lowest (different than zero) is 205. The first 13 variables ranked according to their correlation ratio value have comparable values for their absolute rank correlation. However, their difference in terms of correlation ratio is up to a factor 3.36. Significant differences may also be observed in terms of product moment correlations and rank correlations for each variable to the accident probability indicating a higher degree of monotonic relation rather than simply linear.

In general, it may appear that the human reliability models do not have such a significant influence on the accident in three flight phases. However, this conclusion is also determined by the accident probability distribution which by conditionalizing on a single value for a single conditioning variable (as in table 5.1) the difference in means between conditional and unconditional accident probabilities may be a factor 6 in absolute value. We may conclude that the “most” influential factor is the crew unsuitability at approach and landing flight phase and other factors representing the FC performance model and aircraft generation representing FC performance model and MT performance model.

5.3 Sensitivity analysis of human performance models on TO accident probability

Sensitivity analysis for the predicted variable OUT_TOTake_off with respect to contribution factors representing FC performance model, ATC performance model and MT performance model was also performed. In this sensitivity analysis 100000 samples were used. Figure 5.6 presents the result of these analysis.

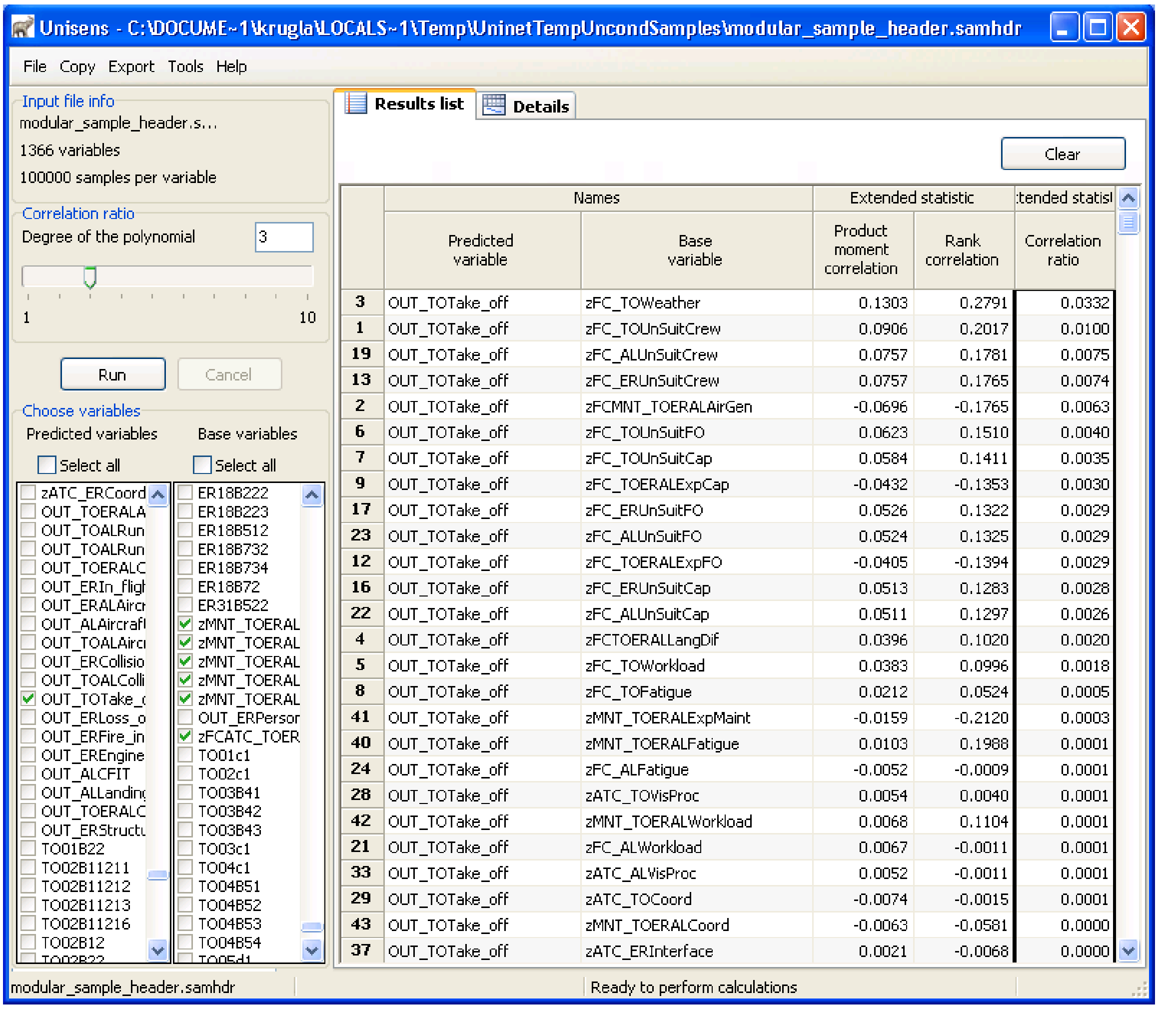


Figure 5.6: Sensitivity analysis of predicted variable OUT_TOTake_off with respect to base variables representing the human reliability models.

As we can observe, the weather during take-off flight phase has the strongest influence on the take-off node. However, the correlation ratio is low and equal to 0.0332. As in the previous sensitivity analysis of the accident node with respect the same contribution factors, we can also notice that the most influential factors on the predicted variable are those representing FC performance model. The next variables with the biggest influence on the take-off node are: crew unsuitability representing the take-off, approach and landing and en-route flight phase, then aircraft generation representing the FC performance model and MT performance model, first officer and captain unsuitability representing the take-off flight phase, and etc.

Also, in this sensitivity analysis the correlation ratio values for all variables are small. The ratio of the highest to lowest (different than zero) in this situation is equal to 332. We can notice that the first 6 variables, ranked according to the correlation ratio value, have rank correlations that differ a factor 1.8 at most. However, the difference between their correlation ratio is up to a factor 8.3. Similarly to the accident node, we can observe significant differences in product moment correlation and rank correlation.

As a final remark, we can conclude that of the human reliability models those representing the influential factors of the FC performance model have the biggest influence on the take-off node. This pattern is similar when analyzing the overall accident probability. Similar kind of analysis as the one presented in this chapter may be conducted for other summary nodes.

In next sections, we will present the sensitivity analysis of base events representing FC errors, ATC errors, MT errors and non-human errors on accident node and as well as on other accident nodes. As an example we choose base events from ESD 30 - Aircraft encounters unexpected wind and is representing approach and landing flight phase. In table 5.3 we present the division of the base events (representing FC errors, ATC errors, MT errors and non-human errors) which will be used later on in the sensitivity analysis.

FC errors	ATC errors	MT errors	Non-human errors
AL30B213	AL30B211	AL30B42	AL30B111
AL30B214	AL30B212		AL30B112
AL30B32			AL30B12
AL30B33			AL30B22
AL30B34			AL30B31
AL30B43			AL30B41

Table 5.3: Division of the base events representing FC, ATC and MT errors and non-human errors.

As we can see from table 5.3, 40% of base events are representing the FC errors, 13.33% of them ATC errors, 6.67% are the MT errors and 40% of the base event are representing the non-human errors.

5.4 Sensitivity analysis of base events representing FC, ATC, MT and non-human errors on accident probability

Figure 5.7 presents the sensitivity analysis of the predicted variable OUT_TOERALAccident with respect to base events representing FC, ATC, MT and non-human errors. In this sensitivity analysis 100000 samples were used.

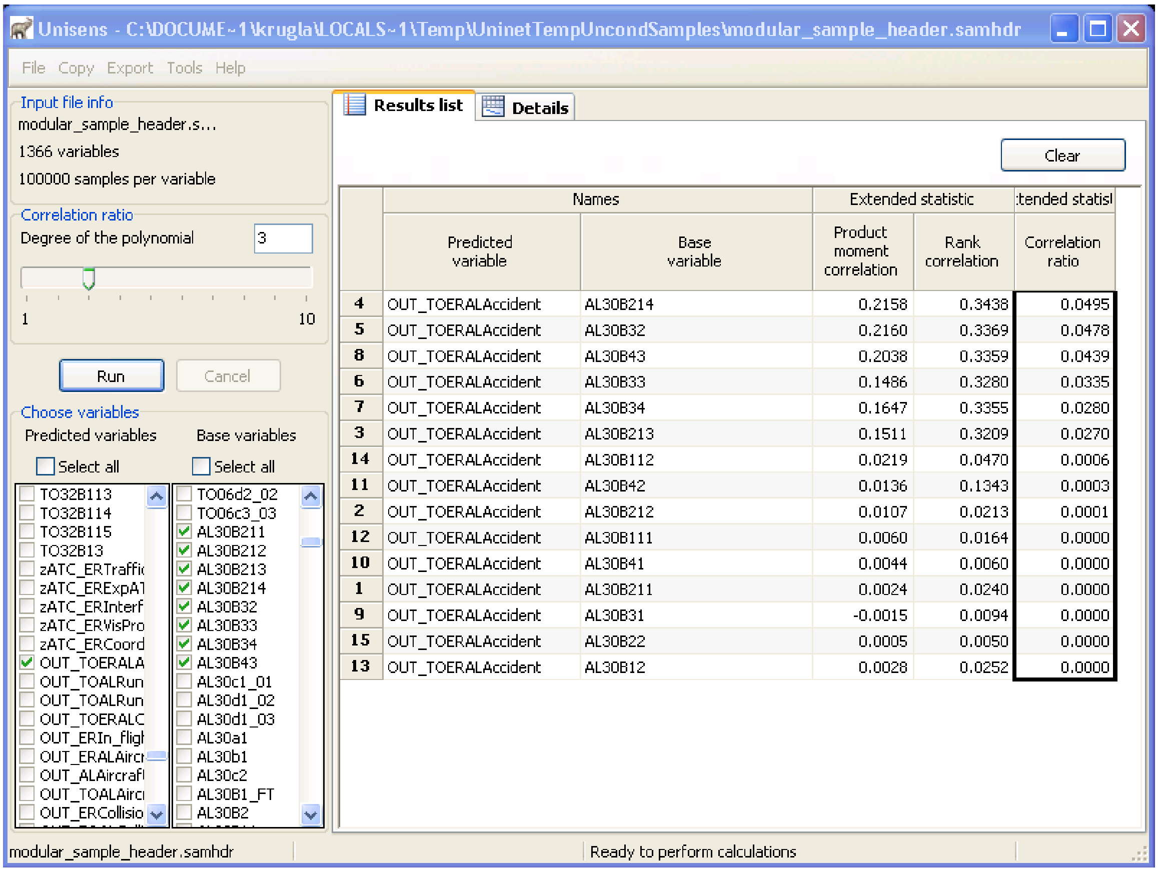


Figure 5.7: Sensitivity analysis of predicted variable OUT_TOERALAccident with respect to base events representing FC, ATC, MT and non-human errors.

As we can observe, the base event AL30B214 defined as "pilot disregards crosswind limit" and representing FC error has the strongest influence on the accident node. However, the correlation ratio is small and equal to 0.0495. The next, big influence on accident node has AL30B32 defined as "lack of control", AL30B43 defined as "brakes not applied correctly". Both of these base events are representing FC error. As we can see the correlation ratio of these base events does not vary significantly. The next three base events which has also big influence on accident node are those representing FC error. They are: AL30B33 defined as "incorrect control", AL30B34 - "insufficient control" and AL30B214 defined as "pilot fails to calculate wind correctly". The base variable AL30B42 defined as "brakes not functioning correctly" and representing MT error does not have a big influence on accident node. The correlation between this variable and accident node is low and equal to 0.0003. It means that the variance of the accident node is explained by 0.03% of variance of the conditional expectation of the accident node given AL30B42.

In general, the correlation ratio values are small for all variables. The ratio of highest to lowest (different than zero) is 495. The first 6 variables ranked according to their correlation ratio value have comparable values for their rank correlation and also for product moment correlation. However, their difference in terms of correlation ratio is up to a factor 1.8333. In general, one may see a differences between the product moment correlation and the rank correlation. The last being larger and indicating a higher degree of monotonic relationship rather than simply linear. As a final remark, we want to notice that base events representing the FC errors have the biggest influence on accident node representing take-off, en-route and approach and landing.

Now , we would like to check which one of events (initiating, pivotal, end or base) representing ESD 30 has the strongest influence on Accident node. To check this we will perform sensitivity analysis of OUT_TOERALAccident with respect to the events representing ESD 30. Also, in this sensitivity analysis we used 100000 samples. Figure 5.8 presents results of these analysis.

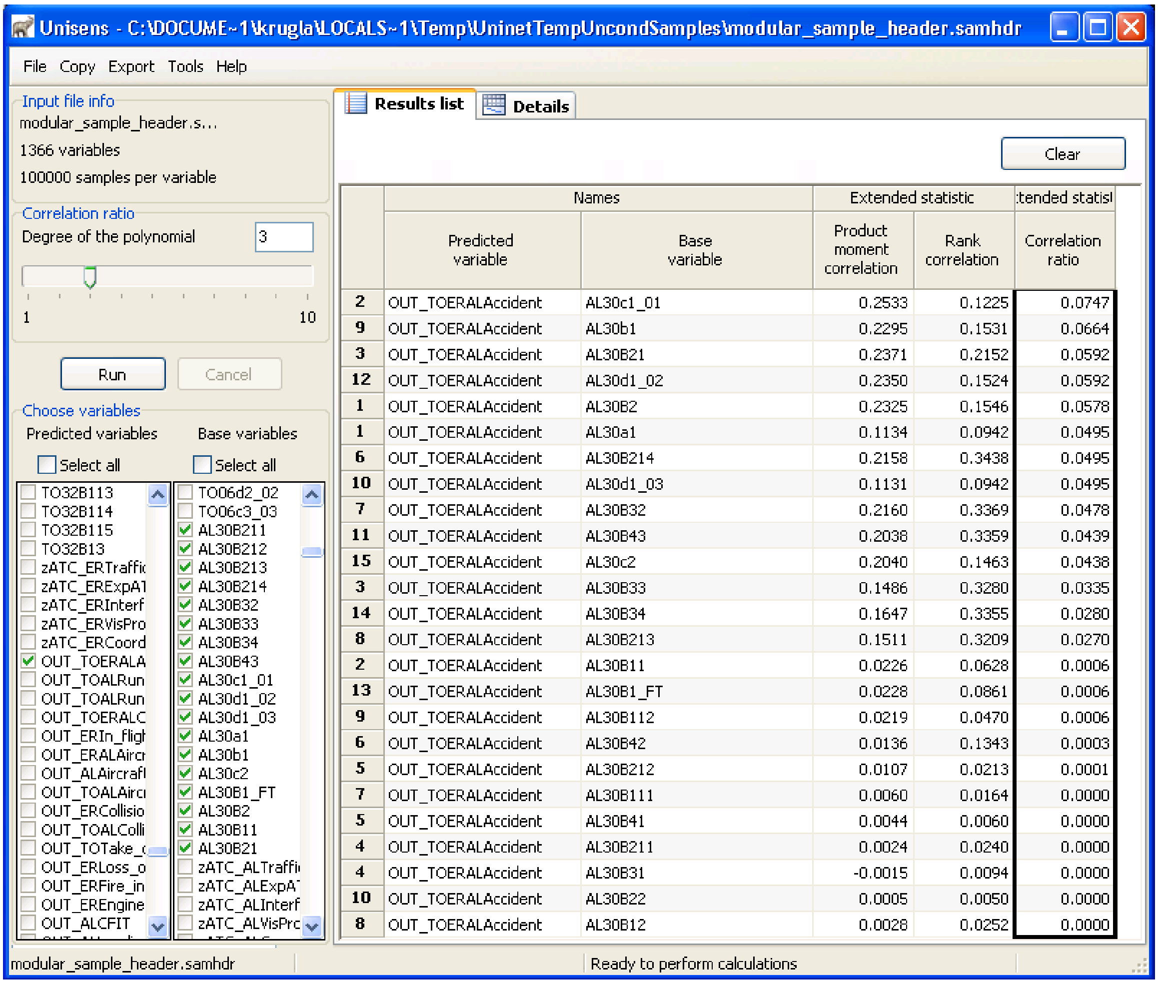


Figure 5.8: Sensitivity analysis of predicted variable OUT_TOERALAccident with respect to ESD 30.

The strongest influence on accident node has end state AL30c1_01 representing the runway veer-of. However, the correlation ratio is low and equal to 0.0747. It means that the variance of accident node is explained by 7.5% of the variance of runway veer-off. The next variable which has the biggest influence on accident node is the pivotal event, AL30b1 defined as "flight crew fails to maintain control". Then is the intermediate event, AL30B21 define as "failure to anticipate severe wind conditions", and the end state, AL30d1_02 representing runway overrun, which have the same correlation ratio. Other variable which has big influence on accident node is the intermediate event, AL30B2 define as "failure to avoid encounter" and 3 variables which have the same correlation ratio, i.e. the pivotal event (AL30a1 - aircraft encounters unexpected wind), the base event (AL30B214 representing the FC error) and end state (AL30d1_04 - aircraft continues landing roll) and etc.

Also in this sensitivity analysis the correlation ratio values for all variables are not high. The ratio between the highest and lowest correlation ratio (different than zero) is equal to 747. We can notice significant differences in rank correlation to the accident probability, whereas the differences in product moment correlation are not that significant.

In the next section, we will present the sensitivity analysis of base events representing FC errors, ATC errors, MT errors and non-human errors on accident node and other accident scenarios, e.g. runway veer-off, runway overrun, aircraft damaged, collision with ground and etc. Also, in this sensitivity analysis we will use base events representing ESD 30 - Aircraft encounters unexpected wind and is representing approach and landing flight phase and division of base events presented in table 5.3.

5.5 Sensitivity analysis of base events representing FC, ATC, MT and non-human errors on different accident scenarios probability

Sensitivity analysis for the different accident scenarios with respect to base events from ESD 30 representing FC, ATC, MT and non-human errors was performed. As in the previous analysis 100000 samples were used. Figure 5.9 presents the results of these analysis.

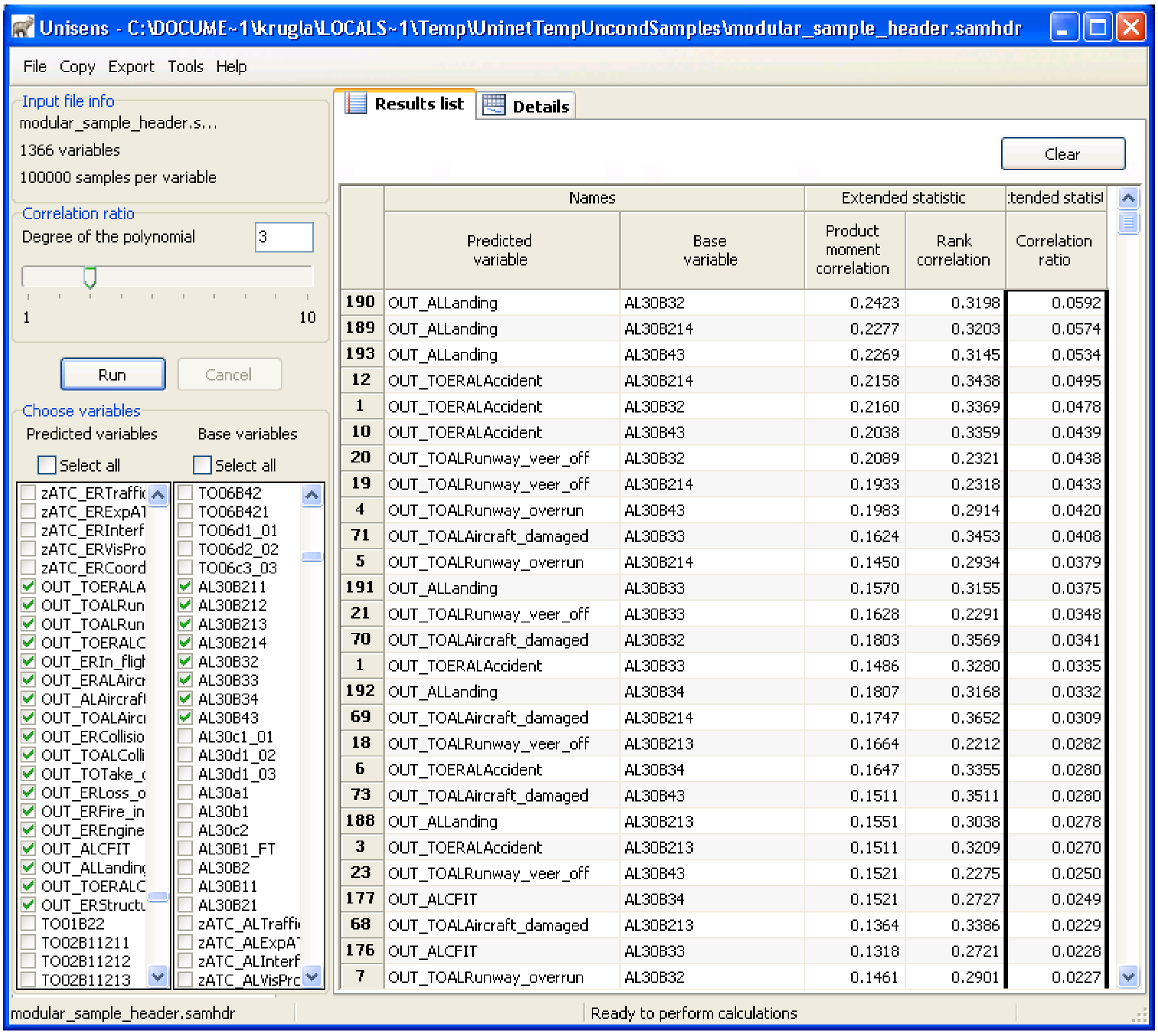


Figure 5.9: Sensitivity analysis of different accident scenarios with respect to base events representing FC, ATC, MT and non-human errors.

In this sensitivity analysis we used 18 different accident scenarios. As it was already mentioned some of these accidents are representing whole flight phases, some of them are specific for flight phase and some of them are represented in 2 flight phases, for example, take-off and approach and landing.

Table 5.4 presents some of the the accidents categories together with corresponding ESDs in which these accidents appear.

Accident category	ESDs
Landing (OUT_ALLanding)	19-30
Accident in TO, ER and AL (OUT_TOERALAccident)	all
Runway veer-off (OUT_TOALRunway_veer_off)	3,4,9, 19-30
Runway overrun (OUT_TOALRunway_overrun)	1-5,9,10,19-26,28-30
Aircraft damaged (OUT_TOALAircraft_damaged)	25

Table 5.4: Accident categories and corresponding ESD in which these accidents appear.

As we can observe, most of the time the same base variables have the strongest influence on the specific accident scenarios. For example, landing in approach and landing flight phase (ALLanding) is influenced the most by the base events AL30B32, AL30B214 and AL30B43 (all of them are representing the FC error). Again, these sensitivity analysis showed that the correlation ratios are not too high. However, the correlation ratios of 3 first variables do not vary significantly. The ratio between the highest and the lowest of these correlation ratios is equal to 1.11. For example, the variance of the ALLanding is explained by 5.9% of the variance of the conditional expectation of ALLanding given AL30B32. We would like to mention that the differences in product moment correlation and rank correlation of 3 first variables are not significant. Moreover, we can notice that the difference in correlation ratio of the 29 variables presented in figure 5.9 is up to 2.61.

Another predicted variable, accident representing take-off, en route and approach and landing is influenced by the same base variables as landing in approach and landing flight phase. Again, to base events representing FC error (AL30B214, AL30B32) and ATC error (AL30B214) have the biggest influence on accident node (as it was showed in section 5.4). Similar explanation may be provided for other predicted variables with respect to base events representing FC, ATC, MT and non-human errors (table 5.3).

We can notice that accident scenarios such as: landing in approach and landing, accident node representing take-off, en route and approach and landing, runway veer-off, runway overrun and aircraft damaged are influenced the most by base events representing the FC error.

Chapter 6

Summary and conclusions

In this chapter we present the summary of the work which was done in context of investigating the aviation risks with continuous-discrete non parametric BBNs.

The following problems have been formulated and solved during the master project:

- developed and built the Maintenance Technician performance model by performed the quantification of the unconditional and conditional rank correlations[1] together with marginal distributions[2]
- explained the structure of the CATS model giving an overview of its different parts, describing the different steps required for constructing the full model and,
- presented analysis with the version of the model that is available at the moment of writing this thesis.

The human error in maintenance may impact on safety and performance in several ways. For this reason while building the Maintenance Technician performance model, which is a part of the Causal Model for Air Transport Safety, we had to point out only these factors which may influence the MT error the most. Each of these factors have to be described with unambiguous and simple definition. For the purpose of the model we chose the following contribution factors: working condition, fatigue, experience, shift overlap time, aircraft generation and workload. Furthermore, the source of the marginal distribution needed to be assign. In case where no data was available the expert judgment procedure was used.

In order to quantify the Maintenance Technician performance model the information about unconditional and conditional rank correlations was required. In this

[1]For this purpose the expert judgment was used.
[2]Based on data or expert judgment.

situation we used expert judgment as well. In practise, the conditional rank correlations are not elicited directly from experts. Instead conditional probabilities of exceedence or ratios of unconditional rank correlations are asked. From these probabilities we can retrieve the rank correlations by assuming a copula. In our situation, we assumed the joint normal copula. We used this kind of copula, because of its useful properties, such as know relationship between rank and product moment correlation and the fact that partial and conditional correlations are equal [2]. Then by integrating appropriate normal distribution the relationship between required conditional rank correlation and the conditional probability is found.

The CATS model combines Event Sequence Diagrams (ESDs), Fault Trees (FTs) and Bayesian Belief Nets (BBNs) into a single structure, i.e. to continuous-discrete non parametric BBN. ESDs are used to represent the top part of the model and are representing accident scenarios. FTs are used to model initial and pivotal events in ESDs in sufficient detail. The initiating and pivotal events in the ESD are the top events in the FTs. The ESDs were quantified by NLR, whereas the FTs were quantified by DNV. The FT contains base events which are connected to BBNs. These represent one of the three human reliability models (FC performance model, ATC performance model and MT performance model).

We want to mention, that 36.2% of base events in the CATS model are represented by the FC error, 7.4% of them are represented by ATC error, 25.4% by the MT error and non-human errors are represented by 31%.

Before working on the CATS model, the model consisted of 1479 arcs and 649 nodes. From these nodes, 332 were functional nodes representing ESDs and FTs as boolean functions and 317 were probabilistic nodes. The 20 ESDs (out of 31 representing the generic accident scenarios) were added to the model. Three phases were represented in the CATS model: Take-off (TO), En-route (ER) and Approach-landing (AL) and also accident node representing these flight phases. After, several implementations have been done in the model, e.g. all missing ESDs and FTs have been added to the model, all end states were connected to the accident node and all base events were connected to specific human performance models, and the data were updated (e.g. more quantiles were added to the Excel file provided by DNV) the minor changes in the CATS model were observed. Finally, The CATS model consists of 4745 arcs and 1366 nodes. From these nodes 532 nodes are functional nodes representing the ESDs and FTs as boolean functions and 834 are probabilistic nodes. Approximately, 3266 more arcs have been added to the model and 717 nodes (probabilistic and functional) in total.

Large graphical structures such as the CATS BBN, can be analyzed (for example, the sensitivity analysis) in reasonable time. For the analysis we used the version of CATS BBN from 27 June 2008.

The sensitivity analysis of the MT performance model on accident in three flight phases confirms that the strongest influence on accident probability has the aircraft generation. This node is representing 2 human reliability models - FC model and MT model.

The sensitivity analysis of human reliability models on accident node showed that the most influential factors on accident probability are those representing FC performance model. Crew unsuitability in approach and landing flight phase was pointed out as the one with the strongest influence on the accident probability.

Also, we performed the sensitivity analysis of the base events representing three human errors and non-human errors on accident in take-of, en route and approach and landing flight phase. Again, this analysis confirms that most of the time the base events expressing the FC errors have to biggest influence on accident node.

The last sensitivity analysis was performed to obtain the dependencies between different types of accident scenarios and the base events from ESD 30 (Aircraft encounters unexpected wind). This analysis confirmed that different accident probabilities are mostly influenced by the same variables, e.g. AL30B214, AL30B32 and AL30B214 (all of them are linked to FC error). We may conclude that the accident scenarios such as: landing in approach and landing, accident in take-off, en route and approach and landing, runway veer-off, runway overrun and aircraft damaged are depending on the base events connected to the FC error.

We may gather that if we take into consideration different base events (those representing different ESDs in flight phases) it may happened that they will influence the most different accident categories. For example, base events representing 3 human and non-human errors in en route flight phase may have strong influence on those accident scenarios which are representing the en route flight phase (fire in flight, engine failure in flight, collision in mid air and etc.).

During preparations to meet with other experts for the quantification of the MT model, it is recommended to apply the expert's suggestions to the questionnaire to make the questions more clear. For example, to be more specific with types of jobs which we have in mind when we are thinking about maintenance technician jobs. Since jobs are different and they could take different amount of time, for example, the complete checking might take 1.5 week. The expert suggests to take, for example, component replacement such as wheel or fuel.

In future work, the CATS model will add two ESDs - ESD36 - Ground collision imminent, representing two flight phases - take-off and approach and landing and ESD37 - Wake vortex encounter, representing en route flight phase.

Appendices

Bibliography

[1] A.M. Hanea, D. Kurowicka, R.M. Cooke, Hybrid Method for Quantifying and Analyzing Bayesian Belief Nets, Delft Institute of Applied Mathematics, Delf University of Technology, The Netherlands

[2] D. Kurowicka, Roger M. Cooke, Uncertainty Analysis with High Dimensional Dependence Modeling, Delf University of Technology, The Netherlands, 2005

[3] A. Roelen, B. van Doorn, J. Smeltink, M. Verbeek, R. Wever, Quantification of event sequence diagrams for a causal risk model of commercial air transport., Tech. Rep. NLR-CR-2006-520, Nationaal Lucht- en Ruimtevaartlaboratorium (National Aerospace Laboratory NLR) (2007).

[4] O. Morales-Napoles, D. Kurowicka, R.M. Cooke, D. Ababei, Continuous-Discrete Non-Parametric Bayesian Belief Nets in Aviation Safety with UniNet, Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands.

[5] A. Roelen, G. van Baren, J. Smeltink, P. Lin, O. Morales, A generic flight crew performance model for application in a causal model of air transport, Tech. Rep. NLR-CR-2007-562, Nationaal Luchten Ruimtevaartlaboratorium (National Aerospace Laboratory NLR) (2007).

[6] O. Morales-Napoles, D. Kurowicka, R.M. Cooke, Description of the Expert elicitation Results for the ATC Performance Model, Delft University of Technology, The Netherlands, 2007

[7] R.M. Cooke, Experts in uncertainty, Oxford University Press, 1991

[8] A.L.C. Roelen, R. Wever, Accident scenarios for an integrated aviation safety model, NLR-CR-2005-560, Nationaal Lucht- en Ruimtevaartlaboratorium (National Aerospace Laboratory NLR) (November 2005)

[9] A.L.C. Roelen, Maintenance technician report (working draft)

[10] I.D. Jagielska, Quantification of non-parametric continuous BBNs with expert judgment, Delft University of Technology, The Netherlands, July 2007.

[11] BASI.1997. Human factors in airline maintenance: a study of incident reports,Department of Transport and Regional Development, Bureau of Air Safety Investigation, Australia.

[12] O. Morales, D. Kurowicka, A. Roelen, Elicitating conditional and unconditional rank correlations from condiotional probabilities, Reliability Engineering and System SafetyDoi: 10.1016/j.ress.2007.03.020.

[13] Civil aviation safety data 1989-2003, brochure

[14] G. van Baren, K. Krugla, O. Morlaes-Napoles, Minutes of meeting - Elicitation of input parameters for the maintenace performance model, Code-/Ordernumber: 1325000.3 (May 2008)

[15] I.D. Jagielska, K. Jozwiak, UniExp - Software Tool for the Elicitation of Conditional Rank Correlations by Experts as Input for Continuous Distribution Free BBNs, MANUAL

[16] 080523 DNV Collected Fault Trees v6_1.xls

[17] B.J.M. Ale, Causal Model for Air Transport Safety, Modified Seventh Interim report (15 March 2007)

[18] A. M. Hanea, D. Kurowicka, Mixed non-parametric continuous and discrete bayesian belief nets, Advances in Mathematical Modeling for Reliability ISBN 978-1-58603-865-6 (IOS Press).

[19] Unisens, MANUAL

[20] Causal Model for Air Transport Safety, Draft of the final report, 31 July 2008

Appendix A

Excitation protocol

Introduction

Thank you for participating in this expert judgment exercise of the probabilistic characterization of the performance of maintenance technicians. Within the CATS project, in which NLR and TU Delft participate, a model has been developed to represent the causal factors that are supposed to influence the probability of making errors by maintenance technicians.

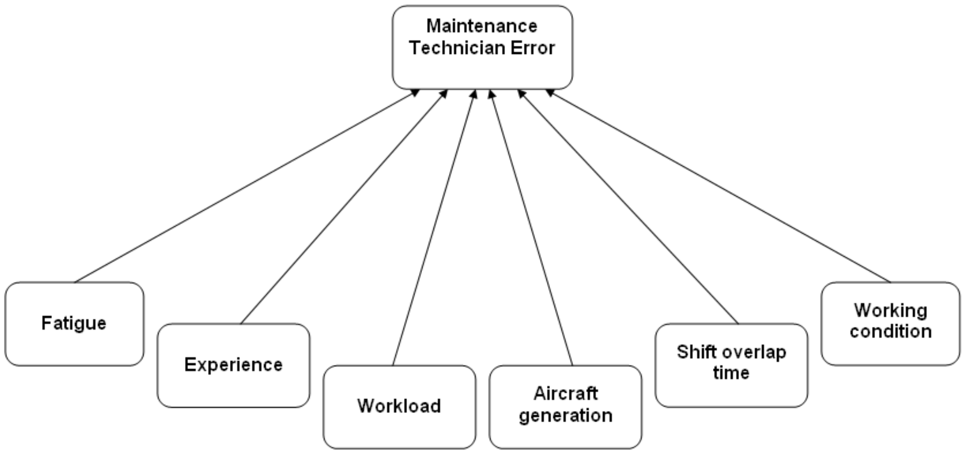
Objective

The objective of this exercise is to gather information on the variables that influence the number of errors that a maintenance technician makes per jobs. The potential variables are working condition, time of day, workload, aircraft generation, etc. This information will be used to quantify the maintenance technician performance model.

In the current version of the model we assume that the influence of a certain variable is independent of the values of other variables.

The Maintenance Technician Performance Model

The picture below displays the structure of the maintenance technician performance model. It obviously is a simplified representation of how the number of maintenance technician errors depends on certain factors. However, these five influential factors (Working conditions, Fatigue, Experience, Shift overlap time, Aircraft generation, and Workload) are considered to be most important.



The model structure itself is not subject of discussion in the current exercise, and should be considered as it is.

The definition of the variables is given in the table below.

Variable	Definition	Unit
Maintenance technician error	Number of unrecovered errors that the maintenance technician makes per # of jobs that are potentially hazardous for the safety of the flight	Number of jobs
Working condition	Whether the work is performed at the ramp (outside - 1) or in the hangar (inside - 2)	1 – 2
Fatigue	Stanford sleepiness scale, where: 1 – Feeling active, vital, alert, or wide awake; 2 – Functioning at high levels, but not at peak; able to concentrate; 3 – Awake, but relaxed; responsive but not fully alert; 4 – Somewhat foggy, let down; 5 – Foggy; losing interest in remaining awake; slowed down; 6 – Sleepy, woozy, fighting sleep; prefer to lie down; 7 – No longer fighting sleep, sleep onset soon; having dream-like thoughts.	1 – 7
Experience	Number of years in current position	Number per years
Shift overlap time	Time available to transfer a job	Minutes
Aircraft generation	Four generations of aircraft, where 4 is the most recent generation of aircrafts	1 – 4
Workload	Estimated delay in release of the aircraft	Minutes

Part I: Elicitation of probability distributions.

All questions in this section refer to a population of maintenance technicians in the Western world (Europe, North America, Australia) and Western-built large aircraft ($> 5,700$ kg Maximum Take-off Weight) currently flying in commercial operations worldwide.

Marginal distributions

When we speak about maintenance technician errors, we think of errors that are not immediately recovered, e.g.,

- Part damaged during repair
- Wrong equipment or part installed
- Panel or system not close
- Material left in aircraft
- Required service not performed

Q0	Job duration	
The number of errors (not immediately recovered) that a Maintenance technician makes are expressed per jobs. The duration of a job may vary from place to place, from time to time, or from position to position. What is your estimate of the average duration in hours of a job in your current position?		
5%	50%	95%

Q1	Number of Maintenance technician jobs	
What are your estimates for the 5th, 50th, and 95th percentiles of the distribution of number of jobs that each Maintenance technician makes during a day?		
5%	50%	95%

Q4	***Shift overlap time***	
What are your estimates for the 5th, 50th, and 95th percentiles of the distribution of the time available to transfer a job (minutes)?		
________ 5%	________ 50%	________ 95%

Q5	***Workload***	
What are your estimates for the 5th, 50th, and 95th percentiles of the distribution of the delay in release of the aircraft (minutes)?		
________ 5%	________ 50%	________ 95%

Q2	***Number of Maintenance technician errors***	
What are your estimates for the 5th, 50th, and 95th percentiles of the distribution of number of errors that each Maintenance technician makes per **median** (50% quantile in previous question) number of jobs * **median** average duration in hrs of a job?		
________ 5%	________ 50%	________ 95%

Q3	***Fatigue***					
Consider 10,000 maintenance technicians taken randomly form population. How many of them (percentage) would fall in each one of the seven following classes?						
1 – Feeling active, vital, alert, or wide awake	2 – Functioning at high levels, but not at peak; able to concentrate;	3 – Awake, but relaxed; responsive but not fully alert;	4 – Somewhat foggy, let down;	5 – Foggy; losing interest in remaining awake; slowed down;	6 – Sleepy, woozy, fighting sleep; prefer to lie down;	7 – No longer fighting sleep, sleep onset soon; having dream-like thoughts.

Q6	***Working Conditions***	
Consider 10,000 maintenance operations taken randomly from your current job. How many of these operations would have to be performed outside the hangar?		
________ 5%	________ 50%	________ 95%

Dependence information.

In this section, we are interested in the relationship between the variable Number of Maintenance Technician Errors and the variables 1-5 in Table 1 (repeated below).

Please assign number 1 to the variable that you consider to have the strongest influence (highest absolute rank correlation coefficient) on the number of Maintenance technician errors, 2 to the next most important and so on until 5 to the least influential variable.

Variable	Definition	Rank
Working condition	Whether the work is performed at the ramp (outside - 1) or in the hangar (inside - 2)	
Fatigue	Stanford sleepiness scale: 1 – Feeling active, vital, alert, or wide awake; 2 – Functioning at high levels, but not at peak; able to concentrate; 3 – Awake, but relaxed; responsive but not fully alert; 4 – Somewhat foggy, let down; 5 – Foggy; losing interest in remaining awake; slowed down; 6 – Sleepy, woozy, fighting sleep; prefer to lie down; 7 – No longer fighting sleep, sleep onset soon; having dream-like thoughts.	
Experience	Number of years in current position	
Shift overlap time	Time available to transfer a job	
Aircraft generation	Four generations of aircraft, where 4 is the most recent generation of aircrafts	
Workload	Estimated delay in release of the aircraft	

For the variable that you ranked highest, we have the following question:

Q7	
Suppose that 20,000 **Maintenance technician** are randomly chosen from our total population. Out of those, 10,000 are selected for which the chosen variable (namely: ....) has values above its median value (or above a certain percentile). What portion of these 10,000 **Maintenance technicians** will commit more than the median number of errors per median (50% quantile in question ***Q1***) jobs. Observe that the median number of Maintenance technician errors was specified in question ***Q2***)?	

For each pair below indicate which of the two effects is the most dominant?	
0	When the work is performed at the ramp (outside - 1), **higher** values of Maintenance technician error are observed, or
0	When the work is performed at the ramp (outside - 1), **lower** values of Maintenance technician error are observed.
0	Low levels of the Stanford sleepiness scale (e.g. feeling active, vital, alert or wide awake) relate to **low** values of Maintenance technician error, or
0	Low levels of the Stanford sleepiness scale (e.g. feeling active, vital, alert or wide awake) relate to **high** values of Maintenance technician error.
0	High values of experience relate to **high** values of Maintenance technician error, or
0	High values of experience relate to **low** values of Maintenance technician error.
0	High values of shift overlap time relate to **high** values of Maintenance technician error, or
0	High values of shift overlap time relate to **low** values of Maintenance technician error.
0	The most recent generation of aircrafts relates to **high** values of Maintenance technician error, or
0	The most recent generation of aircrafts relates to **low** values of Maintenance technician error.
0	High values of workload relate to **high** values of maintenance technician error, or
0	High values of workload relate to **low** values of maintenance technician error.

For the remaining variables, please indicate their influence as a portion of the influence of the variable that you ranked highest:

Node	Rank	Influence as a portion of the influence of the highest ranked variable (0 – 100%) The sum does not necessarily add to 100%
Working condition		
Fatigue		
Experience		
Shift overlap time		
Aircraft generation		
Workload		

Part II: Calibration Variables.

All questions in this section refer to a population of Maintenance technicians in the Western world (Europe, North America, Australia) and Western-built large aircraft ($> 5,700$ kg Maximum Take-off Weight) currently flying in commercial operations worldwide.

To capture your uncertainty, in all questions from this section, we will ask you to provide the 5%, 50% and 95% percentiles of your uncertainty distribution, which can be interpreted as that we ask for your best estimate (50%), the value which would surprise you if the real value would be lower (5%), and the value which would surprise you if the real value would be higher (95%).

CQ1		
How many parts does a Boeing 747-400 have including fasteners?		
5%	50%	95%

CQ2		
What was the average delay per movement, for all causes of delay, for departure traffic in the European Civil Aviation Conference (ECAC) region in 2006? (Give your answer in "minutes".)		
5%	50%	95%

CQ3		
What is the nominal Airbus A-330 nose gear shock absorber gas pressure at a temperature of 20 degrees Celsius? (Give your answer in "bar".)		
5%	50%	95%

CQ4		
What is the maximum capacity of the Boeing-777 200 ER centre fuel tank? (Give your answer in "litres".)		
5%	50%	95%

CQ5		
What was up to December 2002 the in-flight shutdown rate for the 94inch fan PW400 engine?		
5%	50%	95%

CQ6		
What was the average number of daily departures from Schiphol in the 4th quarter of 2006?		
5%	50%	95%

CQ7		
From a survey amongst 1,359 licensed aircraft maintenance engineers in Australia in 1997, what percentage would have indicated that in the last year or so, they had *never* left a tool or torch behind in an aircraft?		
5%	50%	95%

CQ8		
How many flight cycles had been accumulated by the lead-time Fokker 100 on 21 December 2006?		
5%	50%	95%

CQ9		
What is the Maximum Take-off Weight (MTOW) of an Airbus A380? (Give your answer in "kilograms".)		
5%	50%	95%

CQ10		
What is the average of the wing area in Fokker F.28 series (F28-1000, 2000,3000,4000)? (Give your answer in "square meters".)		
5%	50%	95%

CQ11		
What is the average of the minimal passenger capacity of the Fokker F.28 series (F28-1000, 2000,3000,4000)?		
5%	50%	95%

CQ12		
What was the total number of orders of Airbus A380-800 in 2007?		
5%	50%	95%

CQ13		
Consider total F28 sales of 241, including some military customers. How many of these were used as corporate jets until 1998?		
5%	50%	95%

List of Figures

2.1 BBN on 4 variables. 5
2.2 A BBN on 4 variables with associated set of the (conditional) rank correlations. 10
2.3 Relationship between $P(F_{X_4}(X_4) \geqslant 0.5|F_{X_1}(X_1) \geqslant 0.5)$ and the rank correlation $r_{4,1}$. 11
2.4 Relationship between $P(F_{X_4}(X_4) \geqslant 0.5|F_{X_1}(X_1) \geqslant 0.5, F_{X_2}(X_2) \geqslant 0.5)$ and the rank correlation $r_{4,2|1}$. 13
2.5 Example of the sensitivity analysis. 17

3.1 BBN representing the Flight Crew Performance Model. 20
3.2 Conditional distribution of FCError|Weather=0.8. 23
3.3 Conditional distribution of FCError|Weather=0.8, Workload=4000. . . . 24
3.4 Conditional distribution of FCError|Weather=0.8, Workload=4000, AirGen=1. 25
3.5 Sensitivity indices for the predicted variable FC Error and given base variables. 26
3.6 BBN representing the Air Traffic Control Performance Model. 27
3.7 Conditional distribution of ATCError|Interface=4. 29
3.8 Conditional distribution of ATCError|Interface=4, Coord=2. 30
3.9 Conditional distribution of ATCError|Interface=4, Coord=2, Traffic=15. 31
3.10 Sensitivity indices for the predicted variable ATC Error (AL19B8122 in this case) and given base variables. 32
3.11 Two Human Reliability Models for Approach and landing flight phase. From left: Flight Crew performance model and Air Traffic Control performance model. 33
3.12 Structure of Maintenance Technician Performance Model. 35
3.13 Data about experience, where the "solid" curve represents the minimum information solution assuming no MT has more that 40 years experience. 37
3.14 Maintenance Technician Performance Model. 41
3.15 Distribution of shift overlap time obtained from expert's answers. . . . 42

3.16 The panel from UniExp, where the numbers 1-7 corresponds to the following variables: 1 - Fatigue, 2 - Experience, 3 - Workload, 4 - Aircraft generation, 5 - Shift overlap time, 6 - Working condition and 7 - Maintenance technician error. 44
3.17 Relationship between $P(MTError \geqslant median|Fatigue > 4)$ and rank correlation $r_{7,1}$. 45
3.18 Relationship between $P(MTError \geqslant median|Experience \geqslant median)$ and conditional rank correlation $r_{7,2|1}$. 46
3.19 Quantification of the Maintenance Technician Performance Model. 47
3.20 Conditional distribution of X_7|Exp=3, where X_7 refers to MTError. . . . 49
3.21 Conditional distribution of X_7|Exp=3, Fatigue=1, where X_7 refers to MTError. 49
3.22 Conditional distribution of X_7|Exp=3, Fatigue=1, Workload=5.5, ShiftOverTime=6, where X_7 refers to MTError. 50
3.23 Sensitivity indices for the predicted variable Maintenance technician error (AL25B21 in this case) and a given base variable. 51

4.1 Schematic representation of the CATS model with ESDs, FTs and BBNs. 56
4.2 Event Sequence Diagram with green, red and orange paths. 57
4.3 Structure of ESD 33 - Cracks in aircraft pressure cabin. 58
4.4 Quantified ESD 33 - Cracks in aircraft pressure cabin. 59
4.5 Part of FT associated with ESD 33. 61
4.6 Information about human errors. 62
4.7 Information about human errors. 64
4.8 Initiating event, pivotal event and end events in ESD 33. 65
4.9 The representation of ESD 33 with human reliability models (from left: MT performance model and FC performance model). 66
4.10 1 - An Accident/Inicident node, 2 - Boolean functions representing ESDs and FTs, 3 - FC performance model for take-off, 4 - FC performance model for en route, 5 - FC performance model for approach-landing, 6 ATC performance model for take-off, 7 - ATC performance model for approach and landing. 68
4.11 The CATS model. 71
4.12 The CATS model, where: 1 - Maintenance Technician Performance Model and 2 - The Total Transmission Time. 73
4.13 Percentiles of the FT's base event error distribution from DNV. 75
4.14 From the left: "better" example of convergence and "worst" example of convergence. 77

5.1 The empirical cumulative distribution function of accident node before conditioning (solid line) and after conditioning on AirGen=1 (dotted line). 83
5.2 The empirical cumulative distribution function of accident node before conditioning (solid line) and after conditioning on AirGen=3 (dotted line). 84

5.3 The empirical cumulative distribution function of accident node before conditioning (solid line) and after conditioning on AirGen=1 and Exp=1 year (dotted line). 85
5.4 Sensitivity analysis of predicted variable OUT_TOERALAccident with respect to base variables from MT performance model. 86
5.5 Sensitivity analysis of predicted variable OUT_TOERALAccident with respect to base variables representing the human reliability models. . . . 87
5.6 Sensitivity analysis of predicted variable OUT_TOTake_off with respect to base variables representing the human reliability models. 89
5.7 Sensitivity analysis of predicted variable OUT_TOERALAccident with respect to base events representing FC, ATC, MT and non-human errors. 91
5.8 Sensitivity analysis of predicted variable OUT_TOERALAccident with respect to ESD 30. 92
5.9 Sensitivity analysis of different accident scenarios with respect to base events representing FC, ATC, MT and non-human errors. 94

List of Tables

3.1 Dependence Information in the Flight Crew Performance Model. 22
3.2 Dependence Information in the Air Traffic Control Performance Model Model. 28
3.3 Dependence Information for the Total Transmission Time in the Flight Crew Error and the Air Traffic Control Error. 34
3.4 Summary of experience data. 37
3.5 Summary of aircraft generation data. 39
3.6 Variables used in the Maintenance Technician Performance Model. 40
3.7 Ranking of variables according to the expert. 43
3.8 Dependence Information in the Maintenance Technician Performance Model. 47
3.9 Rank correlations for MT model. 48
3.10 Expert's performance. 52

4.1 Number of nodes and arcs in individual stages. 74

5.1 Sample based-conditioning on contribution factors with respect to accident node. 80
5.2 Analytical conditioning of the contribution factors with respect to accident node. 82
5.3 Division of the base events representing FC, ATC and MT errors and non-human errors. 90
5.4 Accident categories and corresponding ESD in which these accidents appear. 95