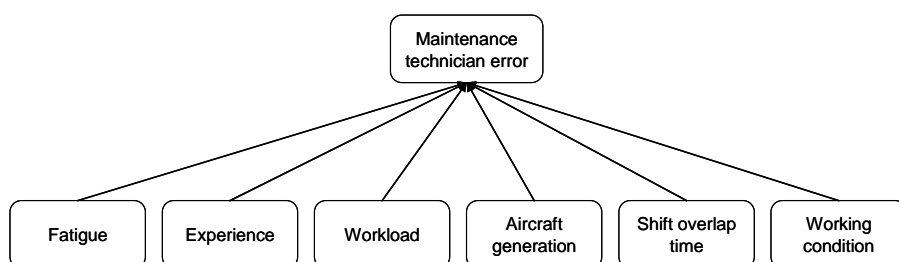


## Executive summary

# A GENERIC MAINTENANCE TECHNICIAN PERFORMANCE MODEL FOR APPLICATION IN A CAUSAL MODEL OF AIR TRANSPORT



### Problem area

The Netherlands Ministry of Transport has initiated a research effort to develop a causal model for aviation safety. The purpose of the model is to describe the air traffic system and its safety functions in such a way that it is possible to analyse risk reduction alternatives and to serve as a means of communication between experts and managers within the industry. The model combines Event Sequence Diagrams, Fault Trees and Bayesian Belief Nets (BBNs). In order to account for the influence of the human operator on accident causation, that role must be properly represented in the causal model. The human operator is best represented in a Bayesian Belief

Net because it involves primarily 'soft' influences rather than deterministic cause-effect relations. The objective of this part of the study is to develop a quantified model for maintenance technician performance for a causal model of air transport safety.

### Description of work

The development of a quantified model of maintenance technician performance is based on existing research on human performance, compatibility with other parts of the causal model, and the mathematical specifics of Bayesian Belief Nets. The BBNs in the causal model use continuous variables rather than 'conventional' BBNs that use discrete variables. The variables

**Report no.**  
NLR-CR-2008-445

**Author(s)**  
A.L.C. Roelen  
G.B. van Baren  
O. Morales  
K. Krugla

**Report classification**  
UNCLASSIFIED

**Date**  
August 2008

**Knowledge area(s)**  
Safety & Security

**Descriptor(s)**  
safety  
risk assessment  
human reliability

are represented by one dimensional marginal distributions and the relations between the variables are represented by rank correlations. Dependencies are quantified by means of expert judgement; the information was obtained according to the protocol developed by Cooke & Goossens [2000].

### **Results and conclusions**

The focus of this study was development of a practicable and working model, ready for use within the causal model. Consequently the model is relatively simple, and some of the influences might require further refinement at a later stage.

NLR-CR-2008-445

## A GENERIC MAINTENANCE TECHNICIAN PERFORMANCE MODEL FOR APPLICATION IN A CAUSAL MODEL OF AIR TRANSPORT

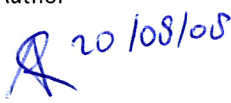
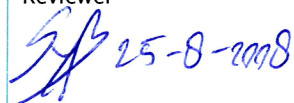
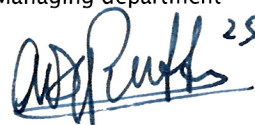
A.L.C. Roelen  
G.B. van Baren  
O. Morales<sup>1</sup>  
K. Krugla<sup>1</sup>

<sup>1</sup> TU Delft

*No part of this report may be reproduced and/or disclosed, in any form or by any means without the prior written permission of the owner.*

<b>Customer</b>	Ministry of Transport, Public Works and Water Management
<b>Contract number</b>	Ra-05.041 / DGL 5.50.2.4019
<b>Owner</b>	Ministry of Transport, Public Works and Water Management
<b>Division</b>	Air Transport
<b>Distribution</b>	Limited
<b>Classification of title</b>	Unclassified August 2008

Approved by:

Author  20/08/08	Reviewer  25-8-2008	Managing department  25/08/08
--	---	--



# CONTENTS

1	INTRODUCTION	7
1.1	Background	7
1.2	Objective and requirements	8
1.3	Research approach	8
2	REVIEW OF EXISTING HUMAN PERFORMANCE MODELS	9
3	LINKING WITH OTHER PARTS OF THE CAUSAL RISK MODEL	11
4	MATHEMATICS OF BAYESIAN BELIEF NETS	13
5	MODEL DEVELOPMENT	15
5.1	General description of aircraft maintenance	15
5.2	Rationale behind PSFs	18
5.2.1	Approach	18
5.2.2	Applicability of PSFs for a maintenance technician	20
5.2.3	Expressing PSFs in units	22
5.3	The conceptual model	29
6	MODEL QUANTIFICATION	31
6.1	Quantification process	31
6.2	Quantified model	32
7	REFERENCES	34
APPENDIX A ELICITATION PROTOCOL		37
A.1	Introduction	37
A.2	Objective	37
A.3	The Maintenance Technician Performance Model	37
A.4	Elicitation Methodology	38
A.5	Information on the probability distributions of variables	38
A.5.1	Information on the influence of variables on the number of errors a maintenance technician makes during jobs.	39

A.6	Confidentiality/ feedback	42
A.7	Marginal distributions.	43
A.8	Dependence information.	44
<b>APPENDIX B</b>	<b>ANSWERS TO CALIBRATION QUESTIONS</b>	<b>48</b>
<b>APPENDIX C</b>	<b>EXPERT ELICITATION RESULTS</b>	<b>50</b>

# I INTRODUCTION

## I.1 BACKGROUND

The Netherlands Ministry of Transport has initiated a research effort to develop a causal model for aviation safety [Ale et al 2005]. The purpose of the model is to describe the air traffic system and its safety functions in such a way that it is possible to analyze risk reduction alternatives and that it will serve as a means of communication between experts and managers within the industry. The model combines Event Sequence Diagrams, Fault Trees and Bayesian Belief Nets into a single structure. The ESD methodology is used for representing accident scenarios. Identified accident scenarios are e.g. abrupt manoeuvre, uncontrolled collision with ground, controlled flight into terrain, forced landing, mid-air collision, collision on ground, structure overload, and fire/explosion. In this way, 33 generic accident scenarios have been developed. An ESD consists of an initiating event, pivotal event(s), and an end state. The initiating and pivotal events are detailed in a submodel which can be a Fault Tree or a Bayesian Belief Net. Initiating and pivotal events may involve a human operator. The human operator plays an essential role at the operational level of any risk bearing activity. In order to account for the influence of the human operator on accident causation, that role must be properly represented in the causal risk model. The human operator is best represented in a Bayesian Belief Net because it involves primarily 'soft' influences rather than deterministic cause-effect relations. More details on the development of the causal model for air transport safety can be found in interim progress reports [Ale 2007]. It is envisioned that three types of human operators will be explicitly represented in the causal model for air transport safety:

- Flight crew,
- Air traffic controller,
- Maintenance technician.

The development of a generic flight crew performance model and a generic air traffic controller model has been reported in Roelen et al [2007a&b] The underlying report describes the development of a maintenance technician performance model.

## 1.2 OBJECTIVE AND REQUIREMENTS

The objective of this part of the study is to develop a quantified model for maintenance technician performance for a causal risk model of air transport safety. The purpose of the human performance model is to describe and quantify the influences on the probability of human failure in the fault trees describing initiating and pivotal events that constitute the ESDs of the model. The ESDs represent generic accident scenarios. A generic maintenance technician performance model is required that can be used to represent the role of the maintenance technician for different situations. The human performance model should focus on those human actions that influence the accident scenarios. The model should remain as simple as possible, focusing on the most important influential factors only. These factors should have a clear and unambiguous definition and rating scale. This is essential as the quantification of the model should be based either on available data or on expert opinion. The model will be a Bayesian Belief Net to be able to represent soft influences rather than deterministic cause-effect relations. Finally, the model should allow representation of managerial and organisational influences on maintenance technician performance.

## 1.3 RESEARCH APPROACH

The development of a quantified model of maintenance technician performance is based on existing research on human performance, the requirement to be able to link-up with other parts of the causal risk model, and the mathematical specifics of Bayesian Belief Nets. Each of these topics has its own characteristics and requirements, which we tried to merge as much as possible into a practicable model. A similar approach was followed in the development of a generic flight crew performance model and air traffic controller performance model [Roelen et al 2007a &b]. The resulting qualitative model was quantified with a combination of existing data and expert judgement. Quantification of the model concerns quantification of the marginal distributions of the nodes in the model as well as of dependencies between the nodes.

## 2 REVIEW OF EXISTING HUMAN PERFORMANCE MODELS

Human error modelling started in earnest in the 1970s. The vast majority of the early development of techniques for hazard identification and hazard representation was based on modelling hardware failures. Since human behaviour and failure is different than hardware performance and failure, new techniques had to be developed to incorporate it. In particular human behaviour is strongly influenced by recovery actions as well as initial failures. A classical approach in human error quantification is to use basic error probabilities that are modified to account for specific circumstances or contexts. Human error probabilities for general types of tasks are adjusted for the influence of possible circumstances or contexts by the application of performance shaping factors [Swain & Guttman 1983]:

$$P(\text{error}) = HEP * \sum_{i=1}^N PSF_i * W_i \quad (1)$$

HEP is the human error probability for a certain generic task,  $PSF_i$  is the  $i$ -th performance shaping factor and  $W_i$  is the weight of  $PSF_i$ . Examples of generic tasks are selecting a control, operating a control, etc. Examples of performance shaping factors are workload, time since last training, etc. Performance shaping factors are normally not seen to be causal in a strict sense, but are generally thought of as factors that contribute to how erroneous actions manifest themselves [Hollnagel 1998]. A strict methodology for deriving performance shaping factors does not exist. They are known, through experience, to have consequences for how the task is carried out and how frequent errors occur. Two of the most often used techniques that follow this approach are HEART (Human Error Assessment and Reduction Technique) [Kirwan 1994] and THERP (Technique for Human Error Rate Prediction) [Swain & Guttman 1983]. Application of these and similar first generation methods usually require a task analysis in combination with engineering techniques such as fault and event tree logic to identify where errors are important. These latter methods are better at dealing with omissions/procedures and are not good at supporting the identification of errors of commission. Errors of commission are particularly difficult to identify with techniques that only look at the logical operation and failures of the system according to design and procedures. Identifying such errors requires special support that is not readily available in the traditional way

that fault trees and event trees were build to include human error [Hale et al 1999]. In particular the construction of fault trees from planned/intended system performance, by systematically incorporating the omission or failure of each intended step, will incorporate many slips and lapses, but miss many mistakes and some violations. A second drawback of the way in which first generation methods combine human error probabilities and performance shaping factors is the assumption that the effects of the context are independent. This is not a very reasonable assumption. As an example, stress cannot be considered independent of workload. Second generation methods have been developed that try to cope with these issues. Examples of second generation methods are ATHEANA (A Technique for Human Event Analysis) [NUREG 2000, 2007] and CREAM (Cognitive Reliability and Analysis Method). The CREAM method starts with identifying the context and common performance conditions [Hollnagel 1998]. The generic task is then associated with a set of context specific error modes.

Some of the limitations of first and second generation human methods can be overcome by using BBNs to represent the influence of the context or performance shaping factors on error probability. Dependencies and interactions between performance shaping factors are easily represented in a BBN structure. In such a BBN, overall error probability is calculated as a conditional probability of a particular set of performance shaping factors rather than as a sum of separate and independent performance shaping factors. It is proposed here to represent the maintenance technician performance model as a Bayesian Belief Net with maintenance technician error as the child node and various performance shaping factors as the parent nodes, see Figure 1.

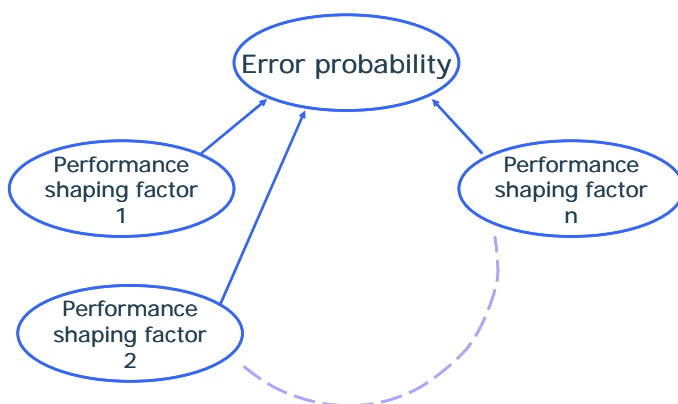


Figure 1: Human performance model schematic

### 3 LINKING WITH OTHER PARTS OF THE CAUSAL RISK MODEL

#### RISK MODEL

The design of the causal risk model for air transport calls for a combination of three modelling techniques in a single model: Event Sequence Diagrams (ESDs), Fault Trees (FTs) and Bayesian Belief Nets (BBNs), see Figure 2 [Ale et al 2007]. The ESDs represent generic accident scenarios at the most abstract level of the model [Roelen & Wever 2005]. Fault Trees link to the initiating events and pivotal events of the ESDs and describe them in a more detailed manner as a sequence of barrier failures. These base events of the fault trees include events representing human reliability. These base events involving human reliability are detailed further as Bayesian Belief Nets.

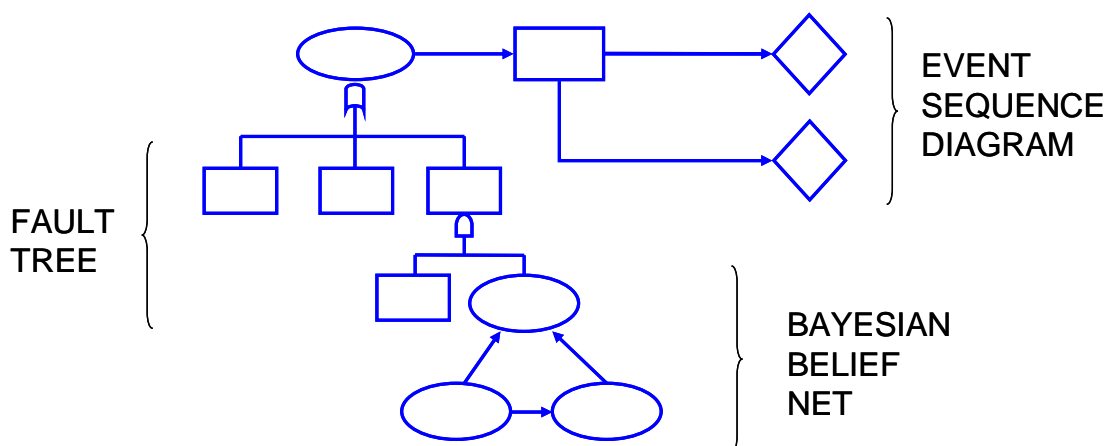


Figure 2: Overall construction of the causal risk model as a combination of ESDs, FTs and BBNs<sup>1</sup>.

The actual maintenance technician error probabilities follow from the fault trees. We propose to use the associated event in the fault tree to estimate the basic human error probability instead of deriving this from a task analysis as is the case in first generation human performance methods. Mapping of accident and incident data on those fault trees will ensure that both errors of omission as well as errors of commission will be captured in the model, because the fault trees contain both types of errors.

<sup>1</sup> In this schematic drawing, only the initiating event of the ESD is described by fault trees and BBNs. In the causal risk model however, the initiating as well as the pivotal events of the ESD are described by fault trees and BBNs.

The maintenance technician performance model must be able to represent managerial and organisational influences on human performance. In the causal risk model, safety management is described as providing the resources and criteria for the frontline workforce to operate safely. Resources and criteria are generically described by delivery systems for both hardware and human elements. The human part is governed by the following delivery systems [Bellamy et al 1999]:

- Procedures, output goals and plans
- Technology interface
- Availability of personnel
- Competence
- Communication and coordination
- Commitment.

In order to represent managerial influences on maintenance technician performance, it is proposed to describe the influences on maintenance technician performance in terms of the delivery systems, see Figure 3.

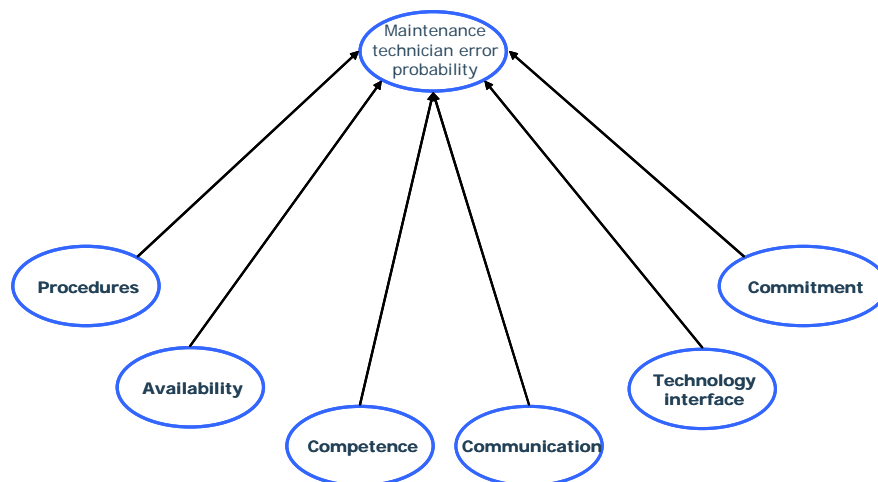


Figure 3: Linking managerial influences to the maintenance technician error probability.

The delivery systems are too generic to act directly as performance shaping factors, but they provide a framework for PSF selection. The aim is to cover all dimensions of safety management by selecting the PSFs in such a way that each of the six delivery systems is represented.

## 4 MATHEMATICS OF BAYESIAN BELIEF NETS

A detailed description of the mathematics of BBNs is provided in Morales et al [2007]. In the underlying report, several elements that are relevant for the choices made in the development of the model are repeated. The Bayesian Belief Nets that are used in the causal risk model are BBNs with continuous variables rather than 'conventional' BBNs that use discrete variables [Ale et al 2007]. When the variables (the nodes in the graph) would be represented as discrete variables and all conditional probabilities would be quantified in tables, the assessment burden can become enormous. When e.g. a child node has 7 parent nodes and each node is discretised to take 3 possible values, a conditional probability table consisting of  $3 \cdot (3^7) = 6561$  entries has to be created, which means that 6561 conditional probabilities have to be assessed. Representation of the variables as continuous variables implied until recently that the distribution of the variables was restricted to a joint normal distribution. Hanea et al [2006] introduced the 'copula-vine' approach. In this approach, the variables are represented by one dimensional marginal distributions (not necessarily normal) and the relations between the variables are represented by rank correlations. The number of rank correlations to be assessed equals the number of arcs in the model. No assumption has to be made on the joint distribution; instead a copula is chosen. A copula is a multivariate cumulative distribution function defined on the n-dimensional unit cube  $[0, 1]^n$  such that every marginal distribution is uniform on the interval  $[0, 1]$  [Bedford & Cooke 2002]. The copula contains all of the information on the nature of the dependence between the random variables that can be given without knowing the marginal distributions, but gives no information on the marginal distributions. In effect the information on the marginal distribution and the information on the dependence are neatly separated from each other. The joint distribution of the BBN is stipulated by Monte Carlo simulation.

Requirements and assumptions are [Morales et al 2008]:

- the copula represents (conditional) independence as zero (conditional) correlation.
- the rank correlations between the variables are constant; i.e. a situation where X and Y are positively correlated when variable Z takes low values, but are negatively correlated when Z takes high values cannot be represented.
- variables have to be monotonic.

Dependencies are quantified by existing data or by means of expert judgement. When expert judgement is involved, elicitation is done according to the protocol developed by Cooke and Goossens [2000]. The information on the dependencies between certain variables is elicited via the rank and conditional rank correlations, i.e. we first ask a ranking of the variables from the largest to the smallest influence. For the 'most influential' variable we then ask for the probability that the variable 'maintenance technician error' is above its median value given that the most influential variable is above its median value. Finally, the experts are asked to estimate the other variables' influence as a fraction of the 'most influential' variable's influence. This information together with the marginal distributions for each node is enough to compute the (un)conditional rank correlations required by the model. The procedure is explained in detail in Appendix E.

A consequence of the use of expert judgement is that the variables under consideration must be expressed in objectively quantifiable units. 'Objectively' denotes that a particular value of that unit has the same meaning for expert A and expert B. This requirement forces us to be very careful in selecting and defining the variables that are included in the model. As an example, consider the variable 'safety culture'. We might be tempted to include this as a variable in the model and apply a 5-point rating scale: very poor, poor, medium, good, very good. We could then ask experts to quantify the effect of safety culture on error probability. The problem here is that as long as we have not strictly defined what we mean with poor safety culture, every expert will interpret it differently and results from different experts cannot be combined.

## 5 MODEL DEVELOPMENT

### 5.1 GENERAL DESCRIPTION OF AIRCRAFT MAINTENANCE

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 business and is characterized by large amounts of regulations, procedures and documentation. This section describes the way in which a maintenance program is created and how it is implemented and executed.

To ensure that the necessary regulations, guidelines and standards are applied and adhered to, and that the correct tasks and inspections are conducted at the correct time, every aircraft type must have an approved aircraft maintenance program. This lays down the mandatory minimum maintenance program and is produced by the Maintenance Review Board (MRB) consisting of manufacturers and aviation authorities, in consultation with the airlines, to ensure the continuous maintenance needs of the aircraft are met. It includes details relating to the minimum maintenance tasks and inspections which have to be carried out on each aircraft at pre-determined times, depending upon either the number of flying hours, number of flights or calendar time. Tasks and inspection above the absolute minimum should normally accompany those which are required and advice is included in the manufacturer’s documentation. Due to an airline’s fleet size, route structure, aircraft utilization etc and from years of operational experience, the maintenance program can be customized for the individual airline.

The actual implementation of the aircraft maintenance schedule is a detailed and complex process. Most airlines have departments which are dedicated to the detailed planning, scheduling and control of aircraft maintenance tasks and inspections. Aircraft maintenance is usually scheduled in ‘checks’ of varying proportions, ranging from walk-around inspections to heavy maintenance ‘D’ checks. The Production Planning / Engineering Department controls the contents of any maintenance check, in accordance with the aircraft maintenance program and the manufacturer’s manual and decides when and where the aircraft input will occur. They therefore plan which particular aircraft is due for which check and prepare and produce all job cards and documentation necessary to complete the check. They also ensure the correct spares, manpower, tools, equipment, hangar space and aircraft are organised for a scheduled check in order that

everything will be ready when the aircraft comes in. This planning process will take about 4-5 weeks for a large commercial aircraft like a Boeing 767.

In the case of heavy maintenance, once the work has been planned, production control teams take over about 4 weeks before the aircraft arrives. Sometimes a pre-input survey of the aircraft will be carried out by qualified personnel, who will establish the condition of the aircraft and highlight any areas where extra work will be needed, e.g. repair of unforeseen minor damage and long standing defects. In the 4 weeks prior to the aircraft coming in, the proposed work schedule will be thoroughly reviewed and all work scheduled to teams of engineers. Milestones are created to enable the progress check to be monitored. The same process would apply for light maintenance except time scales would be significantly reduced since the length of the input and amount of work would be less. Some airlines may try to schedule major work for periods of reduced activity (typically the winter months from September to March), to save on revenue which would be lost in down time.

Once the actual contents of a check are known, resourcing of hangar space, people, materials, tools and equipment also takes place in this department to ensure that the work to be carried out can be achieved. This process is not as straightforward as it might seem, since for many inputs only about 60% of the work will be predictable beforehand; the rest are raised only as the results of inspections or entries made in the aircraft technical log.

A few days before the aircraft is scheduled to arrive in the hangar, the completed work package for that aircraft will be forwarded to the appropriate hangar control centre. A work package is literally the package of work which defines and sets out all the activities, i.e. tasks and inspections which must be conducted for a scheduled check on a particular aircraft. It typically contains a list of contents, job cards, a job card deviation list (to notify planning of jobs not done and therefore carried as Acceptable Deferred Defect ADD), Additional Work Requirements (i.e. non-critical extra tasks such as changing the cabin boarding tape), any relevant Service Bulletins or Airworthiness Directives, a list of parts and equipment needed and ordered, the dispatch reliability list and Certificate of Maintenance release. The work package will be organised and checked by the shift manager who is responsible for monitoring the work progress and output of the work. Shift managers will locate tools, equipment, parts and documentation necessary for the job. They are also responsible for ensuring that they have enough staff on the shift and that such staff has received adequate training. The

shift manager will liaise with Planning to advise on the man hours necessary for particular tasks in order to help plan the work.

Once the aircraft has come into the hangar the shift manager will allocate staff to a particular work area. Typically the resourcing of tasks and work allocation is written up on a control board in the control centre of each hangar, so that each technician will know when and where they will be working.

Tasks are scheduled to technicians via 'job cards' which detail information in relation to an individual task, or a number of tasks depending on the practices of the maintenance organisation. A major check may have as many as 5,000 - 10,000 job cards. Each card contains a description of the task and contain information pertaining to the type of aircraft, the identity of the aircraft, the original work package the card is from and the zone of the aircraft where the task is to be conducted. They also indicate how long the task is expected to take and what license requirements are needed to be able to sign the job off as completed. Although all job cards contain the same basic information, the overall content and format of these cards differs between maintenance organisations. All tasks also have a reference to the relevant section of the Maintenance Manual. This Maintenance Manual contains a description of how to carry out every conceivable task on the aircraft, and as such runs to thousands of pages. It includes detailed diagrams where necessary to aid in the location of components and structures. As previously noted, detailed planning can only be done for those tasks which were scheduled prior to input. The remaining unscheduled defect items will be raised as defect cards from the original scheduled inspections, i.e. any defects which are discovered have job cards raised as they are discovered. During the pre-planning of a check, all job cards are arranged into suitable functional groups ready for the input. The job cards are distributed to the technicians by placing them on racks in designated areas. For example job cards appear on a number of separate display boards depending on the nature of the tasks. The technicians will collect their assigned job cards from these racks to perform the tasks. On completion of the tasks the technician must sign or stamp in the designated place depending on what type of tasks they are licensed to complete. Safety critical tasks will require duplicate inspections. Once signed the job cards are returned to the aircraft's technical records where they are stored for a minimum of two years. [O'Conner and Reynolds 1996, Roelen et al 2006]

For ramp maintenance, i.e. those maintenance tasks that are carried out in between two flights while the aircraft is standing at the gate or on the platform task resourcing and work allocation is a bit more hectic. For typical European

airlines from about 6:00 -22:30 there are some 100 to 200 narrow body departures each day and it is the Traffic Coordinators job to assign technicians to particular aircraft. The traffic coordinator will have a 'time-line' of daily departures with the technicians name, radio call and aircraft registration shown against each time window. If an aircraft is delayed the coordinator will try and reassign the work for an aircraft to another technician. In this way they attempt to keep a technician or team of technicians assigned to an aircraft for the duration of its stay rather than swapping and changing the schedule.

According to European regulations (EC Regulation 2042/2003, Part 66), there are 4 different classes of maintenance licenses:

- A category A aircraft maintenance licence permits the holder to issue certificates of release to service following minor scheduled line maintenance and simple defect rectification within the limits of tasks specifically endorsed on the authorisation. The certification privileges shall be restricted to work that the licence holder has personally performed in a Part-145 organisation.
- A category B1 aircraft maintenance licence permits the holder to issue certificates of release to service following maintenance, including aircraft structure, powerplant and mechanical and electrical systems. Replacement of avionic line replaceable units, requiring simple tests to prove their serviceability, is also included in the privileges. Category B1 automatically includes the appropriate A subcategory.
- A category B2 aircraft maintenance licence permits the holder to issue certificates of release to service following maintenance on avionic and electrical systems.
- A category C aircraft maintenance licence permits the holder to issue certificates of release to service following base maintenance on aircraft.

## 5.2 RATIONALE BEHIND PSFs

### 5.2.1 APPROACH

The general approach in determining relevant performance shaping factors was to follow 'Good practices for implementing human reliability analysis' as prepared by the U.S. Nuclear Regulatory Commission and published as NUREG 1792 [NUREG 2005]. In particular, Appendix B of that report provides 'guidance on consideration of performance shaping factors'. Although that appendix was written 'for the specific purpose of addressing post-initiator HFEs in a risk

assessment for commercial nuclear power plant operations occurring nominally at full power, and for internal initiating events', much of it is considered useful to other modes of operations and for other industry applications.

It is acknowledged in the guidance document that the way in which PSFs are defined may differ somewhat across HRA methods and tools. The guidance document describes 15 Performance Shaping Factors.

1. Applicability and suitability of training / experience
2. Suitability of relevant procedures and administrative goals
3. Availability and clarity of instrumentation
4. Time available and time required to complete the act, including the impact of concurrent and competing activities
5. Complexity of the required diagnosis and response, the need for special sequencing, and the familiarity of the situation
6. Workload, time pressure and stress
7. Team/crew dynamics and crew characteristics
8. Available staffing / resources
9. Ergonomic quality of the human - system interface
10. Environment in which the action needs to be performed
11. Accessibility and operability of the equipment to be manipulated
12. Need for special tools
13. Communications and whether one can be easily heard
14. Special fitness needs
15. Consideration of 'realistic' accident sequence diversions and deviations.

For the purpose of the human performance models of CATS, the 15 PSFs of NUREG 1792 were used as the starting point. For each of the human performance models it was decided which of these PSFs were relevant and which were not, using the guidance on consideration of PSFs in appendix B of NUREG 1792. Next, for the relevant PSFs it was determined how the 'value' of that PSF could best be described.

## 5.2.2 APPLICABILITY OF PSFs FOR A MAINTENANCE TECHNICIAN

This section describes, for each of the 15 performance shaping factors, the considerations for determining their applicability to a maintenance technician.

### Applicability and suitability of training / experience

This is considered to be a relevant factor. The degree of familiarity with the tasks to be performed influences the likelihood of success. This degree of familiarity depends on the training and the experience.

### Suitability of relevant procedures and administrative goals

This is a relevant factor. For the most part, procedures exist to cover many types of sequences and maintenance actions in normal and abnormal situations.

### Availability and clarity of instrumentation

For a maintenance technician the instrumentation that is used is part of the tools and the equipment to be used. Therefore this PSF is combined with PSF 'Accessibility and operability of the equipment to be manipulated' and further discussed under that heading.

### Time available and time required to complete the act.

This factor is considered relevant if there is not enough or barely enough time to act. It is clearly related to PSF 'workload, time pressure and stress'. Therefore this PSF is combined with PSF 'workload, time pressure and stress' and further discussed under that heading.

### Complexity of the required diagnosis and response, the need for special sequencing, and the familiarity of the situation

This factor attempts to measure the overall complexity of the situation at hand and of the action itself. It includes such measures as the number of steps to be performed in rapid succession, the degree of mental effort involved, whether it is a multi-variable or single variable associated task, etc. It is considered relevant.

### Workload, time pressure and stress

This is about the amount of work a maintenance technician has to accomplish in the time available, along with their overall sense of being pressured or threatened in some way with what they are trying to accomplish. If the scenario is familiar, the procedures and training for the scenario are good, and the rate at which the maintenance crews normally implement their procedures allows them to achieve their goal on time, even relatively high expected levels of workload

and stress will not have a significant effect on performance [NUREG 2005]. Therefore, workload is particularly considered relevant for abnormal situations. The factor workload will be integrated with the factor 'available staff / resources' (see under that heading).

#### Team/crew dynamics and crew characteristics

This PSF addresses the degree of independence among team members and their approach for implementing procedures (e.g. aggressive vs. methodical). The factor is considered particularly relevant for unexpected events and distractions.

#### Available staff / resources

This is considered to be a relevant factor, primarily because of its influence on the time required to perform the maintenance actions. If maintenance technicians are not available due to e.g. illness and a timely replacement cannot be obtained, the amount of work for the remaining crew will increase but there will be pressure to maintain the original time schedule. The time required to perform the maintenance actions increases while the time available remains the same. This will then increase workload for the available technicians. Because of the strong link between available staff and workload, the two factors will be integrated.

#### Ergonomic quality of the human system interface.

In the case of maintenance technicians, the human system interface is related to the equipment to be used. This PSF will be integrated with 'Accessibility and operability of the equipment to be manipulated' and further discussed under that heading

#### Environment in which the action needs to be performed

The environment is relevant for maintenance technicians as they frequently have to work outside in all sorts of weather conditions.

#### Accessibility and operability of the equipment to be manipulated

This is considered to be a relevant factor and is integrated with 'Availability and clarity of instrumentation' and 'Ergonomic quality of the human system interface' and 'Need for special tools' as they are all related.

#### Need for special tools

This factor is considered relevant for maintenance technicians as much of the equipment they use are special tools. As a matter of fact, equipment and tools are almost synonymous in maintenance tasks and therefore this factor is

integrated with 'Accessibility and operability of the equipment to be manipulated'

#### Communications and whether one can be easily heard

This is about the possibility that directives are misunderstood. For maintenance technicians this is relevant as they have to work together and coordinate activities within a team and between shifts.

#### Special fitness needs

Physical fitness is not considered an issue for maintenance technicians, but fatigue should be considered. Aircraft maintenance is a 24/7 hour activity and involves shift work. The effects this can have on fatigue and subsequently performance are well documented and are considered relevant for maintenance technicians.

#### Consideration of 'realistic' accident sequence diversions and deviations

This PSF describes the extent to which the actual sequence of events appears to be similar to how it is trained in the simulator. Because simulators are not used for training of maintenance technicians this factor is not considered relevant.

In conclusion, the following PSFs are considered relevant for the maintenance technician:

- Applicability and suitability of training / experience
- Suitability of procedures
- Complexity of the required diagnosis and response
- Workload, time pressure and stress
- Team/crew dynamics and crew characteristics
- Environment in which the actions need to be performed
- Accessibility and operability of the equipment to be manipulated
- Communications and whether one can be easily heard
- Special fitness needs

### 5.2.3 EXPRESSING PSFs IN UNITS

The next step in the process is to establish for the relevant performance shaping factors, the possibilities to express their values in objectively quantifiable units.

#### Applicability and suitability of training / experience

A person can obtain an aircraft maintenance license if requirements for basic knowledge and experience are met. The basic knowledge is usually obtained by

completing an approved training course and experience is obtained by working under supervision as an unlicensed technician on operating aircraft. The required duration of this on the job training depends on the licence category (A, B1, B2, or C) and previous experience and can range from one to five years.

For the license to maintain valid, the maintenance technician must at least have 6 months of maintenance experience in every 2 year period.

Because of the importance of experience, and because expressing the quality of training in an objectively quantifiable unit is considered, for the moment, impossible, the total experience, in years, is selected as the appropriate unit to represent this performance shaping factor.

#### Suitability of procedures

Procedures for aircraft maintenance are laid down in the 'Instructions for Continued Airworthiness' which include also Original Equipment Manufacturers' documents such as the Aircraft Maintenance Manual (AMM), Structural Repair Manual (SRM), Trouble Shooting Manual (TSM) or Fault Isolation Manual (FIM), Wiring Diagram Manual (WDM), Corrosion Prevention Manual (CPM), Non-Destructive Test Manual (NTM), Component Maintenance Manual (CMM), Engine Manual, Illustrated Parts Catalogue (IPC), etc.

System representatives participate in the development of the manuals to provide knowledge on specific systems. Furthermore people with hands-on experience in maintenance are often represented in the process.

There are several standards to support the development of technical manuals. The original ATA 100, and the new ispec 2200 that replaces it, were developed to provide standards for the format and layout of technical manuals as well as an outline for the organization of material within a manual. One of the standards in use is Simplified English which provides a limited and standard vocabulary for technical writers and helps improve the comprehension of documents by non-native English speakers.

Inaccuracies in maintenance documentation are encountered by maintenance organizations. This is illustrated by the number of revisions and temporary revisions that can be found in training bulletins issued in maintenance organizations and by incident and accident reports in which documentation deficiencies are mentioned as causal or contributing factors.

Sometimes procedures are not compatible with the operational practice. Examples are cases where the sequence of tasks is unpractical. Procedure design occasionally makes it difficult to both follow the procedures exactly and carry out the work expeditiously. Such incompatibilities reduce the likelihood of strict adherence to procedures.

Incomplete documentation is mainly caused by inaccuracy during the compilation process of the documentation. A complicating factor is that each airline that orders new aircraft modifies the basic delivery; thereby almost every aircraft that is delivered from the OEM enters into the operational environment with different systems. Time constraints during documentation development may result in omissions, for instance when procedures for one configuration are duplicated for another configuration, without taking sufficient account of differences between the two configurations.

The most frequent form of incorrect documentation are cases where references or codes are incorrect. Examples are references to wrong tools or equipment.

Because of the many procedures, it is practically impossible to assess each procedure on characteristics like ambiguity, compatibility, etc. We could try to look at groups of procedures and consider general differences between for instance Boeing and Airbus procedures, or between different airlines, or between aircraft generations [Roelen et al 2006].

In general aircraft of newer design are easier to maintain than aircraft of older design [Biemans et al 1998], although some maintenance engineers may have a personal preference for older aircraft types. When considering the effect of technological advances on safety of air transport it is common to consider four different generations of aircraft since the introduction of the jet engine. First generation aircraft are typically designed in the 1950s. Most of the aircraft were certified before 1965 according to British Civil Airworthiness Requirements (BCAR's) 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 are the DH Comet, Fokker F-27 and Boeing 707. Second generation aircraft, designed in the 1960s and 1970s, have more reliable engines. The aircraft were certified between 1965 and 1980, 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 aircraft, 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-bypass engines designed according to higher certification standards. Examples of third generation aircraft are Fokker 50 and Boeing 737-700. Fourth generation 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. For this reason 'aircraft generation' is considered to be indicative of the quality of maintenance procedures.

#### Complexity of the required diagnosis and response

This factor attempts to measure the overall complexity of the situation at hand and of the action itself. It describes the extent to which the maintenance activities involve simple straightforward tasks or tasks that the maintenance technician is very familiar with. It includes such measures as the number of steps to be performed in rapid succession, the degree of mental effort involved, whether it is a multi-variable or single variable associated task, etc. However, we have not been able to represent the degree of complexity in objectively quantifiable units. At this stage it is therefore considered impossible to quantify, but it is recommended to continue exploring possible ways to represent this factor in future versions of the model.

#### Workload, time pressure and stress

According to a survey amongst 1,359 licensed aircraft maintenance engineers in Australia, work pressure is the most commonly mentioned factor for safety occurrences [Hobbs and Williamson 1997]. A part of the work pressure is composed of the workload. High workload exists when task demand is close to the operator's maximum capacity, while workload is low when task demand is much below the operator's capacity. Hence, workload is not only sensitive to multiple characteristics of a task, i.e. task demand, but also of the operator, i.e. operator capacity. [Hart 1987, Hancock et al. 1995]. Task demand is determined by [modified from Lysaght et al. 1989]:

- Number of actions
- Sequence of actions
- Time required for action to be completed.
- Type of action.

The number of tasks to be completed by a single technician depends on the total number of technicians available per aircraft, but also by characteristics such as variability of the fleet, average age of the fleet and how tasks are distributed amongst technicians.

There is an abundance of literature on the topic of workload measurements, and various workload measurement schemes are proposed, including such items as eye-movements, heart rate, subjective ratings etc. For the purpose of this model, the deviation of actual versus scheduled task completion as percentage of the overall task schedule is proposed as the unit of measure for workload. As a proxy for this, we propose to use the delay in release to service of the aircraft. Records of this delay are kept by the maintenance organisations.

#### Team/crew dynamics and crew characteristics

The composition of the team of maintenance technicians is expected to influence team member coordination and cooperation and hence affect the maintenance technician error probability. The ability of maintenance technicians to work together as a team depends on cultural and psychosocial factors and can also be influenced by experience, training, and company procedures.

An element that belongs to this factor is the presence of informal rules regarding the way procedural steps are interpreted [NUREG 2005]. From past research, it is known that some maintenance technicians use *blackbooks*, personal (and illegal) notebook in which the technician records useful information for their own reference [Biemans et al 1999]. These blackbooks are most frequently used during line maintenance. A specific line maintenance problem is that maintenance documentation is not available at the location where the task is carried out. Certainly on large airports it takes technicians a large amount of time to go and consult the maintenance documentation in a hangar or office building. It is not uncommon that this requires half an hour, which is, for aircraft like the Boeing 737 and Airbus A-320, very close to the total time planned for the turn-around at the gate. Because the perception of the licensed technician is that he does not need the procedure for a large part of the line maintenance tasks there will be a tendency not to fetch the manual but rather to carry out the task using experience from the past. This may pose a problem for detailed information like torque values, part numbers and limits and therefore many technicians carry a blackbook in which they have recorded this type of information. Although probably most commonly used in line maintenance, blackbooks are sometimes also used during base maintenance [Van Avermaete & Hakkeling-Mesland 2001]. The largest problem with the existence of blackbooks

is a risk that by consulting own private notes or retained copies of maintenance documentation technicians will use outdated information. On the other hand, technicians who do not use a blackbook may perhaps work from memory, which might even be worse.

From past research results it is not possible to give an estimate of the percentage of technicians using a blackbook, but based on interviews it is concluded by Van Avermaete & Hakkeling-Mesland [2001] that they probably exist within every organisation. From an ergonomic survey of maintenance documentation carried out by Trinity College Dublin it was concluded that the use of blackbooks was ubiquitous [McDonald et al, 1999]. Maintenance technicians are aware of the fact that these books are illegal and were found to be reluctant to discuss the topic during the interviews.

In view of the fact that the use of blackbooks can have negative and positive effects on flight safety, and the fact that technicians are reluctant to discuss the issue of blackbooks it is not considered feasible to assess the effect of blackbooks on flight safety from operation experts. Therefore this factor is not included in the model, but it is recommended to continue exploring possible ways to represent this aspect in future versions of the model.

#### Environment in which the actions need to be performed

An FAA evaluation of aviation maintenance working environments concluded that the following environmental factors were most influential on human performance: temperature, noise and lighting [Johnson et al 2001]. In general, aircraft working conditions are more favourable in a hangar environment than on the ramp [Biemans et al 1998]. The ramp is also busy (baggage tugs, catering, refuelling, etc) and because of this it is sometimes difficult to get access to parts if the aircraft. For the purpose of the model, 'working conditions' is considered to be a variable with two possible states: 'platform' and 'hangar'.

#### Accessibility and operability of the equipment to be manipulated

The quality of the interface between machine (the aircraft) and its human operator (the flight crew) has greatly improved over the years. This can be illustrated by comparing the cockpit of a first generation commercial jet transport aircraft like the De Havilland DH-106 Comet, with that of a modern jet airliner like the Boeing 777. When considering the effect of technological advances on safety of air transport it is common to consider four different generations of aircraft since the introduction of the jet engine. Those different aircraft generations provide a convenient classification for the quality of the man-

machine interface. Development of maintenance tools and equipment is assumed to have gone through a similar incremental development as the flight deck. For this reason 'aircraft generation' is also<sup>2</sup> considered to be indicative of the accessibility and operability of maintenance tools and equipment.

#### Communications and whether one can be easily heard

Communication is essential during task- and shift handovers. An infamous example of a catastrophic aircraft accident that resulted from inadequate maintenance shift handover is the Embraer 120 accident on September 11, 1991 in Texas, USA [NTSB 1992]. The primary objective of handovers is to ensure that all necessary information is communicated between the out-going and in-coming personnel. Ideally the procedure should provide for sufficient time to be made available by way of a shift overlap, depending on the complexity of task(s) to be handed over. As a guideline, 20 to 30 minutes could be considered good human factors practise [CAA 2003].

Noise levels are also factors that influence communication. Noise levels at the ramp can be considerable, e.g. when the APU the aircraft that is being serviced is operating and in the case of aircraft with operating engines in the vicinity of the aircraft being serviced. In the hanger, the noise caused by the use of equipment (e.g. tools for riveting) can also hamper communication between two technicians.

Language differences between maintenance technicians may also result in communication difficulties.

Shift hand-over is most often mentioned as the primary communication issue. therefore, and in view of the need to keep the number of nodes in the model limited, the factor communication will be restricted In the model to shift handover. As a proxy for the 'quality' of the shift handover, the shift overlap time is selected

#### Special fitness needs

Fatigue is frequently 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. Maintenance teams therefore work in shifts, and some people have difficulty adjusting to the schedule. A shift system affects the social life as well. Research has shown that human performance is influenced by the time of day. Generally speaking people

---

<sup>2</sup> Note that aircraft generation is also used to represent the influence of the quality of procedures.

perform less well during the early hours in the morning [Rosekind et al 1994]. This is confirmed by a BASI study that showed the greatest relative frequency of incidents occurred during the night shift between 02:00 hrs and 04:00 hrs [BASI 1997].

Because fatigue is less complicated to represent, because of the availability of data, and because time of day and fatigue are correlated, it is proposed to represent only fatigue in the model, and not to represent time of day or any other factor that could influence a maintenance technician’s physical condition.

Fatigue will be quantified using the Stanford Sleepiness Scale (SSS). The result of the SSS is a score with increasing sleepiness from 1 to 7, where 1 signifies “feeling active and vital; wide awake” and 7 stands for “almost in reverie; sleep onset soon; losing struggle to remain awake”.

### 5.3 THE CONCEPTUAL MODEL

Having defined the model variables, their interrelations are defined in the model structure as presented in Figure 4. Each variable described in the previous section is represented as a node in the model. Arrows in Figure 4 indicate dependencies between nodes.

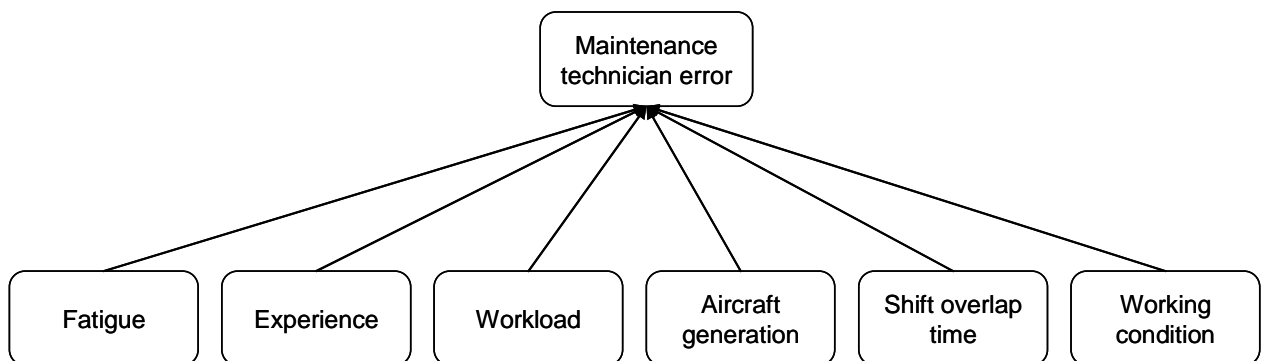


Figure 4: Maintenance technician performance conceptual model

Comparison of this model with the framework proposed in Figure 3 shows that all delivery systems except ‘commitment’ are represented in the model, see table 1. The delivery system ‘commitment’ was also problematic in the performance models for the flight crew [Roelen et al 2007a] and the air traffic controller [Roelen et al 2007b]. The model element ‘working condition’ has not associated delivery system. It is considered to be an external factor, beyond the direct influence of management.

*Table 1: Comparison of model elements with generic delivery systems*

Model element	Associated delivery system
Fatigue	Availability
Experience	Competence
Workload	Availability
Aircraft generation	Procedures / Man-machine interface
Shift overlap time	Communication
Working condition	-

## 6 MODEL QUANTIFICATION

### 6.1 QUANTIFICATION PROCESS

Quantification of the model concerns quantification of the marginal distributions of the nodes of the model as well as of the dependencies between the nodes.

Table 2 lists the nodes, their definition, and how the marginal distributions were derived.

*Table 2: Maintenance technician model nodes and their definitions*

Node	Description	Unit	Source for marginal distribution
Maintenance technician error	Number of unrecovered errors that the maintenance technician makes per job	Number of errors per job	-
Fatigue	7 levels of fatigue, 1 means wide awake, 7 is close to sleep onset.	Stanford Sleepiness Scale	Expert judgement
Experience	Number of years working as a maintenance technician in current position	Years	Bureau of Labour Statistics
Workload	Delay in release to service of the aircraft.	Minutes	Expert judgement
Aircraft generation	Four generations of aircraft, with 4 being the most recent.	1 - 4	Schiphol movement data
Shift overlap time	Overlap time of two subsequent maintenance shifts	Minutes	Expert judgement
Working condition	Whether the work is performed at the ramp (outside) or in the hangar (inside)	Ramp Hangar	Expert judgement

For two variables the data for the marginal distributions were readily available. Data on experience was obtained from the Bureau of Labour Statistics. The results are presented in Table 3. Data on aircraft generation were obtained from Schiphol statistical summary. The results are presented in Table 4.

*Table 3: Maintenance technician experience distribution*

Experience	% of maintenance technicians
3 years or less	22.8 %
4-9 years	28.5%
10-19 years	16.2%
More than 20 years	32.5%
Median = 9.4 years	

Table 4: Aircraft generation distribution

Aircraft generation	% of population
1	0.08 %
2	6.14 %
3	90.78%
4	3.00 %

For the remaining 4 variables the expert judgement procedure was performed. The conditional rank correlations in the model were determined from the conditional probability of experience and ratios of unconditional rank correlation provided by the experts. The order of variables for the elicitation is assumed to be as follows: 2, 3, 6, 5, 4, 1 (fatigue, experience, workload, aircraft generation, shift overlap time, working condition).

A single expert, a maintenance technician of the NLR research aircraft maintenance group, was consulted. The elicitation protocol that was developed for this purpose is represented in Appendix A. All questions in the elicitation protocol 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. In addition to asking questions on the variables of interest, the protocol also included questions on variables for which the correct answers were known to the researchers but not to the experts. Those questions are calibration questions and are important for assessing the performance of the combined experts' assessment. They also form an important part of the feedback to the experts, helping them to gauge their subjective sense of uncertainty against quantitative measures of performance [Cooke and Goossens 2000]. The calibration questions used in this study are presented in Appendix A. Results of the elicitation process are presented in Appendix C. A minimally informative distribution was fitted to the expert's assessment.

## 6.2 QUANTIFIED MODEL

Figure 5 presents the model including dependencies. In Figure 5 these dependencies are represented by their formal notation. For instance  $r_{7,2}$  represents the dependence between 'maintenance technician error' and 'fatigue', and  $r_{7,3|2}$  represents the dependence between 'maintenance technician error' and 'experience' given 'fatigue'.

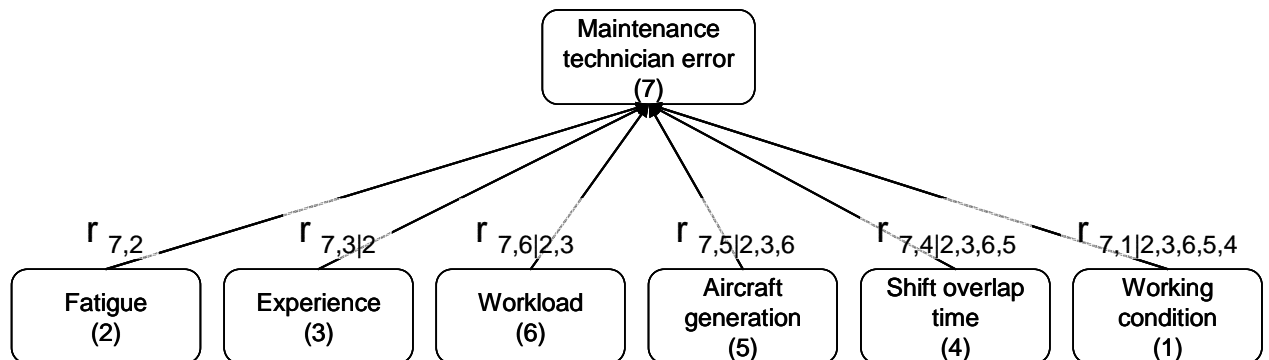


Figure 5: Maintenance technician performance model with dependence notation

The dependence information resulting from the expert elicitation is presented in Table 5. The numbers were obtained using the software tool UniExp, more information on the software is found in Jagielska [2007]. Details on the mathematics involved in the quantification of this model are presented in Krugla [2008].

Table 5: Dependence information

Probability <sup>3</sup>		(Un)conditional rank correlation	
$P_1$	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

<sup>3</sup>  $P_1 = P(\text{maintenance technician error} \geq \text{median} \mid \text{fatigue} \geq 4)$ ,  $P_2 = P(\text{maintenance technician error} \geq \text{median} \mid \text{experience} \geq \text{median})$ ,  $P_3 = P(\text{maintenance technician error} \geq \text{median} \mid \text{workload} \geq \text{median})$ ,  $P_4 = P(\text{maintenance technician error} \geq \text{median} \mid \text{aircraft generation} = 4)$ ,  $P_5 = P(\text{maintenance technician error} \geq \text{median} \mid \text{working condition} = \text{outside})$ .

## 7 REFERENCES

- [1] Ale, B.J.M., Bellamy, L.J., Cooke, R.M., Goossens, L.H.J., Hale, A.R., Kurowicka, D., Roelen, A.L.C., Smith, E. (2005). Development of a causal model for air transport safety, Proceedings of IMECE2005, 2005 ASME International Mechanical Engineering Congress and Exposition, November 5-11, 2005, Orlando, Florida USA, IMEC2005-79374.
- [2] Ale, B.J.M. (2007). Causal Model for Air Transport Safety, Seventh Interim Report. Delft University of Technology, Delft, the Netherlands.
- [3] Ale, B.J.M., Bellamy, L.J., Boom, R. van der, Cooper, J., Cooke, R.M., Goossens, L.H.J., Hale, A.R., Kurowicka, D., Morales, O., Roelen, A.L.C., Spouge, J. (2007). Further development of a causal model for air transport safety (CATS): building the mathematical heart. In Aven & Vinnen (eds), Risk, Reliability and Societal Safety, Taylor & Francis Group, London, UK.
- [4] Avermaete, J.A.G., Hakkeling-Mesland, M.Y. (2001). Maintenance human factors from a European research perspective: results from the Adams project and related research initiatives, 15th Symposium on Human Factors in Aviation Maintenance, London, UK.
- [5] BASI. (1997). Human factors in airline maintenance: a study of incident reports, Department of Transport and Regional Development, Bureau of Air Safety Investigation, Australia.
- [6] Bedford, T.J., Cooke, R.M. (2002). Vines - a new graphical model for dependent Random Variables. *Annals of Statistics*, Vol. 30, No. 4, p. 1031-1068.
- [7] Bellamy, L.J., Papazoglou, I.A., Hale, A.R., Aneziris, O.N., Ale, B.J.M., Morris, M.I., Oh, J.I.H. (1999). Development of an integrated technical and management risk control and monitoring methodology for managing and quantifying on-site and off-site risks, I-Risk main report, May 1999, EU Contract number ENVA CT96-0243.
- [8] Biemans, M.C.M., Avermaete, J.A.G. van, Roelen, A.L.C., Meulen, G.J. van der (1998). ADAMS Aircraft maintenance operational systems, Identification of bottlenecks in the existing maintenance and dispatch system, technical report ADAMS-WP2B-TR01, Issue 1.0. NLR-TR-98120, NLR Amsterdam.
- [9] CAA. (2003). Aviation Maintenance Human Factors (EASA / JAR145 Approved Organisations), Guidance Material on the UK CAA Interpretation of Part-145 Human Factors and Error Management Requirements, CAP 716, Safety Regulation Group, Civil Aviation Authority, London, UK.
- [10] Cooke, R.M., L.J.H. Goossens, L.J.H. (2000). Procedures guide for structured expert judgement, EURATOM document EUR 18820EN, European Commission, Directorate General for Research.

- [11] Hale A.R, Goossens, L.H.J., Bellamy, L.J. (1999). Modelling errors of commission: applicability of ATHEANA in the chemical industry. In: L.H.J. Goossens (eds.); Proceedings Rotterdam '99. Risk Analysis: Facing the New Millennium (Rotterdam, 10/10/99), Delft University Press, Delft, p. 243-246. ISBN: 90-407-1954-3.
- [12] Hancock, P.A., Williams, G., Manning, C.M. (1995). Influence of Task Demand Characteristics on Workload and Performance, *The International Journal of Aviation Psychology*, 5 (1), 63-86.
- [13] Hanea, A.M., Kurowicka, D., Cooke, R.M. (2006). Hybrid method for quantifying and analysing Bayesian belief nets, *Quality and Reliability Engineering Journal*, Vol. 22, Issue 6, p. 709-729.
- [14] Hart, S.G. (1987). The prediction and measurement of mental workload during space operations, paper presented at the NASA Space Life Sciences Symposium, Washington D.C., June 1987.
- [15] Hobbs, A., Williamson, A. (1997). Aircraft maintenance safety survey - results. Department of Transport and Regional Services, Australian Transport Safety Bureau, Australia.
- [16] Hollnagel, E. (1998). Cognitive reliability and error analysis method, Elsevier Science, Oxford.
- [17] Jagielska, I.D. (2007). Quantification on non-parametric continuous BBNs with expert judgement, Master thesis, Delft Institute of Applied Mathematics, Delft University of Technology.
- [18] Johnson, W.B., Mason, F, Hall, S., Watson, J. (2001). Evaluation of aviation maintenance working environments, fatigue, and human performance, Federal Aviation Administration. <http://hfskyway.faa.gov>
- [19] Kirwan, B. (1994). A practical guide to human reliability assessment, Taylor and Francis, London, UK.
- [20] Krugla, K.A. (2008). Aviation risks with continuous / discrete non parametric BBNs, Master thesis, Delft Institute of Applied Mathematics, Delft University of Technology.
- [21] Lysaght, R.J., Hill, S.G., Dick, A.O., Plamondon, B.D., Linton, P.M., Wierwille, W.W., Zaklad, A.L., Bittner, A.C., Wherry, R.J. (1989). Operator Workload: Comprehensive Review and Evaluation of Operator Workload Methodologies, Technical Report 851, United States Army Research Institute for the Behavioral and Social Sciences, USA.
- [22] McDonald, N. et. al. (1999). Human-Centred Management for Aircraft Maintenance, ADAMSWP4A.R5.
- [23] Morales, O., Kurowicka, D., Roelen, A. (2008). Eliciting conditional and unconditional rank correlations from conditional probabilities, *Reliability Engineering and System Safety*, Vol. 93, Issue 5, p. 699-710.

- [24] NTSB. (1992). Britt Airways, Inc., d/b/a Continental Express flight 2574, In-flight structural breakup EMB-120RT, N33701 Eagle Lake, Texas, September 11, 1991, Aircraft Accident Report NTSB/AAR-92/04, National Transportation Safety Board, Washington D.C.
- [25] NUREG. (2000). Technical basis and implementation guidelines for a technique for human event analysis (ATHEANA). NUREG-1624, Rev. 1, U.S. Nuclear Regulatory Commission, Washington, D.C., USA.
- [26] NUREG. (2005). Good practices for implementing human reliability analysis (HRA), NUREG-1792, U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Washington D.C., USA.
- [27] NUREG. (2007). ATHEANA User's guide, NUREG-1880, U.S. Nuclear Regulatory Commission, Office of Nuclear Safety Research, Washington, D.C., USA.
- [28] O'Conner, S., Reynolds, T. (1996). A review of aircraft maintenance and dispatch, ADAMS working document ADAMS-DRA-WD.01
- [29] Roelen, A.L.C., Wever, R. (2005). Accident scenarios for an integrated aviation safety model, NLR-CR-2005-560, NLR Amsterdam.
- [30] Roelen, A.L.C., Bos, T.J.J., Modiri, B. (2006). Aircraft maintenance program development needs improvement; An analysis of the maintenance program development process and its effect on safety for large transport aircraft in commercial service, NLR-CR-2006-351, NLR Amsterdam.
- [31] Roelen, A.L.C., Baren, G.B. van, Smeltink, J.W., Lin, P.H., Morales, O. (2007a). A generic flight crew performance model for application in a causal model of air transport. NLR-CR-2007-562, NLR Amsterdam.
- [32] Roelen, A.L.C., Baren, G.B. van, Lin, P.H., Morales, O. (2007b). A generic air traffic controller performance model for application in a causal model of air transport. NLR-CR-2007-593, NLR Amsterdam.
- [33] Rosekind, M., Gander, P.H., Miller, D.L., Gregory, K.B., Smith, R.M., Weldon, K.J., Co, E.L., McNally, K.L., Lebacqz, J.V. (1994). Fatigue in operational settings: examples from the aviation environment, Human Factors, Vol 36, No. 2, p. 327-338.
- [34] Swain, A.D., Guttman, H.E. (1983). Handbook of human reliability analysis with emphasis on nuclear power plant applications, Final report, NUREG/CR-1278-F, SAND80-0200, Sandia National Laboratories, Albuquerque, NM, USA.

## Appendix A ELICITATION PROTOCOL

### A.1 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.

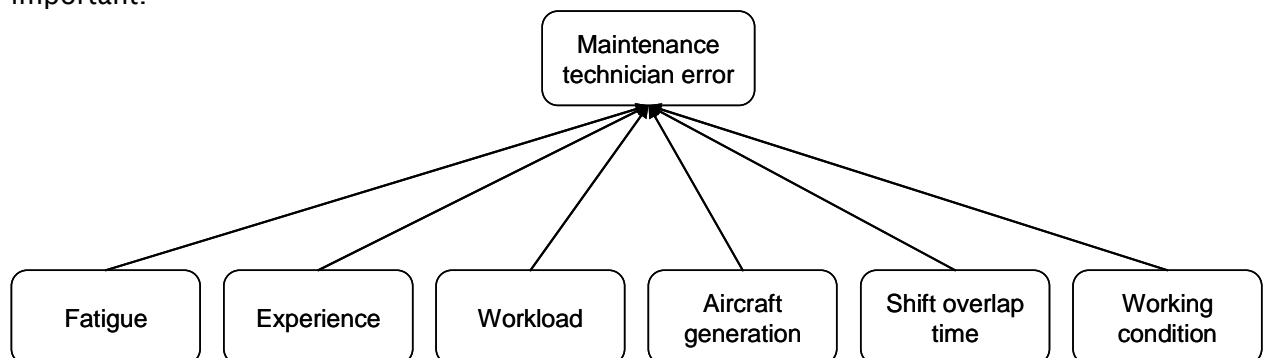
### A.2 OBJECTIVE

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.

### A.3 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 six 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 job	Number of errors per job
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	Years
Shift overlap time	Overlap time of two subsequent maintenance shifts.	Minutes
Aircraft generation	Four generations of aircraft, where 4 is the most recent generation	1 - 4
Workload	Delay in release to service of the aircraft.	Minutes

#### A.4 ELICITATION METHODOLOGY

The method employed is “structured expert judgment”, in which the judgment of operational experts on the values of parameters is gathered in a structured way.

Two kinds of information are gathered:

- information on the probability distributions of variables
- information on the influence of variables on the number of errors an maintenance technician makes during jobs.

#### A.5 INFORMATION ON THE PROBABILITY DISTRIBUTIONS OF VARIABLES

The information on the probability distributions of variables will be elicited via three numbers:

- The median value (50% percentile) of the distribution, i.e. if we would have 101 samples of the variable value, 50 of them should be below and 50 should be higher than the median value.
- The 5% percentile value, which can be interpreted as: it would surprise you if the variable has a value *lower* than this value.
- The 95% percentile value, which can be interpreted as: it would surprise you if the variable has a value *higher* than this value.

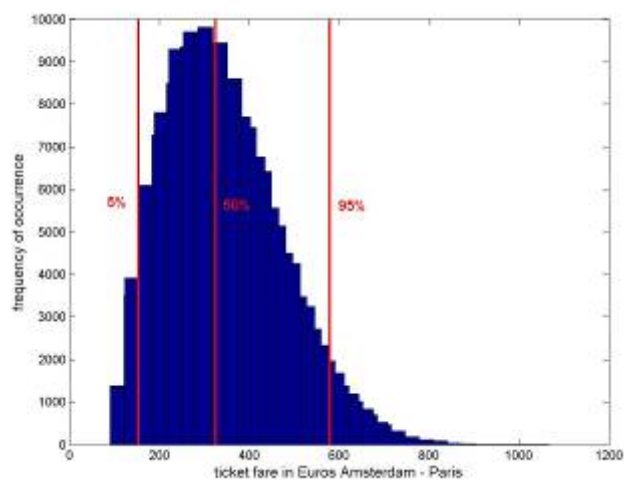
As an example, consider the following question:

*Consider 100,000 people travelling from Amsterdam to Paris by air. Please give your best estimate of the 5%, 50%, and 95% percentiles of the distribution of the*

*ticket fare (including taxes, considering all classes and airlines) that these 100,000 people have paid.*

Your answer could e.g. have been that you estimate that the median value will be 324 Euros, and it would surprise you if there are ticket prices lower than 100 Euros or higher than 500 Euros.

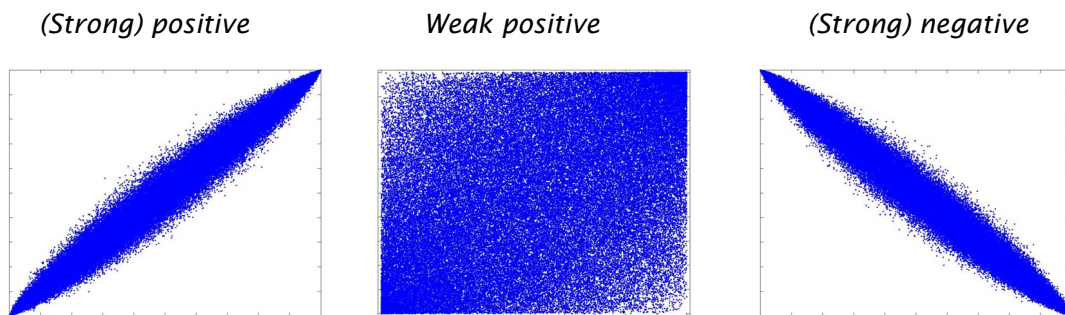
To illustrate the meaning of these percentile values, assume that we have gathered a set of data with 100,000 values of ticket fares. We can make a histogram of the ticket fares as shown in the picture below (this picture is based on artificial data, no real data). The 5%, 50% (median), and 95% percentile values of this data set are indicated by the red lines.



### A.5.1 INFORMATION ON THE INFLUENCE OF VARIABLES ON THE NUMBER OF ERRORS A MAINTENANCE TECHNICIAN MAKES DURING JOBS.

The purpose of the elicitation of dependencies is to obtain ‘rank correlation coefficients’, which we will refer to as ‘correlation’ in the following. Correlation tells something about the relationship between two variables. It can have a value between -1 and +1.

- If *high* values of a variable imply *high* values of another variable, there is a (strong) positive correlation, and the correlation will be close to +1 (the stronger the correlation, the closer it is to +1).
- If *high* values of a variable imply *low* values of another variable, there is a (strong) negative correlation, and the correlation will be close to -1.
- If there appears to be no such relation – the values of a variable do not imply anything on the values of another variable – the variables are called ‘independent’ and the correlation equals 0.



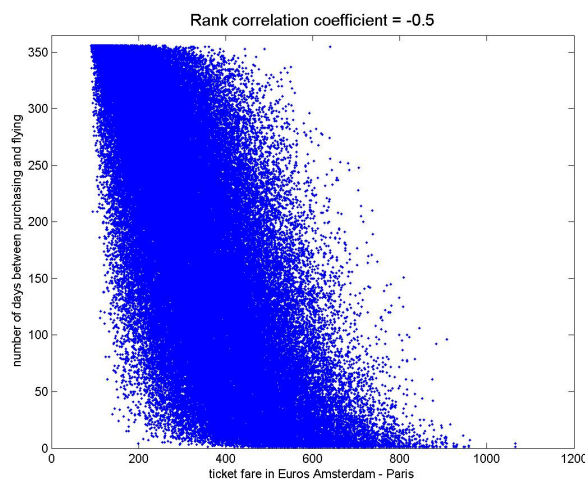
Going back to the example of tickets for Amsterdam – Paris, let’s consider three variables that might have an influence on the ticket fare:

- travel class (economy, business, first class)
- time between purchasing and flying (between 0 and 365 days)
- day of the week at which the flight takes place (1=Monday ... 7=Sunday)

Now, what would you expect about the correlation of each of these variables and the ticket fare? Presumably this:

- a higher travel class will result in a higher ticket fare (positive correlation);
- a higher number of days between purchasing and flying will result in a lower ticket fare (negative correlation);
- the day of the week will hardly have an influence on the ticket fare (correlation close to zero).

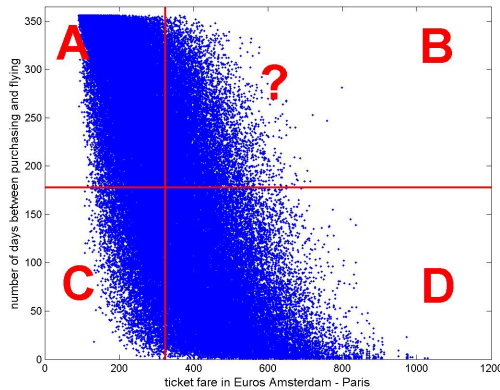
If in addition to the ticket fare we would have gathered the information on the three variables for all 100,000 tickets, we could have made plots like the one below, showing the relations between ticket fare (on the horizontal axis) and the number of days between purchasing and flying (based on artificial data).



As it may sometimes be difficult to directly estimate the correlation, we will use another way of asking, like:

*Suppose that out of the 100,000 tickets, we would have selected 50,000 for which the number of days between purchasing and flying is at least equal to its median value (178 days); what portion of these tickets do you expect to have a ticket fare above its median value (above 324 Euros)?*

Referring to the picture below, this question comes down to  
 “What portion of the blue dots above the red horizontal line lies at the right hand side of the vertical red line (in area B)?”



- A portion of 50% means that the mass (number of dots) in both areas A and B is equal: there is hardly a relation between the two variables (and the correlation will be close to 0);
- A portion between 0 and 50% implies that there is more mass in A than in B: a *higher* number of days between purchasing and flying will more likely result in a *lower* ticket fare: thus a negative relation (and a correlation between -1 and 0);
- A portion between 50% and 100% implies that there is less mass in A than in B: a *higher* number of days between purchasing and flying will more likely result in a *higher* ticket fare: thus a positive relation (and a correlation between 0 and 1);

We will use this way of asking only for one of the variables. We will pick the variable that you think has the strongest influence on the number of errors that a maintenance technician makes per jobs. For the other variables we will ask for the “relative amount of influence”.

Going back to the example of tickets for Amsterdam – Paris once again, this would mean that we first need to order the variables based on the absolute value of the correlation, i.e., make a ranking based on the influence that we think the variables will have, disregarding the ‘sign’ (positive or negative) of the correlation.

From our example data set we found that the correlation of the ticket fare with travel class was 0.84, with number of days between purchasing and flying was -0.50, and with day of the week was 0.19. So the order would be (from strongest to weakest):

- 1) Travel class (absolute correlation = 0.84)
- 2) Number of days between purchasing and flying (absolute correlation = 0.50)
- 3) Day of the week (absolute correlation = 0.19)

Roughly speaking, one could say that the “relative amount of influence” that the second variable has compared to the first one is 60% ( $\sim 0.50/0.84$ ) and of the third variable compared to the first one is 20% ( $\sim 0.19/0.84$ ). Note that the ‘total amount of influence’ is limited. Starting from a certain correlation for the variable with the ‘strongest influence’, it may not be possible

for all other variables to have the same amount of influence and at some point, there may be no more influence left over for a certain variable. Observe also, that the sum of the “relative amount of influence” as expressed by the ratios of each one of the rank correlations to the largest one does not necessarily have to add to 100%.

You may think of the ‘total amount of influence’ as a pie that needs to be distributed over a number of persons. If the strongest one takes a large part of it, the rest of the people have to distribute the remaining small chunk amongst them.

During the exercise, we will check if your estimates for the “relative amount of influence” of the variables are still valid in terms of not exceeding the “total amount of influence”.

#### A.6 CONFIDENTIALITY/ FEEDBACK

After the individual elicitation, we will report the results including the underlying argumentation and any other discussions and these will be send to you for review.

The names and qualifications of the experts will be published in the final report, as will the individual expert assessments and all information relevant, but the link between individual expert assessments and their identity will be removed in the published reports.

## Part II: 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*.

### A.7 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 closed
- Material left in aircraft
- Required service not performed

<i>Q0</i>	<i>Job duration</i>
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?	

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

<i>Q2</i>	<i>Number of Maintenance technician errors</i>	
What are your estimates for the 5 <sup>th</sup> , 50 <sup>th</sup> , and 95 <sup>th</sup> percentiles of the distribution of number of errors that each Maintenance technician makes per <b>median</b> (50% quantile in previous question) number of jobs?		
_____	_____	_____
5%	50%	95%

<i>Q3</i>	<i>Fatigue</i>					
Consider 10,000 maintenance technicians taken randomly from 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.

<i>Q4</i>	<i>Shift overlap time</i>		
What are your estimates for the 5 <sup>th</sup> , 50 <sup>th</sup> , and 95 <sup>th</sup> percentiles of the distribution of the time available to transfer a job (minutes)?			
_____	_____	_____	_____
5%	50%	95%	

<i>Q5</i>	<i>Workload</i>		
What are your estimates for the 5 <sup>th</sup> , 50 <sup>th</sup> , and 95 <sup>th</sup> percentiles of the distribution of the delay in release of the aircraft (minutes)?			
_____	_____	_____	_____
5%	50%	95%	

## A.8 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, 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.	
Experience	Number of years in current position	
Shift overlap time	Overlap time of two subsequent maintenance shifts.	
Aircraft generation	Four generations of aircraft, where 4 is the most recent generation	
Workload	Delay in release to service of the aircraft.	

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

<b>Q6</b>
<p>Suppose that 20,000 <b>Maintenance technician</b> 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).</p> <p>What portion of these 10,000 <b>Maintenance technicians</b> will commit more than the median number of errors per median (50% quantile in question <b>Q1</b>) jobs. Observe that the median number of Maintenance technician errors was specified in question <b>Q2</b>?</p>

<b>For each pair below indicate which of the two effects is the most dominant?</b>	
0	When the work is performed at the ramp (outside - 1), <b>higher</b> values of Maintenance technician error are observed, or
0	When the work is performed at the ramp (outside - 1), <b>lower</b> 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 <b>low</b> 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 <b>high</b> values of Maintenance technician error.
0	High values of experience relate to <b>high</b> values of Maintenance technician error, or
0	High values of experience relate to <b>low</b> values of Maintenance technician error.
0	High values of shift overlap time relate to <b>high</b> values of Maintenance technician error, or
0	High values of shift overlap time relate to <b>low</b> values of Maintenance technician error.
0	The most recent generation of aircrafts relates to <b>high</b> values of Maintenance technician error, or
0	The most recent generation of aircrafts relates to <b>low</b> values of Maintenance technician error.
0	High values of workload relate to <b>high</b> values of maintenance technician error, or
0	High values of workload relate to <b>low</b> 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:

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

### Part III: 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%).

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

<b>CQ2</b>		
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%

<b>CQ3</b>		
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%

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

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

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

<b>CQ7</b>		
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 <i>never</i> left a tool or torch behind in an aircraft?		
5%	50%	95%

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

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

<b>CQ10</b>		
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%

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

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

## Appendix B ANSWERS TO CALIBRATION QUESTIONS

CQ1: How many parts does a Boeing 747-400 have including fasteners?

Answer: 6 million.

Source: [http://www.boeing.com/commercial/747family/pf/pf\\_facts.html](http://www.boeing.com/commercial/747family/pf/pf_facts.html)

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”.)

Answer: 12.4 minutes.

Source: Eurocontrol. (2007). Digest – Annual 2006 Delays to Air Transport in Europe, Central Office for Delay Analysis, Eurocontrol, Brussels.

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”.)

Answer: 17.5 bar.

Source: Airbus A-330 maintenance manual.

CQ4: What is the maximum capacity of the Boeing-777 200 ER centre fuel tank? (Give your answer in “litres”.)

Answer: 98,800 litres

Source: AAIB. (2004). Special bulletin S2/2004, Ref. EW/C2004/06/01, Air Accident Investigation Branch, Aldershot, UK.

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

Answer: 0.006 / 1000 flight hours

Source: UTC. (2002). FAA certifies new compressor case for PW 4000, Press release, United Technologies, East Hartford, Connecticut.

CQ6: What was the average number of daily departures from Schiphol in the 4th quarter of 2006?

Answer: 592.

Source: LVNL.

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?

Answer: 45.8%

Source: Hobbs, A., Williamson, A. (1997). Aircraft maintenance safety survey - results. Department of Transport and Regional Services, Australian Transport Safety Bureau, Australia.

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

Answer: 43,000

Source: Airclaims.

CQ9: What is the Maximum Take-off Weight (MTOW) of an Airbus A380? (Give your answer in “kilograms”.)

Answer: 560,000 kg.

Source: Airbus (2008). A380 Airplane Characteristics, revision 6, Airbus S.A.S., Technical Data Support and Services, Blagnac, France.

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”.)

Answer: 77.7 (76.4, 76.4, 79.0, 79.0 resp.)

Source: Fokker. (1975). Fokker F.28 Engineers Guise, Fokker -VFW International, Schiphol Oost, the Netherlands.

Fokker. (1983). F28 Pilot's Training Manual, Fokker B.V., Flight Crew Training Department, Schiphol, the Netherlands.

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

Answer: 73.5 (65, 79, 65, 85 resp.)

Source: Fokker. (1975). Fokker F.28 Engineers Guise, Fokker -VFW International, Schiphol Oost, the Netherlands.

Fokker. (1983). F28 Pilot's Training Manual, Fokker B.V., Flight Crew Training Department, Schiphol, the Netherlands.

CQ12: What was the total number of orders of Airbus A380-800 in 2007?

Answer: 33

Source: Airbus Annual Review 2007, <http://www.airbus.com/en/presscentre/>.

## Appendix C EXPERT ELICITATION RESULTS

### Elicitation of probability distributions

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?			
<b>5</b> 5%	<b>8</b> 50%	<b>12</b> 95%	

Q1	Number of Maintenance technician jobs		
What are your estimates for the 5 <sup>th</sup> , 50 <sup>th</sup> , and 95 <sup>th</sup> percentiles of the distribution of number of jobs that each Maintenance technician makes during a day?			
<b>1</b> 5%	<b>4</b> 50%	<b>8</b> 95%	

Q2	Number of Maintenance technician errors		
What are your estimates for the 5 <sup>th</sup> , 50 <sup>th</sup> , and 95 <sup>th</sup> percentiles of the distribution of number of errors that each Maintenance technician makes per <b>median</b> (50% quantile in previous question) number of jobs * <b>median</b> average duration in hrs of a job?			
<b>0.5</b> 5%	<b>1</b> 50%	<b>3</b> 95%	

Q3	Fatigue						
Consider 10,000 maintenance technicians taken randomly from the population. How many of them (percentage) would fall in each one of the seven following classes?							
<b>41%</b>	<b>25%</b>	<b>30%</b>	<b>0%</b>	<b>0%</b>	<b>2%</b>	<b>2%</b>	
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.	

From a mathematical point of view the 0% values for the 4<sup>th</sup> and 5<sup>th</sup> class are undesirable. It was therefore decided by the analysts to use the following values in the model:

<b>41%</b>	<b>25%</b>	<b>29%</b>	<b>0.5%</b>	<b>0.5%</b>	<b>2%</b>	<b>2%</b>
------------	------------	------------	-------------	-------------	-----------	-----------

<b>Q4</b>	<b>Shift overlap time</b>		
What are your estimates for the 5 <sup>th</sup> , 50 <sup>th</sup> , and 95 <sup>th</sup> percentiles of the distribution of the time available to transfer a job (minutes)?			
<b>5</b>	<b>10</b>	<b>20</b>	
5%	50%	95%	

<b>Q5</b>	<b>Workload</b>		
What are your estimates for the 5 <sup>th</sup> , 50 <sup>th</sup> , and 95 <sup>th</sup> percentiles of the distribution of the delay in release of the aircraft (hours)?			
<b>½ h</b>	<b>2 h</b>	<b>5 h</b>	
5%	50%	95%	

<b>Q6</b>	<b>Working Conditions</b>		
Consider 10,000 maintenance operations taken randomly from your current job. How many of these operations would have to be performed outside the hangar?			
<b>1</b>	<b>5</b>	<b>1000</b>	
5%	50%	95%	

#### Elicitation of dependence information

In order to gain insight in the relationship between variable **Number of Maintenance Technician Errors** and the variables: Working condition, Fatigue, Experience, Shift Overlap Time, Aircraft generation and Workload, the expert was asked to assign number 1 to the variable that he considered 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. The results are presented in the table below:

<b>Variable</b>	<b>Rank</b>
Working condition	<b>4</b>
Fatigue	<b>1</b>
Experience	<b>1</b>
Shift overlap time	<b>3</b>
Aircraft generation	<b>3</b>
Workload	<b>2</b>

The expert chose Fatigue and Experience as variables which have the strongest influence on the number of Maintenance technician errors. In the further analysis Fatigue was chosen as variable which has the strongest influence on the number of Maintenance technician errors. For this variable the expert was asked the following question:

<b>Q7</b>
<p>Suppose that 20,000 <b>Maintenance technician</b> are randomly chosen from our total population. Out of those, 10,000 are selected for which the chosen variable (namely: <b>Fatigue</b>) has values above its median value (or above a certain percentile).</p> <p>What portion of these 10,000 <b>Maintenance technicians</b> will commit more than the median number of errors per median (50% quantile in question <b>Q1</b>) jobs. Observe that the median number of Maintenance technician errors was specified in question <b>Q2</b>)?</p>
<b>7000</b>

The expert is now asked to express the influence of the remaining variables (other than Fatigue) as a portion of the influence of the variable which our expert ranked the highest i.e Fatigue. Also, the dominant direction (positive or negative) of the correlation is asked. The results are presented in the table below.

<b>Node</b>	<b>Rank</b>	<b>Influence as a portion of the influence of the highest ranked variable (0 – 100%) The sum does not necessarily add to 100%</b>	<b>Direction of the correlation (positive/negative)</b>
Working condition	<b>4</b>	<b>10%</b>	<b>negative</b>
Fatigue	<b>1</b>		<b>positive</b>
Experience	<b>1</b>	<b>100%</b>	<b>negative</b>
Shift overlap time	<b>3</b>	<b>30%</b>	<b>negative</b>
Aircraft generation	<b>3</b>	<b>30%</b>	<b>negative</b>
Workload	<b>2</b>	<b>50%</b>	<b>positive</b>

#### Elicitation of calibration variables

<b>CQ1</b>		
How many parts does a Boeing 747-400 have including fasteners?		
<b>1e6</b>	<b>2e6</b>	<b>5e6</b>
5%	50%	95%

<b>CQ2</b>		
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".)		
<b>5</b>	<b>15</b>	<b>30</b>
5%	50%	95%

<b>CQ3</b>		
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".)		
<b>1750</b>	<b>3000</b>	<b>4000</b>
5%	50%	95%

<b>CQ4</b>		
What is the maximum capacity of the Boeing-777 200 ER centre fuel tank? (Give your answer in "litres".)		
<b>1e4</b>	<b>2.5e4</b>	<b>4e4</b>
5%	50%	95%

<b>CQ5</b>		
What was up to December 2002 the in-flight shutdown rate for the 94inch fan PW400 engine?		
<b>1/5000hours</b>	<b>1/10000hours</b>	<b>1/25000hours</b>
5%	50%	95%

<b>CQ6</b>		
What was the average number of daily departures from Schiphol in the 4 <sup>th</sup> quarter of 2006?		
<b>500</b>	<b>750</b>	<b>1000</b>
5%	50%	95%

<b>CQ7</b>		
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 <i>never</i> left a tool or torch behind in an aircraft?		
<b>80</b>	<b>90</b>	<b>100</b>
5%	50%	95%

<b>CQ8</b>		
How many flight cycles had been accumulated by the lead-time Fokker 100 on 21 December 2006?		
<b>1e4</b>	<b>2.5e4</b>	<b>5e4</b>
5%	50%	95%

<b>CQ9</b>		
What is the Maximum Take-off Weight (MTOW) of an Airbus A380? (Give your answer in "kilograms".)		
<b>4e5</b>	<b>5e5</b>	<b>6e5</b>
5%	50%	95%

<b>CQ10</b>		
What is the average of the wing area in Fokker F.28 series (F28-1000, 2000,3000,4000)? (Give your answer in "square meters".)		
<b>74.99999</b>	<b>75</b>	<b>100</b>
5%	50%	95%

<b>CQ11</b>		
What is the average of the minimal passenger capacity of the Fokker F.28 series (F28-1000, 2000,3000,4000)?		
<b>30</b>	<b>40</b>	<b>50</b>
5%	50%	95%

<b>CQ12</b>		
What was the total number of orders of Airbus A380-800 in 2007?		
<b>50</b>	<b>75</b>	<b>100</b>
5%	50%	95%

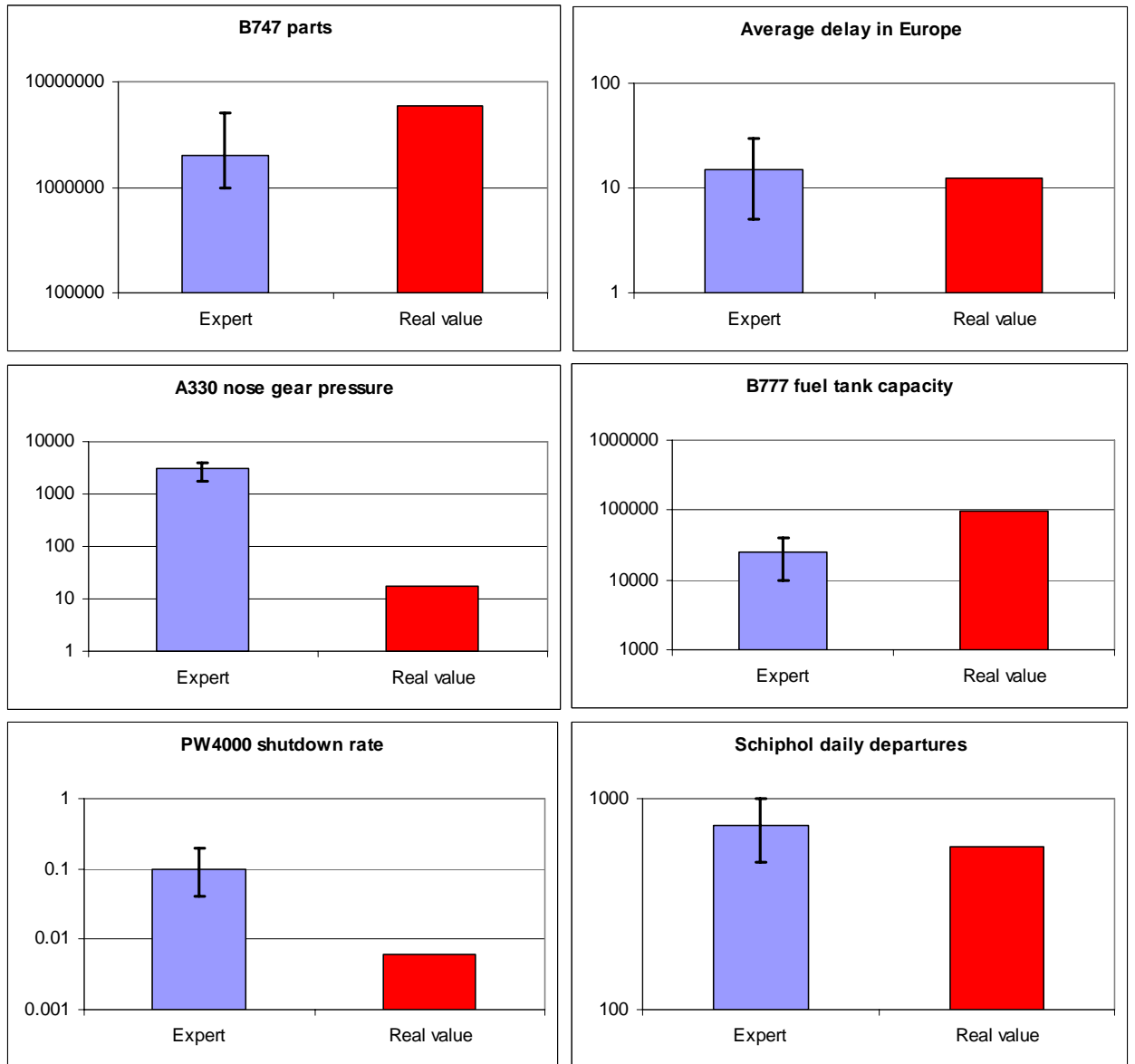


Figure C1: Range graphs for expert's assessments in calibration variables 1-6

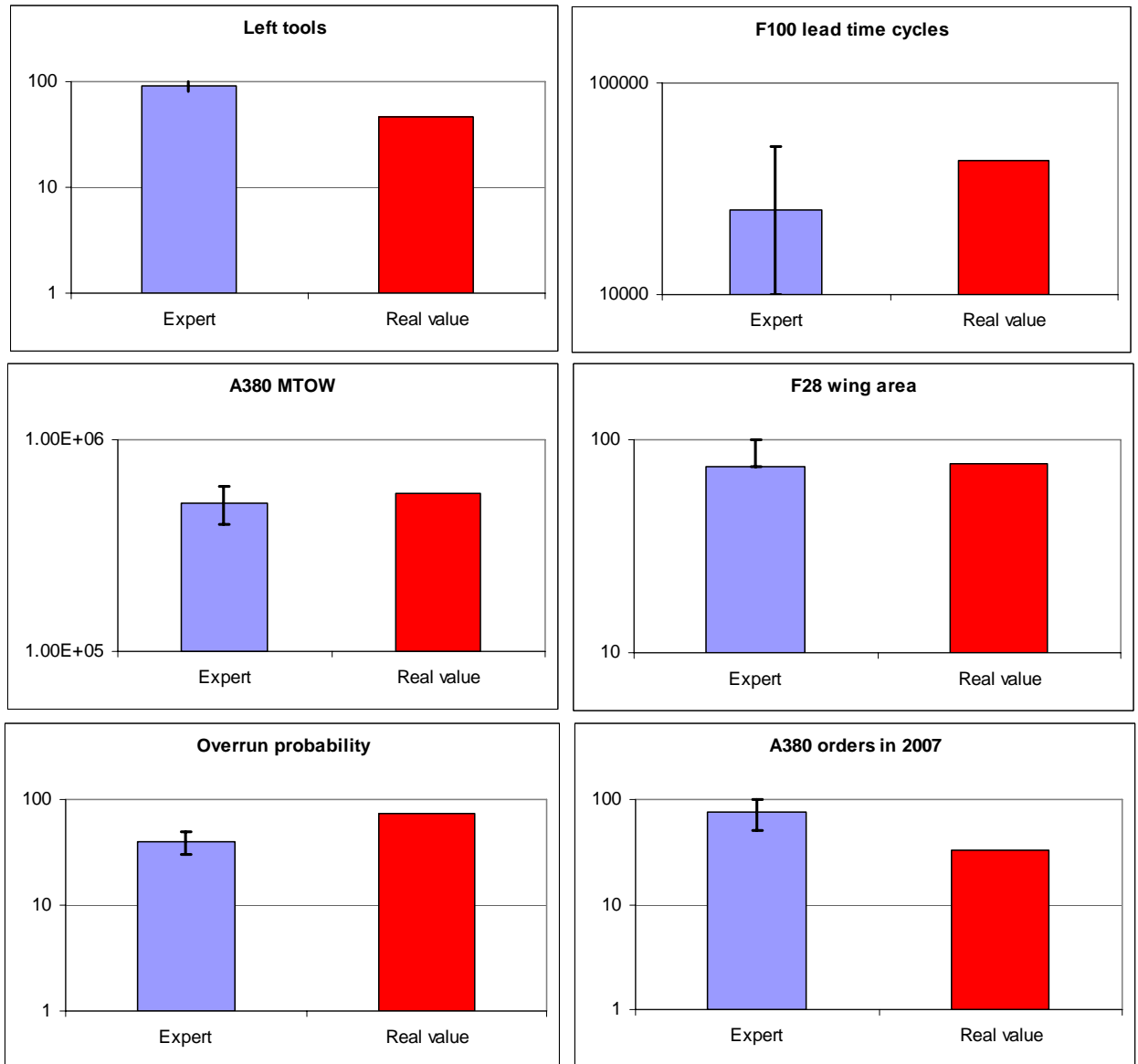


Figure C2: Range graphs for expert's assessments in calibration variables 6-12