

Probabilistic Inversion
in Uncertainty Analysis
and related topics

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van Rector Magnificus prof. dr. ir. J.T. Fokkema,
voorzitter van het College van Promoties,
in het openbaar te verdedigen op 23 april 2002 om 16:00 uur

door

Bernard Cornelis Pieter KRAAN.

Wiskundig Ingenieur
geboren te Den Helder.

Dit proefschrift is goedgekeurd door de promotor
Prof. dr. R.M. Cooke.

Samenstelling promotiecommissie:

Prof. dr. R.M. Cooke, Technische Universiteit Delft, promotor
Prof. dr. T.J. Bedford, University of Strathclyde, Scotland
Prof. dr. F.M. Dekking, Technische Universiteit Delft
Prof. dr. S. Hora, University of Hawaii at Hilo, USA
Dr. ir. L.H.J. Goossens, Technische Universiteit Delft
Dr. M.D. McKay, Los Alamos National Laboratories, USA
Dr. G.N. Kelly, Commission of European Communities, België
Prof. dr. T. Mazzuchi, Technische Universiteit Delft

ISBN 90-9015710-7

Omslagontwerp: Annette Moelker & Ronald Cozijn

Copyright © 2002 by B.C.P. Kraan

All rights reserved. No part of the material protected by copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the author.

Printed by Burghout Drukkerij en Designstudio in Hippolytushoef, the Netherlands.

Summary

Over the last decade uncertainty analysis has become more important in the decision process. It is acknowledged that if decisions are based on mathematical models it is wise to investigate the uncertainty associated with model input parameters, because a slight change in the values of the model input parameters may result in a different decision.

In performing uncertainty analysis a distribution on the values of the uncertain model input parameters is required. Using sampling techniques this distribution is propagated through the model to obtain a distribution on the values of the model output parameters. The crucial part in performing uncertainty analysis is the determination of the distribution on the uncertain model input parameters. If experimental data on these parameters are available this may be used to derive the distribution. In case very little or no experimental data on these parameters are available a different approach is needed to obtain the distribution. This thesis deals with the latter situation.

Confronted with a situation where very little or no experimental data are available it is reasonable to consult experts and take account of their views. Experts will have the best overview of literature, will have access to mathematical models and are able to discriminate, interpolate and extrapolate the experimental data available. In other words it would be helpful if a methodology could be developed in which experts quantify their degrees of belief in a structured, open and defensible way. This led to the development of the Structured Expert Judgment Elicitation Methodology, [12].

In this thesis new mathematical techniques are introduced for the Structured Expert Judgment Elicitation Methodology. The thesis can be divided into four parts: a Probabilistic Inversion part, a part which deals with dependencies in uncertainty analysis, a part which is concerned with performance measures which determine the capability of an expert in how well he/she is able to quantify their degrees of belief and a part which discusses the modeling of uncertainty. Examples from the Joint CEC/USNRC Uncertainty Analysis (Contract F13P-CT92-0023 and 93-ET-001) were used to illustrate the mathematical techniques developed.

Probabilistic inversion

A key element of the Structured Expert Judgment Elicitation Methodology is that experts are queried only about potentially measurable quantities. Hence, it could be that uncertain model input parameters are not suited to be subjected to expert judgment. However, by querying information on potentially measurable quantities and using probabilistic inversion techniques, information on the uncertain model input parameters can be obtained. In Chapter 1 the foundations and implementations of probabilistic inversion techniques are presented.

Dependencies

In order to perform an uncertainty analysis, a distribution on the uncertain model input parameters has to be available. Usually this distribution is constructed using the marginal distributions of uncertain model input parameters and dependency measures among the input parameters. In Chapter 2 a strategy to query dependence from experts is selected which, for reasons explained, is considered to be the most appropriate. Currently the Structured Expert Judgment Elicitation Methodology combines the information of experts for each queried quantity only. Two strategies are introduced which combine expert information and dependence information. Finally it is shown that dependencies between quantities can be incorporated easily in probabilistic inversion.

Calibration with uncertain observations

Under the Structured Expert Judgment Elicitation Methodology, performance measures have been developed which assess/quantify (in a statistical sense) the capability of experts to quantify their degrees of belief. Experts are asked to quantify their knowledge for quantities for which the outcome has been measured by an experiment. This outcome is not available to the expert, but only for the project staff. If a reasonable number of such questions are available, the performance of an expert can be measured statistically. Since reproducibility of experimental results is often not possible and different measurement procedures may result in different outcomes, the outcome of an experiment is uncertain itself. In Chapter 3 the performance measures are extended by allowing the outcome of the experiment to be uncertain as well. The effect of taking account of measurement variability is shown by making use of examples taken from the Joint CEC/USNRC Accident Consequence Code Uncertainty Analysis and a study conducted for the Ministry of Transport, Public Works and Water Management.

Modeling Uncertainty

Modeling uncertainty deals with questions regarding the quantification of uncertainty of model input parameters and the effect/considerations on how to reduce the complexity of a mathematical model before performing uncertainty analysis. Special attention is given to a special class of acyclic compartmental models (ACMs). ACMs are widely used in environmental modeling and due to their graphical representation may look easy to use. However ACMs need special attention when performing uncertainty analysis. Like ACMs, influence diagrams (IDs) are acyclic graphs as well. In Chapter 4 the relationship between ACMs and IDs is investigated, which resulted into a decomposition strategy of complex ACMs. The decomposition will influence the line of questioning. However, the line of questioning should be in compliance with the Structured Expert Judgment Elicitation Methodology at all times.

Application of Results

Each chapter introduces new mathematical techniques to support the Structured Expert Judgment Elicitation Methodology and the majority of which have been validated and shown practicable in the Joint CEC/USNRC Uncertainty Analysis [25], [26], [27], [28], [29], [30], a study for the Ministry of Transport, Public Works and Water Management [21] and in the PhD thesis ‘Uncertainty in predictions of thermal comfort in buildings’ [19]. In performing an uncertainty analysis, these mathematical techniques will be used in the reverse order as presented. Firstly, one has to think deeply and possibly use the considerations introduced in Chapter 4 on how the uncertainty can be modeled satisfactorily. Secondly, if experts are used to quantify the uncertainty and their performance is measured using experimental results, it is only fair, for reasons explained in Chapter 3, to take account of the measurement variability in these experimental results. Thirdly, the expert’s assessments will be combined using the elicited dependence information and strategies introduced in Chapter 2. And if necessary, the probabilistic inversion techniques developed in Chapter 1 can be used to obtain a distribution on target variables.

Table of Contents

Acknowledgements	1
Introduction	3
1 Probabilistic Inversion	9
1.1 Foundations	9
1.1.1 The measure $\tilde{\mu}$	13
1.1.2 Approximating $\tilde{\mu}$ by $\tilde{\mu}_k$	17
1.1.3 Approximating $\tilde{\mu}_k$ by $\tilde{\mu}_{n',k}$	22
1.1.4 Example: revisited	25
1.2 Probabilistic Inversion Implementations	25
1.2.1 PARFUM	26
1.2.2 PREJUDICE	28
1.2.3 PARFUM versus PREJUDICE	32
1.3 Background measure	36
1.3.1 Uniform background measure	36
1.3.2 Results	38
1.4 Iterative PREJUDICE	40
1.4.1 Iterative PREJUDICE: Uniform background measure	45
1.4.2 Results	48
1.5 CP solvers	58
1.6 Efficient version of PREJUDICE	60
1.7 Conclusions	63
2 Dependencies	65
2.1 Dependence elicitation techniques	65
2.1.1 Overview literature	65
2.1.2 Dependence Elicitation and Structured Expert Judgment Elicitation Methodology	68
2.1.3 $\pi_{\frac{1}{2},\frac{1}{2}}(Y_1, Y_2)$ and Spearman's ρ_{Y_1, Y_2}	70
2.2 Combining Expert Judgments	71
2.2.1 Strategy 1	71
2.2.2 Strategy 2	73

2.2.3	Performance based weights	74
2.2.4	Examples	74
2.3	Dependencies and Probabilistic inversion	81
2.3.1	Example	81
2.4	Conclusions	84
3	Calibration with Uncertain Observations	89
3.1	Step 1: The distribution of Z_j	91
3.1.1	Relinf Y : Minimum relative information for Y	92
3.1.2	Relinf Z : Minimum relative information for Z_j	93
3.1.3	Limiting behavior	93
3.2	Step 2: Scoring calibration based on Z_j	94
3.2.1	Scoring with individual values	94
3.2.2	Scoring with sampling distributions	96
3.3	Examples	97
3.3.1	Dispersion example	97
3.3.2	Ministry of Transport, Public Works and Water Management	102
3.4	Conclusion	105
4	Modeling Uncertainty	109
4.1	How to quantify uncertainties?	109
4.2	Uncertainty Capture	111
4.3	ACMs and influence diagrams	114
4.3.1	Example	114
4.3.2	Systemic retention of Pu in the human body	118
4.4	Cyclic Compartmental Models	119
4.5	Implications to Methodology	119
4.6	Conclusions	120
5	Conclusions	121
	Reference	127
	Appendices	
A	ϵ-neighborhood vs. Bin combinations	133
A.1	ϵ -neighborhood sampling scheme	133
A.2	Bin-combination sampling scheme	135
A.3	Conclusion	136
B	Determination of M.	139
B.1	Dispersion coefficient example	140
B.2	Results: Step 1	140
B.2.1	Results: Step 2	141
B.2.2	Discussion of results	142

B.3	Lung morbidity	144
B.3.1	Model	144
B.3.2	Target variables & Elicitation variables	144
B.3.3	Results: Step 1	144
B.3.4	Results: Step 2	146
B.3.5	Discussion of results	148
B.4	Systemic retention of Sr in the Human body	150
B.4.1	Compartmental model	150
B.4.2	Target variables & Elicitation variables	151
B.4.3	Results: Step 1	151
B.4.4	Results: Step 2	155
B.4.5	Discussion of results	155
B.5	Conclusion	158
C	Reduction of Dimension	161
D	Minimal solution	163
E	Dependencies: assessments and results.	165
E.1	The $\rho\pi$ -table	165
E.2	Example 1	167
E.3	Example 2	171
F	Calibration with Uncertain observations: data	177
F.1	Dispersion example	177
F.2	Public Works and Water Management example	184
	Samenvatting	193
	Dankwoord	197
	Curriculum Vitae	201



Acknowledgements

The author would like to acknowledge Prof dr. R.M. Cooke and dr. ir. L.H.J. Goossens for their invaluable contributions and support during the period that this thesis was written. Prof. dr. T. Bedford of the University of Strathclyde, Scotland is acknowledged for his contribution in Chapter 1. A large amount of time has been devoted on formulating probabilistic inversion in the framework of measure theory. Thanks to the investigations of drs. N.P. van Elst and drs. F.L. Härte, I was able to make a jump start on the issues discussed in Chapter 3. The contribution on Prof. dr. F.M. Dekking of Delft University of Technology is acknowledged for pointing out and fixing the proof in Chapter 4. The author would also like to acknowledge dr. M. McKay of Los Alamos National Laboratories for his comments and the interesting/revealing discussions on the different subjects. From the National Radiological Protection Board, the numerous discussion with dr. J.A. Jones and Mrs. dr. J.Brown are greatly appreciated and were regarded as fruitful, very challenging. Dr. G.N. Kelly from the Commission of European Communities is acknowledged for making it possible to become a PhD student at all and to participate in the challenging Joint CEC/USNRC Uncertainty Analysis project.

Special thanks go out to Annette Moelker and Ronald Cozijn for their beautiful cover design.

The author also greatly appreciate the assistance of Ms. C. Bosman and Ms. D. Droog of Delft University of Technology.

Last but not least, the author would like to acknowledge his parents for their devoted support. This one is for you.



Introduction

It is becoming common practice to perform uncertainty analysis on complex models. In an uncertainty analysis one works with a joint probability distribution over the model input parameters instead of ‘nominal values’ for these parameter. With distributional input, the model does not make deterministic predictions, rather a distribution over the model outcomes is obtained. Concepts like model validation, model calibration and sensitivity of model output to input, take on a new meaning when viewed from the perspective of uncertainty analysis. A growing field of physicist, engineers, mathematicians and modelers is focusing on this cluster of issues [36]. The question remains how to determine the distribution on the uncertain model input parameters. In many situations the uncertain model input parameters do not correspond to physically measurable/observable quantities. Moreover the uncertainties may exhibit strong dependencies. In case uncertain model input parameters do not correspond to physically measurable/observable quantities, the model outcome is usually a physically measurable/observable quantity, however, there may be situations where the experimental data available on model outcomes are scarce, conflicting and exhibit wide uncertainty.

A possible way to proceed in this type of situation is to consult experts in the field. They will have the best overview of the literature, will have access to mathematical models and are able to discriminate, interpolate and extrapolate the experimental results available. In other words, experts can be used to express/quantify their views in cases where experimental data on model outcomes are scarce or conflicting. In order to make this process transparent and defensible the Structured Expert Judgment Elicitation Methodology [12] has been developed. A key element of the Structured Expert Judgment Elicitation Methodology is that experts can express their degrees of belief only for quantities which are observable; a quantity is regarded as observable if there is a physically (not necessarily practical) procedure for determining its value. The variables for which the experts quantify their degrees of beliefs are called *elicitation* variables and the uncertain model input parameters are termed *target* variables. In case the target variables are regarded as non-observable, elicitation variables will have to be formulated which are related to the target variables and which

are familiar to the experts.

Since experts may have different viewpoints on how to value experimental results, what model is appropriate etc., it is recommended to have a number of experts who individually express their degrees of beliefs for the elicitation variables. By allowing a diversity of viewpoints, a better understanding of the ‘true’ uncertainty associated with the elicitation variable is obtained. Under the Structured Expert Judgment Elicitation Methodology the experts’ uncertainty distributions are aggregated to obtain the uncertainty distribution of the Decision Maker (DM); if $N_{\text{exp.}}$ experts have assessed elicitation variable Y , the Decision Maker (DM) ’s density is given by:

$$f_{\text{DM,eq},Y} = \sum_{j=1}^{N_{\text{exp.}}} w_j f_{j,Y} \quad (1)$$

where w_j and $f_{j,Y}$ are the weight of expert j and the density associated with expert j ’s assessments for elicitation variable Y , respectively.

This thesis introduces new mathematical techniques for the Structured Expert Judgment Elicitation Methodology, which have been developed and applied under the joint CEC/USNRC Accident Consequence Code Uncertainty Analysis using Expert Judgment (Contract F13P-CT92-0023 and 93-ET-001) [25], [26], [27], [28], [29], [30]. The newly developed mathematical techniques concern Probabilistic Inversion, the elicitation and application of dependencies and the development of a special set of weights w_j to be used in Equation 1. Finally, in the course of the Joint CEC/USNRC Uncertainty Analysis a large variety of models was encountered for which the uncertainty had to be quantified. The final chapter of this thesis is a discussion on experiences gained and introduces a mathematical technique which may be applied in the future on how to model uncertainty.

Probabilistic Inversion

Chapter 1 introduces the mathematical technique Probabilistic Inversion (PI). In short, the objective of PI is to take a distribution representing the uncertainty on certain observables, and to translate this uncertainty to uncertainty on the target variables of a certain model. Note that the distribution on certain observables does not have to be obtained from structured expert judgment but also may be obtained from experimental data. In Section 1.1, PI is formulated in terms of a measure theoretical problem. Based on this formulation, certain approximations are introduced on which the implementation will be based. In sections 1.2.1 and 1.2.2 the implementations PARFUM (PARAmeter Fitting for Uncertain Models) and PREJUDICE (PRocessing Expert JUDgment Into Code paramETers) are introduced, respectively. PARFUM is based on an idea of Cooke [10] and PREJUDICE is based on an idea proposed by Hora and Young [27] and PARFUM. The

dispersion coefficient example from [27] will be used to illustrate the different steps involved in PI; lateral plume spread σ_y is modeled as a power law function of downwind distance z from the source of a release:

$$\sigma_y(z) = A_y z^{B_y} \quad (2)$$

where the coefficients A_y and B_y depend on the stability of the atmosphere at the time of the release. Clearly there will be more variables, like wind variation, plume meander, surface roughness and vertical wind profile, which influence lateral plume spread. Although it hasn't been derived from underlying physical laws and the coefficients are fit to data from tracer experiments, Expression 2 is recognized to capture the uncertainty associated with lateral plume spread well enough. In performing uncertainty analysis on Expression 2, a joint distribution on A_y and B_y is required; with the use of this distribution the uncertainty on σ_y for *any* desired downwind distance z can be determined. In the dispersion coefficient example A_y and B_y are the target variables. Since the dimension of A_y must be [meters]^{1-B_y}, these parameters do not have any obvious physical interpretation, so that no experimental data are available to directly determine a distribution on A_y and B_y . According to the Structured Expert Judgment Elicitation Methodology, the target variables A_y and B_y are non observables and cannot serve as elicitation variables. However, the lateral plume spread σ_y is an observable and could be determined from tracer experiments, hence σ_y became the elicitation variable. Experts were queried to quantify their uncertainty on $\sigma_y(z_i)$ for down wind distances z_1, \dots, z_n . Note that since the expert is only asked about uncertainty on observables, he/she may use their own preferred mathematical model for lateral plume spread if he/she wishes. The expert data do not rely on any particular model. This way of querying information from experts is also known as model independent elicitation. Finally, it is noted that the uncertainty on the target variables, obtained by PI, will include both parameter uncertainty and model uncertainty.

Summarized, the PI problem for the dispersion coefficient example is to translate the distributions on $\sigma_y(z_i)$ into a distribution on (A_y, B_y) using Expression 2, such that the push-forward distributions $A_y z_i^{B_y}$ comply with the distributions on $\sigma_y(z_i)$ for $i = 1, \dots, n$. To numerically illustrate this PI problem, the uncertainty distributions for $\sigma_y(z_i)$ ($i = 1, \dots, 5$) for the DM have been taken from [27], see Table 1. The uncertainty distributions are available in the form of 5%, 50% and 95% quantile points and concern stability category C according to the Pasquill-Gifford classification scheme¹.

¹Stability category C of the Pasquill-Gifford classification scheme corresponds to neutral (nor stable, nor unstable) weather conditions.

Quantile	$\sigma_y(z_i)$				
	500 m.	1 km.	3 km.	10 km.	30 km.
5%	3.30e+1	6.48e+1	1.75e+2	4.48e+2	1.10e+3
50%	9.49e+1	1.72e+2	4.46e+2	1.22e+3	2.82e+3
95%	1.95e+2	3.46e+2	1.04e+3	3.37e+3	8.25e+3

Table 1: *Dispersion coefficient example (stability class C): Assessments of Decision Maker for σ_y .*

Dependencies

From PI the distribution on (A_y, B_y) is obtained. Based on this distribution different dependence measures between A_y and B_y can be determined. Recall that the target variables A_y and B_y were non observables and hence could not be elicitation variables. But what to do in case a set of target variables are observables? Since a joint distribution on the target variables of the model has to be specified before performing the uncertainty analysis, how shall we obtain the dependence between target variables which are observable. Chapter 2 starts off with an overview of available dependence elicitation techniques. Based on the considerations that the dependence elicitation technique should be easy understandable to the expert, and should fit in the Structured Expert Judgment Elicitation Methodology, a dependence elicitation technique is selected. It is stressed that the selected technique may not be the most rigorous way of eliciting dependence, but it is satisfactorily for reasons explained. Experts had little difficulty understanding the selected of dependence elicitation technique and acknowledged its importance. Next the usage of dependence information in combining expert assessments and in PI is investigated.

Calibration with uncertain observations

Table 1 lists quantile information of the DM for σ_y . These quantiles are obtained by combining the uncertainty distributions of the N_{exp} experts using equal weights, i.e. the weights w_j of Equation 1 are set to $w_j = \frac{1}{N_{\text{exp}}}$ for $j = 1, \dots, N_{\text{exp}}$. At Delft University of Technology procedures have been developed to assign weights to experts based on statistical performance on so-called seed variable questions; seed variable questions are questions in the experts' field of expertise for which the answer/experimental result is known to the project staff, but unknown to the expert. Project oversight authorities of the Joint CEC/USNRC project Uncertainty Analysis requested to investigate the effect of measurement variability in determining the experimental result on performance based weights. Different experimental procedures may have different strengths and weaknesses. Since the elicitation does not specify an experimental procedure, the values used to score the expert's performance might be contaminated by a measurement vari-

ability which the experts could not take into account. In Chapter 3 different methods are presented which take account of measurement variability.

Modeling Uncertainty

As mentioned earlier there are many variables which influence lateral plume spread, but power law model 2 is regarded to capture the uncertainty associated satisfactorily. The focus of Chapter 4 is to provide guidelines in developing/extracting models which capture the associated uncertainty. The use of mathematical models to capture uncertainty rather than to make predictions requires experts and decision makers to think about these models in a new and different ways. In particular a special class of compartmental models are studied and a relationship between acyclic compartmental models and influence diagrams is established.

Application of Results

As mentioned earlier, the aim of this thesis has been the development of new mathematical techniques which support the Structured Expert Judgment Elicitation Methodology. In performing an uncertainty analysis, these mathematical techniques have to be used in the reverse order as presented. Firstly, the model under consideration has to be investigated if it can be simplified using the considerations of Chapter 4. Based on the simplified model the target variables are identified, upon which the elicitation variables will be based. If data are scarce, experts are queried on the marginal distribution of the elicitation variables and the dependence among them. Based on these results, the expert's calibration is scored using the techniques developed in Chapter 3. Based on the expert's calibration scores the marginal distributions and dependence information are aggregated using the strategies of Chapter 2. If necessary, the probabilistic inversion techniques of Chapter 1 will be used on certain elicitation variables to obtain the distribution on target variables. The end result will be a joint distribution on all relevant model input parameters of the model. The final step of the uncertainty analysis will be the analysis of the model output and to investigate which uncertain model input parameters influence the model output uncertainty the most.

Chapter 1

Probabilistic Inversion

Many scientific applications involve the use of computer modeling to predict the course of a physical process, such as lateral plume spread. These models typically contain lots of parameters whose precise values are unknown. Probabilistic Inversion (PI) is used to establish a subjective probability distribution over the parameters of such models, and is needed when the model input parameters do not have a direct physical interpretation. The method works by taking subjective probability distributions on the observables and inverting them through the model. Mathematically, PI means that a probability distribution in the domain of a certain mapping is re-weighted using a Radon-Nikodym derivative so that its push-forward distribution on the image space of the mapping equals a given required distribution. This chapter start off with an investigation on the foundations of PI [33], which result in two implementation schemes [36],[38]. Crucial will be the determination of the domain on which the measure of the uncertain input parameters will be determined. To ensure the correctness of the determination of the domain an iterative version of an implementation has been developed which checks at each iteration whether or not the domain has to be extended. Finally, based on a theoretical result an efficient implementation is given which is able to deal with large problems.

1.1 Foundations

Let $(\mathbb{R}^m, \mathcal{B}, \lambda)$ and $(\mathbb{R}^n, \mathcal{B}, \nu)$ be two Borel probability spaces and $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ a continuous mapping. The mapping T will represent a physical model that gives an output representing an observable quantity when an input, specifying the parameters of the physical model, is given; in case of the dispersion coefficient example (see Introduction); if information on one downwind distance x is available we might choose $T(z) := \sigma_y(z) = A_y z^{B_y}$. Based on practical considerations, a compact subset $N \subset \mathbb{R}^n$ is identi-

fied as the set of *potential observations* in the image space. A compact set $M \subset \mathbb{R}^m$ ($\lambda(M) > 0$) is identified as the set of potential input parameter values and we say that M is *observationally complete* if N is in the image of M , $N \subseteq T(M)$. This means that all potentially observable outputs can be reproduced by some combination of model input parameter values. Here it is assumed that M is such that it is observationally complete.

The measure λ on M (normalized if necessary, so $\lambda(M) = 1$) will be termed the background measure and the variable μ will be used to denote another probability measure on M . Furthermore, the measure γ on N is defined as $\gamma := \lambda \circ T^{-1}$; that is, γ is the push-forward of λ . Finally, assume that the measure ν on N is given and $\nu(N) = 1$. Intuitively, the background measure plays a similar role as the prior distribution in Bayesian Analysis.

PI can be expressed mathematically as follows: a measure μ on M is sought such that

$$\forall B \in \mathcal{B}(N) \quad \mu \circ T^{-1}(B) = \nu(B) \quad (1.1)$$

where $\mu \circ T^{-1}(B) = \mu(T^{-1}(B))$ and $\mathcal{B}(N)$ is the restriction of the Borel σ -algebra to N .

Figure 1.1 illustrates the relation between the various objects.

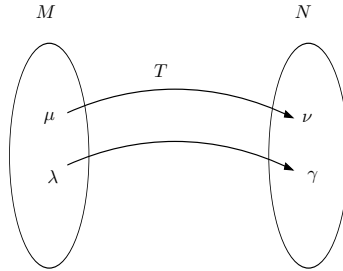


Figure 1.1: *Graphical illustration of relation between measures.*

The following examples illustrate that the uniqueness and existence of μ are not guaranteed.

Example 1: If μ exists it may not be unique. Suppose $T : \mathbb{R}^2 \rightarrow \mathbb{R}$, with $T(x_1, x_2) = x_1 - x_2$ and $\nu = \delta_0$, then any probability μ supported on $\{(x_1, x_2) | x_1 = x_2\}$ satisfies $\mu \circ T^{-1}(B) = \nu(B)$ for all Borel sets B . See Figure 1.2.

Example 2: The probability measure μ need not exist. Suppose $T : \mathbb{R} \rightarrow \mathbb{R}^2$, with $T(x) = (x, \frac{1}{2})$ and $N = [0, 1]^2$ with ν being uniform on N . For any μ on M , $\mu \circ T^{-1}$ is supported on $T(\mathbb{R}) = \{(x, \frac{1}{2}) | x \in \mathbb{R}\}$ which has ν -measure 0 in N . So Equation 1.1 cannot hold. Note that T is not surjective onto N . See Figure 1.3.

Example 3: Even if T is surjective onto N , the push-forward of an arbitrary measure μ need not be absolutely continuous w.r.t. ν . Suppose

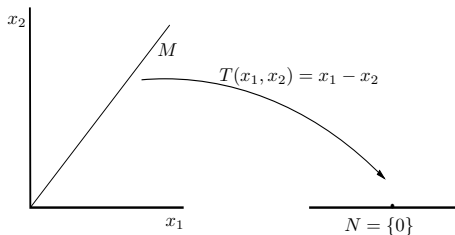


Figure 1.2: Graphical illustration of Example 1.

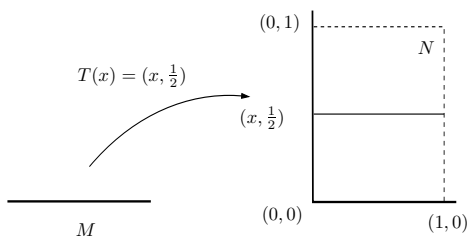


Figure 1.3: Graphical illustration of Example 2.

$T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, with $T(x, y) = (x, y)$. Let $M = N = [0, 1]^2$, then M is observationally complete. Let ν be uniform on the diagonal $D = \{(x, x) | x \in [0, 1]\}$ and μ be uniform on M . Then $\mu \circ T^{-1}$ is uniformly distributed on N , and $\mu \circ T^{-1}(D) = 0$ but $\nu(D) = 1$. See Figure 1.4

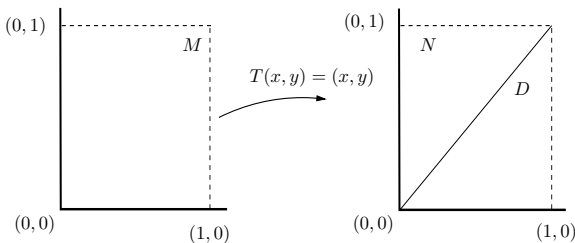


Figure 1.4: Graphical illustration of Example 3.

It can be shown that under certain conditions a probability measure μ can be found solving Equation 1.1. It is assumed that $\nu \ll \gamma$, which is stronger than assuming M is observationally complete. Under these conditions the Radon-Nikodym Theorem [3] (Theorem 32.2 p.443) permits us to assert the existence of a measurable non-negative function $g : N \rightarrow \mathbb{R}$, unique up to sets of γ -measure 0, such that

$$\nu(B) = \int_B g(y) d\gamma(y) \quad \forall B \in \mathcal{B}(N). \tag{1.2}$$

The function g is γ -integrable, since ν is finite, but it is assumed also that g is continuous, $g(y) > 0$ for $y \in N$ and $\int_N g(y) \log g(y) d\gamma(y) < \infty$. Define $\tilde{f} : M \rightarrow \mathbb{R}$ as $\tilde{f} := g \circ T$, and $\tilde{\mu} = \tilde{\mu}_{T, \nu, \gamma}$ as a new probability measure on $(M, \mathcal{B}(M))$ by $\tilde{f} = \frac{d\tilde{\mu}}{d\lambda}$, that is

$$\tilde{\mu}(A) := \int_A \tilde{f}(x) d\lambda(x) \quad \forall A \in \mathcal{B}(M) \quad (1.3)$$

so $\tilde{\mu} \ll \lambda$. Hence \tilde{f} is a continuous function unique up to sets of λ -measure 0, $\tilde{f}(x) = \tilde{f}(x')$ when $T(x) = T(x')$ and $\tilde{f}(x) > 0$ for $x \in T^{-1}(N)$. Since

$$\begin{aligned} \tilde{\mu} \circ T^{-1}(B) = \tilde{\mu}(T^{-1}(B)) &= \int_{T^{-1}(B)} \tilde{f}(x) d\lambda(x) = \int_{T^{-1}(B)} g \circ T(x) d\lambda(x) \\ &= \int_B g(y) d\gamma(y) \\ &= \nu(B) \end{aligned}$$

the induced Borel measure $\tilde{\mu} \circ T^{-1}$ equals ν . In particular $\tilde{\mu}$ is a probability measure satisfying Equation 1.1.

The measure $\tilde{\mu}$ is not necessarily the only measure satisfying Equation 1.1 (see Example 1). However, measure $\tilde{\mu}$ is the preferred one in the sense (as described in [40], p.4-5) that it uniquely minimizes the relative information¹ with respect to λ amongst all those measures satisfying Equation 1.1. The motivation for relative information is briefly recalled. Consider measures μ_i ($i = 1, 2$) satisfying Expression 1.1 and such that $\mu_i \ll \lambda$, and denote the Radon-Nikodym derivatives by f_i . As stated in [40], if H_i is the hypothesis that X is from the statistical population with measure μ_i , then it follows from Bayes theorem that

$$P(H_i|X = x) = \frac{P(H_i)f_i(x)}{P(H_1)f_1(x) + P(H_2)f_2(x)}[\lambda], \quad i = 1, 2 \quad (1.4)$$

from which it follows:

$$\log \frac{f_1(x)}{f_2(x)} = \log \frac{P(H_1|X = x)}{P(H_2|X = x)} - \log \frac{P(H_1)}{P(H_2)}[\lambda] \quad (1.5)$$

where $[\lambda]$ means that the assertion is unique up to sets of λ -measure 0. Furthermore, $P(H_i)$ ($i = 1, 2$) is the prior probability of H_i and $P(H_i|X = x)$ is the posterior probability of H_i , or the conditional probability of H_i given $X = x$. The right-hand side of Equation 1.5 can be viewed as a measure of the difference between the logarithm of the odds in favor of H_1 after observing $X = x$ and before the observation. This difference, which may be negative or positive, may be considered as the information resulting from the observation $X = x$. The relative information is defined

¹Relative information is referred to frequently as the Kullback-Leibler divergence.

as the logarithm of the likelihood ratio, $\log \frac{f_1(x)}{f_2(x)}$ when $X = x$, for the discrimination in favor of H_1 against H_2 . The mean relative information for discrimination in favor of H_1 against H_2 given $x \in A \in \mathcal{B}(M)$, for μ_1 , is

$$I(\mu_1|\mu_2) := \begin{cases} \frac{1}{\mu_1(A)} \int_A f_1(x) \log \frac{f_1(x)}{f_2(x)} d\lambda(x) & \text{for } \mu_1(A) > 0 \\ 0 & \text{for } \mu_1(A) = 0. \end{cases} \quad (1.6)$$

A necessary, but not sufficient, condition for $I(\mu_1|\mu_2)$ to be finite is if μ_1 and μ_2 are absolutely continuous with respect to one another, i.e. $\mu_1 \ll \mu_2$ and $\mu_2 \ll \mu_1$. Furthermore, relative information is non-negative, that is $I(\mu_1|\mu_2) \geq 0$, with equality if and only if $\mu_1 = \mu_2$, see [40] (Theorem 3.1 p. 14).

In general, the initial (background) measure λ does not have to satisfy Equation 1.1,

$$\forall B \in \mathcal{B}(N) \quad \lambda \circ T^{-1}(B) \neq \nu(B) \quad (1.7)$$

Probabilistic inversion is concerned with the determination of a measure $\tilde{\mu}$ on M , which satisfies Equation 1.1. Under the conditions we have imposed the existence of a solution to Equation 1.1 is guaranteed. A uniqueness property will be obtained by searching for a minimal relative informative measure satisfying Equation 1.1.

1.1.1 The measure $\tilde{\mu}$

Let $(X, \mathcal{S}, \alpha_1)$ and $(Y, \mathcal{T}, \beta_i)$, $i = 1, 2$ be probability spaces, where X is a topological space with \mathcal{S} the Borel σ -algebra, (Y, \mathcal{T}) is a Polish space with \mathcal{T} the Borel σ -algebra, and let S be a continuous mapping from X to Y . These probability spaces are introduced because Theorem 1.1.4 will be applied to various probability spaces. Suppose $\beta_1 = \alpha_1 \circ S^{-1}$ and assume $\beta_2 \ll \beta_1$, then a measurable non-negative function s_β exists, such that $\beta_2(B) = \int_B s_\beta(y) d\beta_1(y)$ for all $B \in \mathcal{T}$. Figure 1.5 illustrates the relation between the different objects.

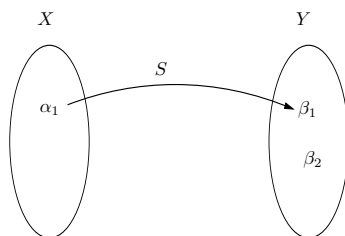


Figure 1.5: Graphical illustration of relation between α_1 , β_1 and β_2 .

Additionally it is assumed that s_β is continuous, $s_\beta(y) > 0$ for $y \in Y$ and $\int_Y s_\beta \log s_\beta d\beta_1(y) < \infty$. Define $\tilde{s} := s_\beta \circ S$ (\tilde{s} is continuous and $\tilde{s}(x) > 0$

for $x \in X$) and $\tilde{\alpha}(A) := \int_A \tilde{s}(x) d\alpha_1(x)$ for all $A \in \mathcal{S}$ so that $\tilde{\alpha}$ is absolutely continuous to α_1 , $\tilde{\alpha} \ll \alpha_1$.

Let $A_{\alpha_1, \beta_2, S} = \{\alpha | \alpha \ll \alpha_1, \alpha \circ S^{-1} = \beta_2\}$ be the set which contains measures α which are a solution to

$$\forall B \in \mathcal{T} \quad \alpha \circ S^{-1}(B) = \beta_2(B). \quad (1.8)$$

Lemma 1.1.1. *The set $A_{\alpha_1, \beta_2, S}$ is non-empty and convex.*

Proof: To show that $A_{\alpha_1, \beta_2, S}$ is non-empty, we prove that $\tilde{\alpha}$ is an element. Clearly $\tilde{\alpha} \ll \alpha_1$ and

$$\begin{aligned} \tilde{\alpha} \circ S^{-1}(B) = \tilde{\alpha}(S^{-1}(B)) &= \int_{S^{-1}(B)} \tilde{s}(x) d\alpha_1(x) = \int_{S^{-1}(B)} s_\beta \circ S(x) d\alpha(x) \\ &= \int_B s_\beta(y) d\beta_1(y) \\ &= \beta_2(B). \end{aligned}$$

In order to demonstrate convexity, let $\alpha', \alpha'' \in A_{\alpha_1, \beta_2, S}$ and $\delta \in [0, 1]$. Because $\alpha' \ll \alpha_1$ and $\alpha'' \ll \alpha_1$ it follows that $\delta\alpha' + (1 - \delta)\alpha'' \ll \alpha_1$. Let $B \in \mathcal{T}$,

$$\begin{aligned} (\delta\alpha' + (1 - \delta)\alpha'') \circ S^{-1}(B) &= (\delta\alpha' + (1 - \delta)\alpha'')(S^{-1}(B)) \\ &= \delta\alpha'(S^{-1}(B)) + (1 - \delta)\alpha''(S^{-1}(B)) \\ &= \delta\beta_2(B) + (1 - \delta)\beta_2(B) = \beta_2(B). \end{aligned}$$

□

In the remainder assume that $A_{\alpha_1, \beta_2, S}$ contains a measure with finite relative information with respect to α_1 . Presently, it is shown that $\tilde{\alpha}$ is the unique measure in $A_{\alpha_1, \beta_2, S}$ minimizing relative information with respect to α_1 , hence $I(\tilde{\alpha} | \alpha_1) < \infty$.

Let $\alpha \in A_{\alpha_1, \beta_2, S}$ and define $\alpha_{1,y}$ and α_y to be the conditional measures of α_1 and α , respectively, on X induced by conditioning on $S(x) = y$. By Theorem 10.2.2 (p. 270) from [20], these conditional measures exist for β_1 almost all y . Denote the Radon-Nikodym derivative of α with respect to α_1 by s . Lemma 1.1.3 will show that $\alpha_y \ll \alpha_{1,y}$ by explicitly giving a version of $\frac{d\alpha_y}{d\alpha_{1,y}}$. Theorem 10.2.1 taken from [20] (p. 269) will be used in Lemma 1.1.3:

10.2.1 Theorem *Let $(X \times Y, \mathcal{A}, P)$ be a probability space where (X, \mathcal{S}) and (Y, \mathcal{T}) are measurable spaces and \mathcal{A} is the product σ -algebra. Denote the points of $X \times Y$ by (x, y) . Suppose there exist conditional distributions P_y of x for $y \in Y$. Let $\beta := \int_X P(\bar{x}, y) d\bar{x}$ be the projection of P onto Y . Then for any integrable function g for P ,*

$$\int_{X \times Y} g dP = \int_X \int_Y g(x, y) dP_y(x) d\beta(y).$$

Let P be the push-forward measure of α onto D , and P_1 be the push-forward measure of α_1 onto D , where $D = \{(x, y) \in X \times Y | S(x) = y\}$. Note that for any measurable integrable function $g(x, y)$ and any measure α on X we have

$$\int g(x, y) d\alpha \circ S^{-1}(x, y) = \int g(x, S(x)) d\alpha(x). \quad (1.9)$$

See Figure 1.6 for a graphical illustration on how the measures, mapping and set D are connected.

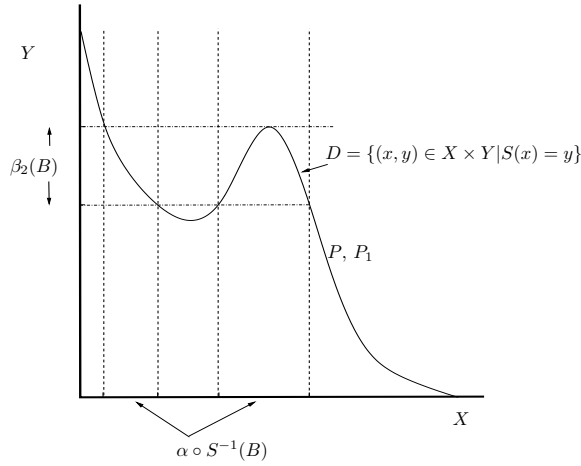


Figure 1.6: Illustration of measures, mapping S and set D .

Lemma 1.1.2. *The measure P is absolutely continuous with respect to P_1 , with Radon-Nikodym derivative*

$$\frac{dP}{dP_1}(x, y) = \begin{cases} s(x) & \text{if } y = S(x) \\ 0 & \text{otherwise.} \end{cases}$$

Proof: For any measurable integrable function g ,

$$\begin{aligned} \int_{A \times B} g(x, y) s(x) dP(x, y) &= \int_{(A \cap S^{-1}(B)) \times B} g(x, y) s(x) dP(x, y) \\ &= \int_{A \cap S^{-1}(B)} g(x, S(x)) s(x) d\alpha(x) \\ &= \int_{A \cap S^{-1}(B)} g(x, S(x)) d\alpha_1(x) \\ &= \int_{A \times B} g(x, y) dP_1(x, y). \end{aligned}$$

□

Lemma 1.1.3. *The conditional measure P_y is absolute continuous with respect to $P_{1,y}$, for β_2 almost all y , with Radon-Nikodym derivative*

$$\frac{dP_y}{dP_{1,y}}(x) = \begin{cases} \frac{s(x)}{s_\beta(S(x))} & \text{if } y = S(x) \\ 0 & \text{otherwise.} \end{cases}$$

Proof: We have

$$\begin{aligned} \int g(x, y) \frac{s(x)}{s_\beta(S(x))} dP_{1,y}(x) d\beta_2(y) &= \int g(x, y) s(x) dP_{1,y}(x) d\beta_1(y) \\ &= \int g(x, y) s(x) dP_1(x, y) \\ &= \int g(x, y) dP(x, y) \\ &= \int g(x, y) dP_y(x) d\beta_2(y) \end{aligned}$$

for any measurable integrable function g . □

Finally, remark that since P is the push-forward of α , and P_1 of α_1 , that $P_y = \alpha_y$ and $P_{1,y} = \alpha_{1,y}$.

Theorem 1.1.4. *For any $\alpha \in A_{\alpha_1, \beta_2, S}$,*

$$I(\alpha|\alpha_1) = \int_Y I(\alpha_y|\alpha_{1,y}) d\beta_2(y) + I(\beta_2|\beta_1)$$

Proof: Recall $s = \frac{d\alpha}{d\alpha_1}$, then $I(\alpha|\alpha_1)$ can be written as

$$\begin{aligned} I(\alpha|\alpha_1) &= \int_X \log s(x) d\alpha(x) \\ &= \int_Y \int_{S^{-1}(y)} \log s(x) d\alpha_y(x) d\beta_2(y). \end{aligned}$$

Using Lemma 1.1.3 we write

$$\begin{aligned} I(\alpha|\alpha_1) &= \int_Y \int_{S^{-1}(y)} \log \left(\frac{d\alpha_y}{d\alpha_{1,y}}(x) s_\beta(y) \right) d\alpha_y(x) d\beta_2(y) \\ &= \int_Y \int_{S^{-1}(y)} \log \frac{d\alpha_y}{d\alpha_{1,y}}(x) d\alpha_y(x) d\beta_2(y) + \int_Y \log s_\beta(y) d\beta_2(y) \\ &= \int_Y I(\alpha_y|\alpha_{1,y}) d\beta_2(y) + I(\beta_2|\beta_1). \end{aligned}$$

□

Corollary 1.1.5. *The probability measure $\tilde{\alpha}$ is the unique element of $A_{\alpha_1, \beta_2, S}$ minimizing relative information w.r.t. α_1 , $I(\tilde{\alpha}|\alpha_1) = I(\beta_2|\beta_1)$, and $I(\tilde{\alpha}|\alpha_1) < \infty$.*

Proof: For $\tilde{\alpha}$, by construction $\frac{d\tilde{\alpha}_y}{d\alpha_{1,y}}(x) = \frac{\tilde{s}(x)}{s_\beta(y)} = \frac{s_\beta \circ S(x)}{s_\beta(y)} = \frac{s_\beta(y)}{s_\beta(y)} = 1$ $[\alpha_1]$ almost surely, is constant. Hence $\tilde{\alpha}_y = \alpha_{1,y} [\beta_2]$ almost surely and $I(\tilde{\alpha}_y|\alpha_{1,y}) = 0$ $[\beta_2]$ almost surely. To show that $\tilde{\alpha}$ is unique, suppose that the measure $\alpha \in A_{\alpha_1, \beta_2, S}$ satisfies $I(\alpha|\alpha_1) = I(\beta_2|\beta_1)$. We then have $I(\alpha_y|\alpha_{1,y}) = 0$ $[\beta_2]$ almost surely. Hence $\alpha_y = \alpha_{1,y} [\beta_2]$ almost surely and so $\alpha_y = \tilde{\alpha}_y [\beta_2]$ almost surely, from which $\alpha = \tilde{\alpha}$ follows. In order to conclude that $I(\tilde{\alpha}|\alpha_1) < \infty$, we remark that $I(\beta_2|\beta_1) = \int_Y s_\beta \log s_\beta d\beta_1(y) < \infty$.

□

Specializing Theorem 1.1.4 and Corollary 1.1.5 to the setting of probabilistic inversion as introduced in the previous section gives the following two results.

Corollary 1.1.6. *For any $\mu \in A_{\lambda, \nu, T}$,*

$$I(\mu|\lambda) = \int_N I(\mu_y|\lambda_y) d\nu(y) + I(\nu|\gamma)$$

Corollary 1.1.7. *The probability measure $\tilde{\mu}$ is the unique element of $A_{\lambda, \nu, T}$ minimizing relative information w.r.t. λ , $I(\tilde{\mu}|\lambda) = I(\nu|\gamma)$, and $I(\tilde{\mu}|\lambda) < \infty$.*

In many applications, only partial information on the measure ν is available. For example, under the Structured Expert Judgment Elicitation Methodology [12], experts provide information on the measure ν by specifying a number of quantile points of their distribution. In case of the dispersion coefficient example, the experts provided the 5%, 50% and 95% quantile point. Section 1.1.2 discusses the implications of the availability of partial information ν_k for ν .

1.1.2 Approximating $\tilde{\mu}$ by $\tilde{\mu}_k$

Let \mathcal{B}_k ($k = 1, 2, \dots$) be a sequence of finite Borel partitions of N such that

$$\sigma(\cup_{k=1}^{\infty} \mathcal{B}_k) = \mathcal{B}(N)$$

and \mathcal{B}_{k+1} refines \mathcal{B}_k . For simplicity, it is assumed that the cardinality of \mathcal{B}_k is equal to k , and $\max_{B_i \in \mathcal{B}_k} \text{diam}_{B_i} B_i \rightarrow 0$. We write $\mathcal{B}_k = \{B_1, \dots, B_k\}$. Clearly $\mathcal{A}_k = T^{-1}(\mathcal{B}_k)$ is a finite Borel partition of M , $\mathcal{A}_k = \{A_1, \dots, A_k\}$ where $A_i = T^{-1}(B_i)$.

Define $\phi_k : N \rightarrow \{1, \dots, k\}$, by $\phi_k(y) := i$ when $y \in B_i$. Furthermore, define $T_k : M \rightarrow \{1, \dots, k\}$ as the composition $T_k(x) := \phi_k \circ T(x)$. Let $\gamma_k = \gamma \circ \phi_k^{-1}$ and $\nu_k = \nu \circ \phi_k^{-1}$ be the push-forwards of γ and ν , respectively, onto $\{1, \dots, k\}$.

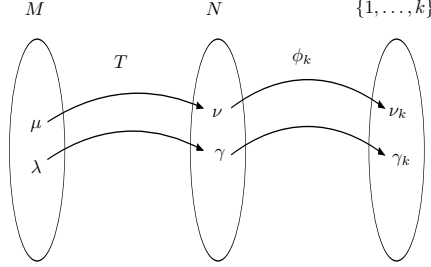


Figure 1.7: *Graphical illustration of extended relation between measures.*

See Figure 1.7 for a graphical illustration of how the different measures are connected.

This models the situation in which a limited number of probability ‘bins’ are quantified by expert opinion. In principle, this can be refined as much as required by doing more expert assessments. This corresponds to refining the partition \mathcal{B}_k .

Lemma 1.1.8. *The condition $\nu \ll \gamma$ implies that for all k ,*

$$\nu_k \ll \gamma_k.$$

Proof: Let $i \in \{1, \dots, k\}$ and suppose that $\gamma_k(i) = 0$. Then clearly $\gamma(B_i) = 0$, so that $\nu(B_i) = 0$, and $\nu_k(i) = 0$.

□

By the Radon-Nikodym theorem, the density $g_k : \{1, \dots, k\} \rightarrow \mathbb{R}$ given by $g_k = \frac{d\nu_k}{d\gamma_k}$ exists. Define $\tilde{f}_k : M \rightarrow \mathbb{R}$ as $\tilde{f}_k := g_k \circ T_k$ and $\tilde{\mu}_k(A) = \int_A \tilde{f}_k(x) d\lambda(x)$.

Corollary 1.1.9. *For any $\mu_k \in A_{\lambda, \nu_k, T_k}$,*

$$I(\mu_k | \lambda) = \sum_{i=1}^k I(\mu_k | \lambda_k | i) \nu_k(\{i\}) + I(\nu_k | \gamma_k)$$

Proof: Apply Theorem 1.1.4 with $\alpha = \mu_k$ and $S = T_k$ between $(M, \mathcal{B}(M), \lambda)$ and $(\{1, \dots, k\}, \mathcal{B}(\{1, \dots, k\}), \gamma_k, \nu_k)$.

□

Corollary 1.1.10. *The probability measure $\tilde{\mu}_k$ is the unique element of A_{λ, ν_k, T_k} minimizing relative information w.r.t. λ , $I(\tilde{\mu}_k | \lambda) = I(\nu_k | \gamma_k)$, and $I(\tilde{\mu}_k | \lambda) < \infty$.*

Since \mathcal{B}_{k+1} refines \mathcal{B}_k , $\sigma(\mathcal{B}_k) \subset \sigma(\mathcal{B}_{k+1})$ from which follows $\sigma(\mathcal{A}_k) \subset \sigma(\mathcal{A}_{k+1})$. Hence $\sigma(\mathcal{A}_k)$ is a sub σ -algebra of $\sigma(\mathcal{A}_{k+1})$. Our objective is to

show that the measures $\tilde{\mu}_k$ will converge to $\tilde{\mu}$ as $k \rightarrow \infty$. To show this it will be convenient to show that the densities \tilde{f}_k form a right-closable martingale.

Theorem 1.1.11. *The sequence $\{\tilde{f}_k, \sigma(\mathcal{A}_k)\}$ is a right-closable martingale.*

Proof: Recall the increasing sequence of partitions \mathcal{B}_k , and measurable partitions \mathcal{A}_k on M defined by $\mathcal{A}_k = T^{-1}(\mathcal{B}_k) = \{A_i | A_i = T^{-1}(B_i), B_i \in \mathcal{B}_k\}$. By construction \tilde{f}_k is constant on elements on \mathcal{A}_k . Hence \tilde{f}_k is $\sigma(\mathcal{A}_k)$ -measurable since $\sigma(\tilde{f}_k) \subset \sigma(\mathcal{A}_k)$. We claim that $\{\tilde{f}_k, \sigma(\mathcal{A}_k)\}$ is a right-closable martingale sequence.

Firstly, \tilde{f}_k and \tilde{f} are densities, hence \mathcal{L}^1 -functions and $\sigma(\mathcal{A}_k) \subset \sigma(\mathcal{A}_{k+1})$ for all k . Secondly for any $k, l \geq k$ and $A_i \in \mathcal{A}_k$

$$\begin{aligned} E_\lambda(\tilde{f}_l | A_i) &= \frac{1}{\lambda(A_i)} \int_{A_i} \tilde{f}_l(x) d\lambda(x) \\ &= \frac{\tilde{\mu}_l(A_i)}{\lambda(A_i)} \\ &= \frac{\nu(B_i)}{\lambda(A_i)} \\ &= \frac{\tilde{\mu}_k(A_i)}{\lambda(A_i)} \\ &= \tilde{f}_k(x) \quad \text{for any } x \in A_i \end{aligned}$$

This shows that $\tilde{f}_k = E_\lambda(\tilde{f}_l | \sigma(\mathcal{A}_k))$. Similarly $\tilde{f}_k = E_\lambda(\tilde{f} | \sigma(\mathcal{A}_k))$. Hence $\{\tilde{f}_k, \sigma(\mathcal{A}_k)\}$ is a right-closable martingale sequence, with extension $\tilde{f}_\infty = \tilde{f}$ and $\sigma(\mathcal{A}_\infty) = T^{-1}(\mathcal{B}(N))$.

□

Note that $\{f_k, \sigma(\mathcal{A}_k)\}$, with f_k not constant on elements of \mathcal{A}_k , is not a martingale sequence. By applying a martingale convergence theorem 10.5.1 (p.285) of [20], it follows that $f_k(x) \rightarrow f(x)$ $[\lambda]$ almost surely as $k \rightarrow \infty$. However a stronger convergence result is needed for Lemma 1.1.15 and Theorems 1.1.13 and 1.1.18, namely that $\frac{\tilde{f}(x)}{\tilde{f}_k(x)} \rightarrow 1$ $[\lambda]$ uniformly² as $k \rightarrow \infty$.

Proposition 1.1.12. $\lim_{k \rightarrow \infty} \frac{\tilde{f}(x)}{\tilde{f}_k(x)} = 1$ $[\lambda]$ uniformly.

Proof: N is compact, hence \tilde{f} is uniformly continuous on $T^{-1}(N)$. Furthermore, since $\tilde{f}(x) > 0$, there must exist an $L > 0$ such that $\tilde{f}(x) > \frac{1}{L}$ for all $x \in T^{-1}(N)$. Hence we also have $\tilde{f}_k = E(\tilde{f} | A_i) > \frac{1}{L}$. Additionally, since g is uniformly continuous on N , for any $\epsilon > 0$ there exists a $\delta > 0$ such that $|g(y_1) - g(y_2)| < \frac{\epsilon}{L}$ for all $y_1, y_2 \in N \subset \mathbb{R}^n$ satisfying $\|y_1 - y_2\|_{\max} < \delta$, where $\|y_1 - y_2\|_{\max} = \max_{j=1, \dots, n} \{|y_{1,j} - y_{2,j}|\}$.

² $\forall \epsilon > 0, \exists k_1, M_\epsilon$ such that $\lambda(M_\epsilon) = 1$ and $|\frac{\tilde{f}(x)}{\tilde{f}_k(x)} - 1| < \epsilon \forall x \in M_\epsilon$ and $k > k_1$.

Suppose $x \in A_i = T^{-1}(B_i)$ for some $B_i \in \mathcal{B}_k$, choose $\delta > 0$ such that $\max_{y^* \in B_i} |g(y) - g(y^*)| < \frac{\epsilon}{L}$ for all $y \in B_i$ then

$$\begin{aligned}
 \left| \frac{\tilde{f}(x)}{\tilde{f}_k(x)} - 1 \right| &\leq \left| \frac{1}{\tilde{f}_k(x)} \right| \left| \tilde{f}(x) - \tilde{f}_k(x) \right| & (1.10) \\
 &< L |\tilde{f}(x) - E_\lambda(\tilde{f}|A_i)| \\
 &\leq L \max_{x^* \in A_i} \left| \tilde{f}(x) - \tilde{f}(x^*) \right| \\
 &= L \max_{y^* \in B_i} |g(y) - g(y^*)| < \epsilon.
 \end{aligned}$$

□

Theorem 1.1.13. $\lim_{k \rightarrow \infty} \tilde{\mu}_k = \tilde{\mu}$ weakly.

Proof: It is sufficient to prove that $\int_M h(x) d\tilde{\mu}_k(x) \rightarrow \int_M h(x) d\tilde{\mu}(x)$ as $k \rightarrow \infty$ for any continuous bounded function $h : M \rightarrow \mathbb{R}$, or equivalently

$$\int_M h(x) \tilde{f}_k(x) d\lambda(x) \rightarrow \int_M h(x) \tilde{f}(x) d\lambda(x).$$

Since $h(x)$ is bounded, the above is implied by

$$\int_M |\tilde{f}_k(x) - \tilde{f}(x)| d\lambda(x) \rightarrow 0.$$

We have

$$\begin{aligned}
 \int_M \left| \tilde{f}_k(x) - \tilde{f}(x) \right| d\lambda(x) &\leq \int_M \left| \frac{\tilde{f}(x)}{\tilde{f}_k(x)} - 1 \right| d\lambda(x) \sup_{x \in M} \left| \tilde{f}_k(x) \right| \\
 &\leq \int_M \left| \frac{\tilde{f}(x)}{\tilde{f}_k(x)} - 1 \right| d\lambda(x) \sup_{x \in M} \left| \tilde{f}(x) \right| \rightarrow 0
 \end{aligned}$$

as $k \rightarrow \infty$ by Proposition 1.1.12.

□

Lemma 1.1.14. For every $k > 0$ we have $\tilde{\mu} \ll \tilde{\mu}_k$.

Proof: It is sufficient to prove that $\tilde{f}_k(x) = 0 \Rightarrow \tilde{f}(x) = 0$. By construction, if $\tilde{f}_k(x) = 0$ then $\tilde{f}_k(x^*) = 0$ for all $x^* \in A_i$. Hence $g_k(i) = g_k \circ T_k(x^*) = \tilde{f}_k(x^*) = 0$. But then $\nu_k(\{i\}) = 0$ which implies $\nu(B_i) = 0$. Then $g(y) = 0$ for all $y \in B_i$, from which is concluded that $\tilde{f}(x) = 0$.

□

Lemma 1.1.15. $\lim_{k \rightarrow \infty} I(\tilde{\mu}|\tilde{\mu}_k) = 0$.

Proof: Since $\lim_{k \rightarrow \infty} \frac{\tilde{f}(x)}{\tilde{f}_k(x)} = 1$ $[\lambda]$ uniformly, then for all $\epsilon > 0$, there exist k_1, M_ϵ such that $\lambda(M_\epsilon) = 1$ and $|\frac{\tilde{f}(x)}{\tilde{f}_k(x)} - 1| < \epsilon$ for all $x \in M_\epsilon$ and $k > k_1$. So

$$\begin{aligned} I(\mu|\mu_k) &= \int_{M_\epsilon} \tilde{f}(x) \log \frac{\tilde{f}(x)}{\tilde{f}_k(x)} d\lambda(x) \\ &\leq \int_{M_\epsilon} \tilde{f}(x) \log(1 + \epsilon) d\lambda(x) \quad k > k_1 \\ &= \log(1 + \epsilon). \end{aligned}$$

Hence $I(\mu|\mu_k) \rightarrow 0$ as $k \rightarrow \infty$.

□

The next lemma shows that the relative information increases when a partition is refined.

Lemma 1.1.16. For partition \mathcal{B}_k and \mathcal{B}_l , with $k \geq l$, $I(\nu_k|\gamma_k) \geq I(\nu_l|\gamma_l)$.

Proof: Recall $(\{1, \dots, i\}, \mathcal{B}(\{1, \dots, i\}), \gamma_i, \nu_i)$ for $i \in \{k, l\}$ are probability spaces and define $\pi_{k,l} : \{1, \dots, k\} \rightarrow \{1, \dots, l\}$ as the projection map such that $\pi_{k,l} \circ \phi_k = \phi_l$ be a projection from elements of \mathcal{A}_k to \mathcal{A}_l . Application of Theorem 1.1.4 between $(\{1, \dots, k\}, \mathcal{B}(\{1, \dots, k\}), \gamma_k)$ and $(\{1, \dots, l\}, \mathcal{B}(\{1, \dots, l\}), \gamma_l, \nu_l)$, with $\alpha = \nu_k$ and $S = \pi_{k,l}$. Since the first term of the right hand-side will consist of positive terms only, it follows that $I(\nu_k|\gamma_k) \geq I(\nu_l|\gamma_l)$.

□

Lemma 1.1.17. $I(\tilde{\mu}_k|\lambda) \leq I(\tilde{\mu}|\lambda) < \infty$.

Proof: By Corollary 1.1.10, $I(\tilde{\mu}_k|\lambda) = I(\nu_k|\gamma_k)$ and Lemma 1.1.16 states $I(\nu_k|\gamma_k) \geq I(\nu_l|\gamma_l)$ for $k \geq l$, hence $I(\tilde{\mu}_k|\lambda)$ is an increasing function in k , which is bounded above by $I(\mu|\lambda)$ since application of Theorem 1.1.4 between $(M, \mathcal{B}(M), \lambda)$ and $(\{1, \dots, k\}, \mathcal{B}(\{1, \dots, k\}), \gamma_k, \nu_k)$ with $\alpha = \tilde{\mu}$ and $S = T_k$ results in $I(\tilde{\mu}|\lambda) \geq I(\nu_k|\gamma_k) = I(\tilde{\mu}_k|\lambda)$. In particular, it is assumed that $A_{\lambda, \nu, T}$ contains a measure with finite relative information with respect to λ , $I(\tilde{\mu}|\lambda) < \infty$, hence $I(\tilde{\mu}_k|\lambda)$ is bounded.

□

Theorem 1.1.18. $\lim_{k \rightarrow \infty} I(\tilde{\mu}_k|\lambda) = I(\tilde{\mu}|\lambda)$.

Proof: Write

$$\begin{aligned} I(\tilde{\mu}|\lambda) &= \int_M \tilde{f}(x) \log \tilde{f}(x) d\lambda(x) \\ &= \int_M \tilde{f}(x) \log \frac{\tilde{f}(x)}{\tilde{f}_k(x)} d\lambda(x) + \int_M \tilde{f}(x) \log \tilde{f}_k(x) d\lambda(x). \end{aligned}$$

Consider

$$\begin{aligned} I(\tilde{\mu}|\lambda) - I(\tilde{\mu}_k|\lambda) &= \int_M \tilde{f}(x) \log \frac{\tilde{f}(x)}{\tilde{f}_k(x)} d\lambda(x) + \\ &\quad + \int_M (\tilde{f}(x) - \tilde{f}_k(x)) \log \tilde{f}_k(x) d\lambda(x). \end{aligned}$$

Taking the absolute value of the equation above and applying the triangle inequality,

$$\begin{aligned} |I(\tilde{\mu}|\lambda) - I(\tilde{\mu}_k|\lambda)| &\leq \left| \int_M \frac{\tilde{f}(x)}{\tilde{f}_k(x)} \log \frac{\tilde{f}(x)}{\tilde{f}_k(x)} \tilde{f}_k(x) d\lambda(x) \right| + \\ &\quad + \left| \int_M (\tilde{f}(x) - \tilde{f}_k(x)) \log \tilde{f}_k(x) d\lambda(x) \right|. \end{aligned}$$

Since $\frac{\tilde{f}(x)}{\tilde{f}_k(x)} \rightarrow 1$ [λ] uniformly, it follows for sufficiently large k , there exists an $\epsilon > 0$ such that $|\tilde{f}(x) - \tilde{f}_k(x)| < \epsilon |\tilde{f}_k(x)|$

$$\begin{aligned} |I(\tilde{\mu}|\lambda) - I(\tilde{\mu}_k|\lambda)| &\leq \left| \int_M \frac{d\tilde{\mu}}{d\tilde{\mu}_k} \log \frac{d\tilde{\mu}}{d\tilde{\mu}_k} d\tilde{\mu}_k \right| + \epsilon \left| \int_M \tilde{f}_k(x) \log \tilde{f}_k(x) d\lambda(x) \right| \\ &= I(\tilde{\mu}|\tilde{\mu}_k) + \epsilon I(\tilde{\mu}_k|\lambda) \\ &\leq I(\tilde{u}|\tilde{u}_k) + \epsilon I(\tilde{u}|\lambda) \end{aligned} \tag{1.11}$$

by Lemma 1.1.17. It follows from Lemma 1.1.15 that $I(\tilde{\mu}|\tilde{\mu}_k) \rightarrow 0$ as $k \rightarrow \infty$.

□

In the implementation of the probabilistic inversion scheme, discrete probability approximations will be considered. Section 1.1.3 explores the implications of approximating $\tilde{\mu}_k$ by $\tilde{\mu}'_{n',k}$, where n' denotes the total number of samples.

1.1.3 Approximating $\tilde{\mu}_k$ by $\tilde{\mu}'_{n',k}$

Assume that $\{x_1, \dots, x_{n'}\} \subset M$ is an i.i.d. sample of size n' from the background measure λ . Let $\lambda_{n'}$ be the empirical background measure of the sample $\{x_1, \dots, x_{n'}\}$. Recall $\gamma_k = \lambda \circ T_k^{-1}$, we now define $\gamma_{n',k} := \gamma_{n'} \circ \phi_k^{-1}$ as the sample push-forward background measure of $\lambda_{n'}$.

Lemma 1.1.19. *For fixed k , $\lim_{n' \rightarrow \infty} \gamma_{n',k} = \gamma_k$ $[\lambda]$ almost surely.*

Proof: Since \mathcal{A}_k is a finite measurable partition of M , the strong law of large numbers says that, for $[\lambda]$ almost surely, the sequence X_1, X_2, \dots

$$\lambda_{n'}(A_i) = \frac{\#\{j | X_j \in A_i, 1 \leq j \leq n'\}}{n'} \rightarrow \lambda(A_i) \quad (1.12)$$

where $i = 1, \dots, k$. Hence $\lambda_{n'}(A_i) \rightarrow \lambda(A_i)$ for $i = 1, \dots, k$ almost surely, or equivalently $\gamma_{n',k}(\{i\}) \rightarrow \gamma_k(\{i\})$ $[\lambda]$ almost surely. \square

Since $\{1, \dots, k\}$ is discrete, a sample of size n' may be taken for which $\gamma_{n',k} \equiv \gamma_k$, i.e. for large n' : $\gamma_{n',k} \ll \gamma_k$ and $\gamma_k \ll \gamma_{n',k}$. Let n' be large enough that $\gamma_{n',k} \equiv \gamma_k$, whence by Lemma 1.1.8, $\nu_k \ll \gamma_{n',k}$. Then $g_{n',k} = \frac{d\nu_k}{d\gamma_{n',k}}$ exists by the Radon-Nikodym Theorem and define $\tilde{f}_{n',k} := g_{n',k} \circ T_k$. Finally, define $\tilde{\mu}_{n',k}(A) := \int_A \tilde{f}_{n',k}(x) d\lambda_{n'}(x)$ for $A \subset M$.

Lemma 1.1.20. *For fixed k and $i \in \{1, \dots, k\}$, $\lim_{n' \rightarrow \infty} \frac{g_{n',k}(i)}{g_k(i)} = 1$ $[\lambda]$ almost surely.*

Proof: Let k be fixed and choose $i \in \{1, \dots, k\}$. For all large n' ,

$$\frac{g_{n',k}(i)}{g_k(i)} = \frac{d\nu_k(i)}{d\gamma_{n',k}(i)} \frac{d\gamma_k(i)}{d\nu_k(i)} = \frac{d\gamma_k(i)}{d\gamma_{n',k}(i)} = \frac{\gamma_k(\{i\})}{\gamma_{n',k}(\{i\})}.$$

From Lemma 1.1.19 it follows that $\frac{g_{n',k}(i)}{g_k(i)} \rightarrow 1$ $[\lambda]$ almost surely as $n' \rightarrow \infty$. \square

Theorem 1.1.21. *For fixed k , $\lim_{n' \rightarrow \infty} \tilde{\mu}_{n',k} = \tilde{\mu}_k$ weakly $[\lambda]$ almost surely.*

Proof: Recall $\tilde{\mu}_{n',k}(A) = \int_A \tilde{f}_{n',k}(x) d\lambda_{n'}(x)$,

$$\begin{aligned} \tilde{\mu}_{n',k}(A) &= \sum_{i=1}^k \tilde{\mu}_{n',k}(A \cap A_i) \quad A_i \in \mathcal{A}_k \\ &= \sum_{i=1}^k \int_{A \cap A_i} \tilde{f}_{n',k}(x) d\lambda_{n'}(x). \end{aligned}$$

Note that $\tilde{f}_{n',k}$ is constant on $A \cap A_i$. For each i choose $x_i \in A_i \cap A$ then

$$\begin{aligned} \tilde{\mu}_{n',k}(A) &= \sum_{i=1}^k \tilde{f}_{n',k}(x_i) \lambda_{n'}(A \cap A_i) \\ &= \sum_{i=1}^k g_{n',k}(T_k(x_i)) \lambda_{n'}(A \cap A_i). \end{aligned}$$

The Strong Law of Large Numbers implies that as $n' \rightarrow \infty$ then $\lambda_{n'} \rightarrow \lambda$ almost surely. Application of Lemma 1.1.20,

$$\begin{aligned} \tilde{\mu}_{n',k}(A) &\rightarrow \sum_{i=1}^k g_k(T_k(x_i))\lambda(A \cap A_i) \\ &= \sum_{i=1}^k \tilde{f}_k(x_i)\lambda(A \cap A_i) \\ &= \tilde{\mu}_k(A). \end{aligned}$$

□

Theorem 1.1.22. $\lim_{k \rightarrow \infty} \lim_{n' \rightarrow \infty} \tilde{\mu}_{n',k} = \tilde{\mu}$ weakly, $[\lambda]$ almost surely.

Proof: Follows from applying Theorem 1.1.21 first, followed by applying Theorem 1.1.13.

□

Proposition 1.1.23. For fixed k , $\lim_{n' \rightarrow \infty} I(\tilde{\mu}_{n',k}|\lambda_{n'}) = I(\tilde{\mu}_k|\lambda)$.

Proof: Let $A_i \in \mathcal{A}_k$ and $x_i \in A_i$, then

$$\begin{aligned} I(\tilde{\mu}_{n',k}|\lambda_{n'}) &= \sum_{i=1}^k \tilde{f}_{n',k}(x_i) \log \tilde{f}_{n',k}(x_i) \lambda_{n'}(A_i) \\ &= \sum_{i=1}^k g_{n',k}(T_k(x_i)) \log g_{n',k}(T_k(x_i)) \lambda_{n'}(A_i). \end{aligned}$$

The Strong Law of Large Numbers implies that as $n' \rightarrow \infty$ then $\lambda_{n'} \rightarrow \lambda$ almost surely. Application of Lemma 1.1.20,

$$\begin{aligned} I(\tilde{\mu}_{n',k}|\lambda_{n'}) &\rightarrow \sum_{i=1}^k g_k(T_k(x_i)) \log g_k(T_k(x_i)) \lambda(A_i) \\ &= \sum_{i=1}^k \tilde{f}_k(x_i) \log \tilde{f}_k(x_i) \lambda(A_i) \\ &= I(\tilde{\mu}_k|\lambda). \end{aligned}$$

□

Theorem 1.1.24. $\lim_{k \rightarrow \infty} \lim_{n' \rightarrow \infty} I(\tilde{\mu}_{n',k}|\lambda_{n'}) = I(\tilde{\mu}|\lambda)$.

Proof: First apply Proposition 1.1.23, followed by Theorem 1.1.18.

□

Theorem 1.1.24 says that the sequence thus obtained also approximates $\tilde{\mu}$ in terms of relative information.

1.1.4 Example: revisited

In the dispersion coefficient example, $M \subset \mathbb{R}^2$ and $N \subset \mathbb{R}^5$; the target variables \mathbf{X} are (A_y, B_y) and $(\sigma_y(z_1), \sigma_y(z_2), \dots, \sigma_y(z_n))$ are the elicitation variables \mathbf{Y} . Quantile assessments for the measure ν on N are given in Table 1. Let $I_{j,l}$ denote the l -th interquantile interval of the lateral plume spread assessments for the j -th elicitation variable ($l = 1, \dots, 4$, $j = 1, \dots, 5$), for example:

$$\begin{aligned} I_{1,1} &= [0, 33] \\ I_{4,2} &= (448, 1220] \end{aligned}$$

Taking the product of all such interquantile intervals, the partition elements ('observable hypercubes') B_i are defined as,

$$B_i := I_{1,l_1} \times I_{2,l_2} \times \dots \times I_{5,l_5} \quad B_i \in \mathcal{B}_k$$

where $j \in \{1, \dots, 5\}$, $l_j \in \{1, \dots, 4\}$. The number of partition elements equals $k = 4^5$. The mapping ϕ_k of Section 1.1.2, links the interquantile intervals with the partition element k and is defined here as

$$\phi_k(y_1, \dots, y_5) := \sum_{j=1}^5 \sum_{l=1}^4 l 4^{j-1} 1_{I_{j,l}}(y_j)$$

and thus associates with each partition element a number in base 4. Note that experts provided information in terms of information on the interquantile intervals ($\nu(I_{j,l_j})$) rather than on the partition elements ($\nu(B_i)$); in this example, an interquantile interval is the union of 4 partition elements. By asking experts about other quantiles, a sequence of partitions is obtained as in Section 1.1.2.

1.2 Probabilistic Inversion Implementations

In the implementation of Probabilistic Inversion, the discrete approximation $\tilde{\mu}_{n',k}$ is determined in two steps. Firstly, based on a Monte-Carlo heuristic a sample set in the target variable space is generated, which is considered to approximate the i.i.d. sample of size n' from background measure λ . The resulting sample set is identified as domain M upon which in Step 2 the measure $\tilde{\mu}_k$ will be determined. The target variables (uncertain input parameters) are represented by $\mathbf{X} = (X_1, \dots, X_m)$ with realizations $\mathbf{x} = (x_1, \dots, x_m)$ and the elicitation variables (observables) are represented by $\mathbf{Y} = (Y_1, \dots, Y_n)$ with realizations $\mathbf{y} = (y_1, \dots, y_n)$. The two steps of the implementation are illustrated graphically using the diffusion coefficient example for 2 observables³. For notational convenience, the measures $\tilde{\mu}_{n',k}$, $\lambda_{n'}$ will be denoted by $\tilde{\mu}_k$, λ , respectively.

³It must be stressed that the figures illustrating the different steps are used only for the purpose of illustration and are not based on actual data.

Two different implementations will be presented; PARFUM and PREJUDICE. Each implementation has advantages and disadvantages which will be discussed in Section 1.2.3.

1.2.1 PARFUM

The acronym PARFUM stands for PARAMeter Fitting for Uncertain Models and is described in [10], [36]. The current computer implementation of PARFUM is suitable only for linear models, or models that become linear when transformed, whereas the implementation of PARFUM can be applied to non-linear models as well.

Step 1 Determination of M : for each elicitation variable Y_j ($j = 1, \dots, n$) the k^* -th quantile point is chosen ($k^* = 1, \dots, K$),

$$\mathbf{y} = (y_{1,k^*}, \dots, y_{n,k^*}) \quad (1.13)$$

where $y_{1,1}$ would represent the 5% quantile point of Y_1 and K the total number of quantile points elicited. Next the vector \mathbf{y} is checked if it is potentially observable. If \mathbf{y} is potentially observable, it will be called a *scenario* and it will be added to the set of potentially observable scenarios N .

For each $\mathbf{y} \in N$, define $\mathbf{x}_{\mathbf{y}} := (x_{\mathbf{y},1}, \dots, x_{\mathbf{y},m})$ as the model inversions which minimize,

$$\min \sum_{j=1}^n (y_{j,k^*} - T_{(j)}(\mathbf{x}_{\mathbf{y}}))^2 \quad (1.14)$$

where $T_{(j)} : \mathbb{R}^m \rightarrow \mathbb{R}$ is the mapping into j -th dimension of the observable space⁴; for example, in case of the dispersion coefficient example $T(2) = A_y z_2^{B_y}$.

For each target variable X_i ($i = 1, \dots, m$) the minimum and maximum value, represented by \min_i and \max_i respectively, are determined. The interval $[\min_i, \max_i]$ is extended with a certain percentage of its range on both sides to determine the interval M_i on which the distribution for X_i will be specified. Next M_i is discretized by a number of equally spaced points. The domain M will be the Cartesian product of the discretized versions of M_i ;

$$M = M_1 \times \dots \times M_m \quad (1.15)$$

The domain M is propagated through $T_{(j)}$ to obtain n observable spaces, see Figure 1.8; in case of the dispersion coefficient example

⁴The mapping $T_{(j)}$ should not be mistaken by the mapping T_k , since $T_{(j)} : \mathbb{R}^m \rightarrow \mathbb{R}$ and $T_k : \mathbb{R}^m \rightarrow \{1, \dots, k\}$ (see Section 1.1.2).

the realizations of (A_y, B_y) are represented as (a_y, b_y) and the ‘dots’ in the target variable space represent target variable realizations and the ‘dots’ above the lines in the observable space represent the push-forward outcomes of the target variable realizations.

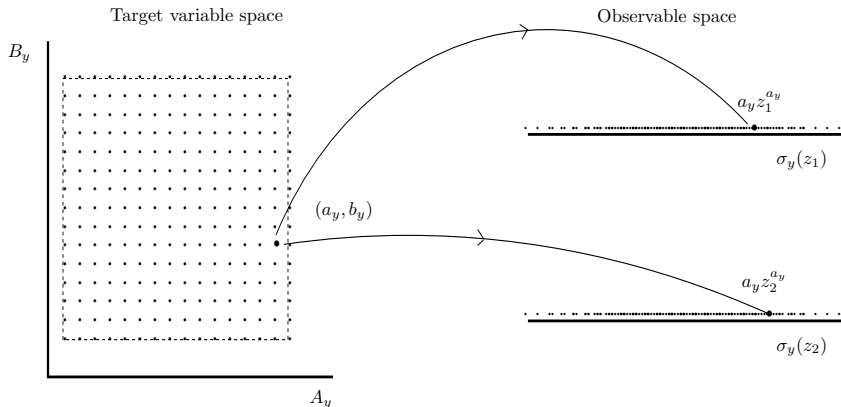


Figure 1.8: *PARFUM: Propagation of samples.*

Step 2 Determination of distribution: For each elicitation variable Y_j , the conditional probability mass function $f_{k|j}$ of conditional measure $\mu_{k|j}$ for $\mathbf{x} \in M$ is defined as

$$f_{k|j}(\mathbf{x}|T(\mathbf{x}) \in I_{j,l_j}) := \frac{\nu(I_{j,l_j})}{\sum_{\mathbf{x} \in M} 1_{T(j)}(\mathbf{x}) \in I_{j,l_j}(\mathbf{x})} \quad (1.16)$$

where $l_j = 1, \dots, K + 1$ and $\nu(I_{j,l_j})$ has to be interpreted as the probability of the l_j -th interquantile interval of elicitation variable Y_j under measure ν ; intuitively the push-forward outcomes receive probabilities depending on the probability associated with interquantile interval I_{j,k_j} and the number of push-forward samples contained in I_{j,k_j} . These probabilities are ‘pulled-back’ onto the corresponding target variable realizations.

To summarize, for each elicitation variable Y_j , a measure $\mu_{k|j}$ on M is obtained. See Figure 1.9.

Next, the measure $\tilde{\mu}_k$ on M has to be determined, which *best fits* the measures $\mu_{k|1}, \dots, \mu_{k|n}$. Relative information arguments are used to solve the problem ([40], for a discussion in this context see [10]): find the measure $\tilde{\mu}_k$ on M for which $\sum_{j=1}^n I(\mu_{k|j}|\tilde{\mu}_k)$ is minimal, where

$$I(\mu_{k|j}|\tilde{\mu}_k) = \sum_{\mathbf{x} \in M} f_{k|j}(\mathbf{x}) \ln \frac{f_{k|j}(\mathbf{x})}{\tilde{f}_k(\mathbf{x})}. \quad (1.17)$$

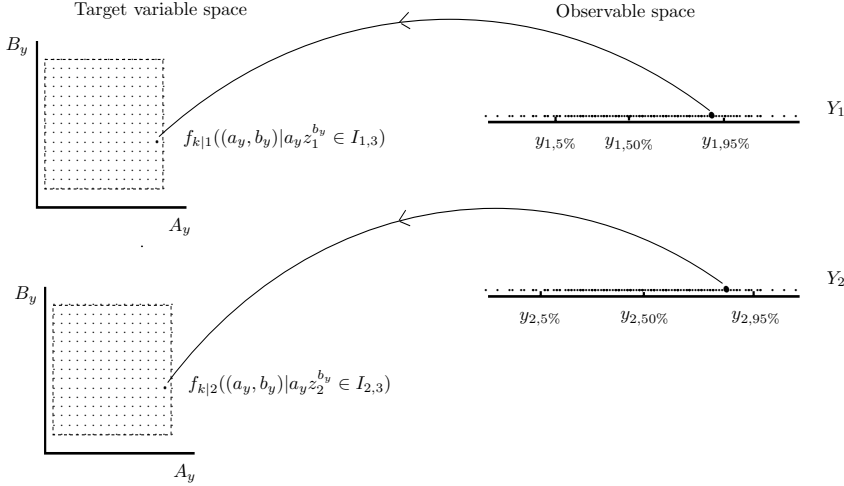


Figure 1.9: *PARFUM: Determination of distribution.*

The $\tilde{\mu}_k$ solving the minimization problem is just the average of the measures $\mu_{k|j}$, $j = 1, \dots, n$. Other solution concepts could be used leading to other solutions, for details see [10].

1.2.2 PREJUDICE

The solution scheme implemented in PREJUDICE is based on an idea proposed by Hora and Young, [27]. The acronym PREJUDICE stands for PRocessing Expert JUDgment Into Code paramEters.

Step 1 Determination of M : all possible combinations of quantile points y_{j,k_j} among the elicitation variable Y_j ($j = 1, \dots, n$) will be generated:

$$\mathbf{y} = (y_{1,k_1}, \dots, y_{n,k_n}) \quad (1.18)$$

where $k_j \in \{1, \dots, K\}$ and K the total number of quantile points elicited. Next each combination \mathbf{y} is checked if it is potentially observable. If the combination is potentially observable, it will be called a scenario and it will be added to the set of potentially observable scenarios N .

For each $\mathbf{y} \in N$ define $\mathbf{x}_{\mathbf{y}} := (x_{\mathbf{y},1}, \dots, x_{\mathbf{y},m})$ as the model inversions which minimize,

$$\min \sum_{j=1}^n (y_{j,k_j} - T_{(j)}(\mathbf{x}_{\mathbf{y}}))^2 \quad (1.19)$$

Expression 1.19 is evaluated for each scenario in N .

The main difference between PARFUM and PREJUDICE is the dimension of the observable space; PARFUM considers an observable space for each elicitation variable, PREJUDICE considers one observable space which consists of all elicitation variables. Another difference between PREJUDICE and PARFUM concerns the construction of domain M . In case the number of target variables is large, the construction of M as done in PARFUM leads to combinatorial problems. PREJUDICE allows the user to choose from the following schemes: to illustrate the sampling scheme it is assumed that the set of model inversions obtained from minimizing Expression 1.19 are the same for both sampling schemes.

ϵ -neighborhood : For $i = 1, \dots, m$, choose $\epsilon_i > 0$ and for each model inversion $\mathbf{x}_y, \mathbf{y} \in N$, sample randomly from $M_{\mathbf{x}_y}$ defined as:

$$M_{\mathbf{x}_y} := M_{x_{y,1}} \times \dots \times M_{x_{y,m}} \quad (1.20)$$

where $M_{x_{y,i}} := [(1 - \epsilon_i) x_{y,i}, (1 + \epsilon_i) x_{y,i}]$. The samples are propagated through T , see Figure 1.10.

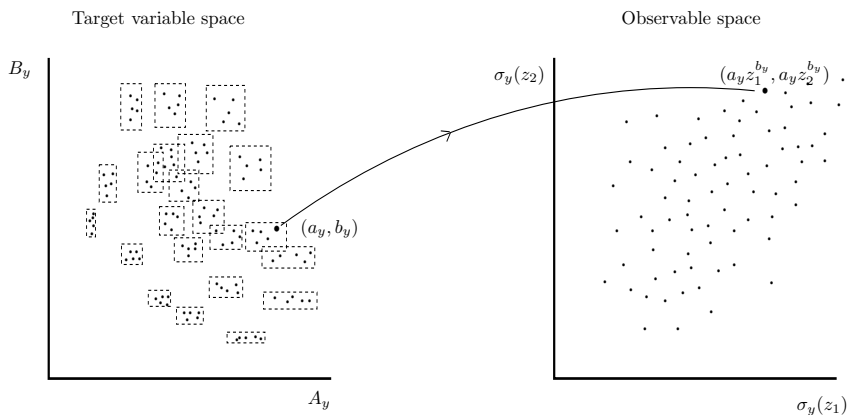


Figure 1.10: *PREJUDICE: Propagation of samples (ϵ -neighborhood).*

Bin combinations : for each target variable X_i ($i = 1, \dots, m$) the minimum and maximum value are represented by \min_i and \max_i respectively. The interval $[\min_i, \max_i]$ is extended with a certain percentage of its range on both sides to determine the interval M_i on which the distribution for X_i will be specified. For each target variable X_i , bins $B_{i,p}$ ($p = 1, \dots, C_i$) are constructed such that

$$\cup_{p=1}^{C_i} B_{j,p} = [\min M_i, \max M_i] \quad (1.21)$$

and $B_{i,p} \cap B_{i,q} = \emptyset$ for $p, q = 1, \dots, C_i$, $p \neq q$. For each target variable, the number of bins C_i depends on the spread of the model inversions \mathbf{x}_y .

For each \mathbf{x}_y , it is determined in which bin B_{i,p_i} ($p_i \in \{1, \dots, C_i\}$) the model inversion $x_{y,i}$ is positioned ($i = 1, \dots, m$). The sample region $M_{\mathbf{x}_y}$ is defined as the Cartesian product of the bins B_{i,p_i} , in formula

$$M_{\mathbf{x}_y} := B_{1,p_1} \times \dots \times B_{m,p_m} \quad (1.22)$$

Next, samples are drawn randomly from $M_{\mathbf{x}_y}$ and propagated through T . See Figure 1.11.

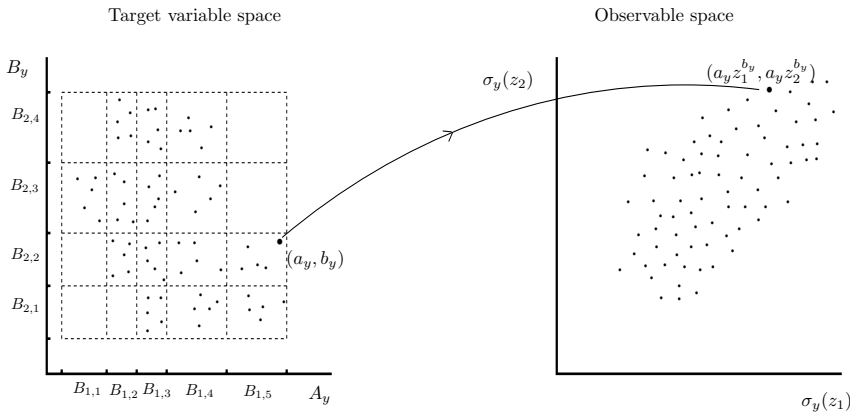


Figure 1.11: *PREJUDICE: Propagation of samples (bin combinations).*

Appendix A illustrates the difference between the ϵ -neighborhood and bin-combinations sampling schemes for the dispersion coefficient example. The crucial step in the probabilistic inversion implementation is the determination of M , which depends heavily on the minimization of Expression 1.19; Appendix B investigates the influence of various global optimization routines, different characteristics of T and choice of starting points for 3 cases taken from the Joint CEC/USNRC Uncertainty Analysis.

Compared to Section 1.1, the sample set is actually the support of λ , but by the results of Section 1.1 we can approximate M by this set. For simplicity the sample set of size n' will be called M .

Step 2 Determination of distribution: briefly, a measure $\tilde{\mu}_k$ (with probability mass function $\tilde{f}_k \in \mathbb{R}^{n'}$) on domain M is determined, which has minimum relative information with respect to a background measure λ (with probability mass function $f_\lambda \in \mathbb{R}^{n'}$) under the condition

that its push-forward measure complies with the quantile information available on the elicitation variables Y_j , i.e.

$$\sum_{\mathbf{x} \in M} \tilde{f}_k(\mathbf{x}) 1_{\{\mathbf{x}|T_{(j)}(\mathbf{x}) \in I_{j,l_j}\}}(\mathbf{x}) = \nu(I_{j,l_j}) \quad j = 1, \dots, n. \quad (1.23)$$

The probability mass function \tilde{f}_k can be looked at as the solution of the Convex Programming (CP) Problem:

$$\begin{aligned} \min_{f_k} \quad & \sum_{\mathbf{x} \in M} f_k(\mathbf{x}) \log \frac{f_k(\mathbf{x})}{f_\lambda(\mathbf{x})} \\ \text{s.t.} \quad & A_M f_k = \mathbf{b} \\ & f_k(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in M \end{aligned} \quad (1.24)$$

where $f_k \in \mathbb{R}^{n'}$ and $A_M \in \mathbb{R}^{n \cdot (K+1) \times n'}$ contains the information based on the indicator function of Expression 1.23, in case n' samples are drawn from M , and $\mathbf{b} \in \mathbb{R}^{n \cdot (K+1)}$ is defined as

$$\mathbf{b} := (\nu(I_{1,1}), \dots, \nu(I_{1,K+1}), \dots, \nu(I_{n,1}), \dots, \nu(I_{n,K+1}))^T.$$

Note that $f_k \in \mathbb{R}^{n'}$, hence the dimension of f_k depends on the samples drawn from M .

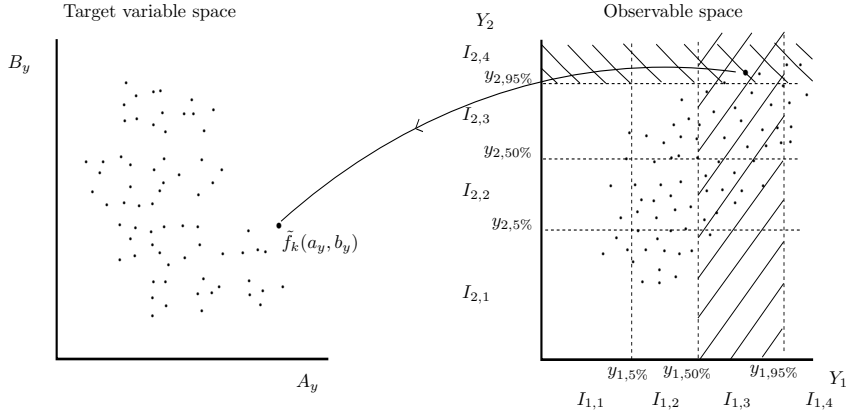


Figure 1.12: *PREJUDICE*: determination of probability mass function \tilde{f}_k .

Looking at Figure 1.12, the sample (a_y, b_y) from the target variable space maps into $(a_y z_1^{b_y}, a_y z_2^{b_y})$ in the observable space, which is contained in partition element $B_{19} = I_{1,3} \times I_{2,4}$. Hence, the assignment of $\tilde{f}_k(a_y, b_y)$ is such that the probability of the push-forward samples contained in $I_{1,3}$ adds up to 0.45 (the // -area in Figure 1.12) and be such that the probability of the push-forward samples contained in $I_{2,4}$ adds up to 0.05 (the \\\ -area in Figure 1.12).

Let PI problem (1.24) be the primal formulation of the PI problem. The dual formulation of PI problem (1.24) can be obtained by looking at the Lagrangian, and results is an unconstrained CP problem:

$$\max_{\mathbf{y}_M} \mathbf{y}_M^T \mathbf{b} - f_\lambda^T \exp(A_M^T \mathbf{y}_M - \mathbf{e}) \quad (1.25)$$

where $\mathbf{y}_M \in \mathbb{R}^{n \cdot (K+1)}$ is the dual variable corresponding to the dual formulation of the PI problem (1.24) and $\mathbf{e} \in \mathbb{R}^{n'}$ the unit vector. Hence, in this thesis a PI problem is always a CP problem.

Lemma 1.2.1 and Corollary 1.2.2 list some well-known duality properties of PI problem (1.24), see [44].

Lemma 1.2.1 (Weak duality). *If f_k is feasible for PI problem (1.24) and \mathbf{y} is feasible for PI problem (1.25), then*

$$\sum_{\mathbf{x} \in M} f_k(\mathbf{x}) \log \frac{f_k(\mathbf{x})}{f_\lambda(\mathbf{x})} \geq \mathbf{y}^T \mathbf{b} - \sum_{\mathbf{x} \in M} f_\lambda^T \exp(A_M^T \mathbf{y} - \mathbf{e}) \quad (1.26)$$

with equality if and only if for all $\mathbf{x} \in M$,

$$f_k = f_\lambda \cdot \exp(A_M^T \mathbf{y} - \mathbf{e}). \quad (1.27)$$

where \cdot represents the element-wise multiplication of vectors.

Looking at Expression 1.27, the vector $\exp(A_M^T \mathbf{y} - \mathbf{e})$ can be looked at as the vector with which the background probability mass function f_λ will be re-weighted. Let the feasible regions of PI problems (1.24) and (1.25) be denoted by \mathcal{P} and \mathcal{D} , respectively.

Corollary 1.2.2. *If Expression 1.27 holds for some $f_k \in \mathcal{P}$ and $\mathbf{y} \in \mathcal{D}$ then they are both optimal and the duality gap⁵ is zero.*

Hence, necessary and sufficient conditions for optimality for PI problem (1.24) are obtained if $f_k = f_\lambda \cdot \exp(A_M^T \mathbf{y}_M - \mathbf{e})$. Furthermore, the primal formulation of PI problem (1.24) always has an optimal solution, if it is feasible, and the objective function is bounded.

1.2.3 PARFUM versus PREJUDICE

Both implementations, PARFUM and PREJUDICE, were applied to the dispersion coefficient example. The number of scenarios considered by PARFUM and PREJUDICE were 5 and 50, respectively, and the domain M consisted in both cases of 900 samples. In order to avoid numerical problems at $f_k = 0$, the interior point method contained in the commercial

⁵The duality gap is defined as the difference between the left hand side and right hand side of Inequality 1.26.

optimization package MOSEK (www.mosek.com) was used. Quantile information and Spearman's rank correlations of the distribution on the target variables are given in Table 1.1. The results of PREJUDICE are based on the uniform background distribution on M . Figure 1.13 presents the marginal distributions of the target variables, graphically.

	Quantile	PARFUM	PREJUDICE
A_y	5%	1.06e-1	2.24e-2
	50%	3.11e-1	4.19e-1
	95%	9.86e-1	3.48
B_y	5%	8.09e-1	5.87e-1
	50%	8.84e-1	8.63e-1
	95%	9.66e-1	1.19
ρ_{A_y, B_y}		-6.09e-2	-8.99e-1

Table 1.1: *Dispersion coefficient example (stability class C): quantile information and Spearman's rank correlations on target variables from PARFUM and PREJUDICE.*

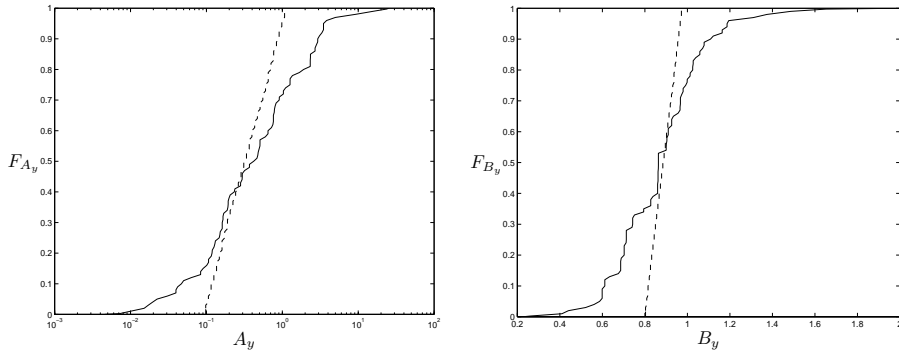


Figure 1.13: *Dispersion coefficient example (stability class C): graphical comparison of distributions of A_y (left) and B_y (right) as determined by PARFUM (- -) and PREJUDICE (-).*

The most important step in the implementation of Probabilistic Inversion is the determination of the domain M in the target variable space. The Monte Carlo heuristics used in determining M consist of three elements: (i) the assessments of the Decision Maker (DM), (ii) the mapping T , and (iii) the physics underlying the problem. These three elements meet in Expression 1.19 of **Step 1** of the implementation. Since the dispersion coefficient example has two target variables only, it is possible to display the domain M . Figure 1.14 shows the domain M_{PARF} and M_{PRED} . The domain M_{PRED} is more spread out than M_{PARF} . This is explained by the manner in which scenarios are generated in PREJUDICE and PARFUM.

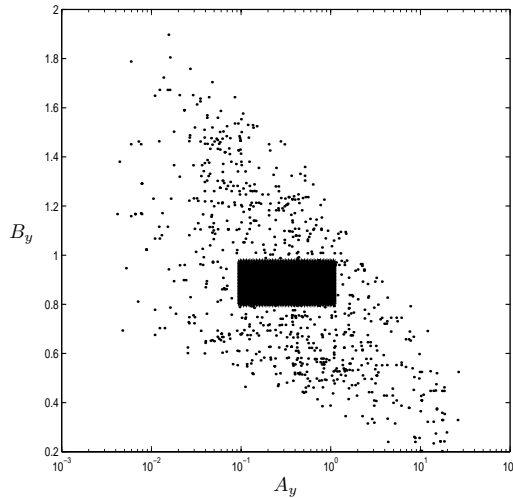


Figure 1.14: *Dispersion coefficient example (stability class C): graphical comparison of domain M as determined by PARFUM (black area) and PREJUDICE (\cdot).*

To measure the performance of PARFUM and PREJUDICE, their distributions on M are pushed through the power law function and compared to the DM distributions. Table 1.2 gives quantile information of the DM distributions and the push-forward distributions of PARFUM and PREJUDICE.

On the basis of comparing quantile information of the DM distribution and the push-forward distributions, it is concluded that the push-forward distributions of PREJUDICE resemble the DM distributions very well. The push-forward of PARFUM has the tendency to slightly overestimate the uncertainty expressed by the DM, whereas the medians are comparable to the DM medians.

The differences between the PARFUM and PREJUDICE solution schemes can be summarized as:

1. Generation of scenarios in PARFUM is based only on a certain quantile of the distribution of the elicitation variables. This way of generating scenarios is considered to result in a set of potentially observable scenarios which is too sparse. Taking account of combinations of different quantile points of the distributions of the elicitation variables leads to a larger set of potentially observable scenarios.
2. The construction of domain M in PARFUM may lead to computational problems in case of a large number of target variables. Therefore different sampling schemes were introduced for PREJUDICE.

Distance		$\sigma_y(z_i)$		
		DM	PARFUM	PREJUDICE
500 m.	5%	3.30e+1	3.32e+1	3.38e+1
	50%	9.49e+1	8.72e+1	9.55e+1
	95%	1.95e+2	2.09e+2	1.94e+2
1 km.	5%	6.48e+1	5.92e+1	6.57e+1
	50%	1.72e+2	1.61e+2	1.74e+2
	95%	3.46e+2	3.90e+2	3.41e+2
3 km.	5%	1.75e+2	1.54e+2	1.77e+2
	50%	4.46e+2	4.28e+2	4.53e+2
	95%	1.04e+3	1.10e+3	1.03e+3
10 km.	5%	4.48e+2	4.37e+2	4.49e+2
	50%	1.22e+3	1.20e+3	1.26e+3
	95%	3.37e+3	3.20e+3	3.36e+3
30 km.	5%	1.10e+3	1.10e+3	1.11e+3
	50%	2.82e+3	3.30e+3	2.83e+3
	95%	8.25e+3	8.90e+3	8.17e+3

Table 1.2: *Dispersion coefficient example (stability class C): quantile information comparison between DM distributions vs. PARFUM and PREJUDICE ‘push-forward’ distributions.*

3. The optimal fitting for PARFUM is done in the target variable space, whereas PREJUDICE performs the optimal fitting in the observable space. As a result of this, the push-forward distributions of PREJUDICE resemble the distributions of the DM better than the push-forward distributions of PARFUM.
4. PARFUM is easily implemented and nearly always feasible because each elicitation variable has its own observable space, whereas PREJUDICE is not easily implemented and not always feasible, because the observable space is constructed of all elicitation variables.

From the above it is concluded that PREJUDICE performs better than PARFUM, even though it may be hampered by infeasibilities. Hereafter the attention is focused on PREJUDICE. Special attention will be given to develop a procedure to check if M is observationally complete and how to deal with a Probabilistic Inversion problem which is infeasible.

So far no attention has been given to the background measure λ . Since the probability mass function f_k is a re-weighted version of f_λ (see Expression 1.27), it is important to determine a background measure which is representative. As mentioned earlier the background measure λ plays a similar role as the prior distribution in Bayesian Analysis. The next section addresses the determination of background measure λ .

1.3 Background measure

Recall that the domain M is based on the model inversions, as determined in the **Step 1** of the implementation. In many cases λ is constructed by combining (dependent or independent) marginal background measures λ_i of target variables X_i . Research into the implications of taking dependency information into account is still ongoing. In the remainder of this thesis the background measure λ is constructed by combining the marginal measures λ_i independently.

The assignment of a marginal background measures λ_i to a target variable X_i depends on numerous factors; results reported in literature, previous studies to quantify its uncertainty, spread in its model inversions and the physics underlying the problem. For certain target variables the spread in model inversions can be the only information available. For this type of situation, the following is suggested; if the spread in model inversions for a target variable covers two orders of magnitude or less, the uniform background distribution is assigned to the model inversions⁶. In case the spread of model inversions covers more than two orders of magnitude, the uniform background distribution is assigned to the orders of magnitude. Hence, the (positively oriented) target variable is considered to be log-uniformly distributed over the area.

1.3.1 Uniform background measure

Besides being a suitable background measure in certain situations, the uniform background measure has some useful computational features as well.

Given λ is the uniform background measure, PI problem (1.24) can be written as

$$\begin{aligned} \min_{f_k} \quad & \sum_{\mathbf{x} \in M} f_k(\mathbf{x}) \log f_k(\mathbf{x}) & (1.28) \\ \text{s.t.} \quad & A_M f_k = \mathbf{b} \\ & f_k(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in M. \end{aligned}$$

The contribution of the uniform background measure to the objective function of PI problem (1.28) is omitted as it is a constant ($\log(n')$). Finally note that minimizing the objective function of PI problem (1.28) is equivalent to maximizing the entropy of probability mass function f_k ; the entropy of probability mass function f_k is defined as $H(f_k) := -\sum_{\mathbf{x} \in M} f_k(\mathbf{x}) \log f_k(\mathbf{x})$.

Suppose sample $\mathbf{x}' \notin M$ is added to PI problem (1.28) which satisfies $T_k(\mathbf{x}') \in T_k(M)$, where $T_k(M) := \cup_{\mathbf{x} \in M} T_k(\mathbf{x})$; \mathbf{x}' maps into a hypercube which has already been covered by the propagation of M . PI problem (1.28)

⁶Intuitively, any sample from that area is considered to be as reasonable as any other sample from the same area. This assumption is reflected by the uniform distribution.

is extended with \mathbf{x}'

$$\begin{aligned} \min_{f_k} \quad & \sum_{\mathbf{x} \in M} f_k(\mathbf{x}) \log f_k(\mathbf{x}) + f_k(\mathbf{x}') \log f_k(\mathbf{x}') & (1.29) \\ \text{s.t.} \quad & A_M f_k + A_{\mathbf{x}'} f_k(\mathbf{x}') = \mathbf{b} \\ & f_k \geq 0. \end{aligned}$$

The column vector $A_{\mathbf{x}'} \in \mathbb{R}^{n \cdot (K+1)}$ contains the information based on the indicator function of Expression 1.23 for \mathbf{x}' . Lemma 1.3.1 is specific for the uniform background measure.

Lemma 1.3.1. *Given λ is uniform, if $T_k(\mathbf{x}') = T_k(\mathbf{x})$ then $f_k(\mathbf{x}') = f_k(\mathbf{x})$.*

Proof: Since λ is uniform, it follows $f_\lambda(\mathbf{x}') = f_\lambda(\mathbf{x})$. Furthermore, if $T_k(\mathbf{x}') = T_k(\mathbf{x})$ then column vectors $A_{\mathbf{x}'}$ and $A_{\mathbf{x}}$ are equal. From the above and Expression 1.27,

$$\begin{aligned} f_k(\mathbf{x}') &= f_\lambda(\mathbf{x}') \exp(A_{\mathbf{x}'}^T \mathbf{y} - 1) = \\ &= f_\lambda(\mathbf{x}) \exp(A_{\mathbf{x}}^T \mathbf{y} - 1) = f_k(\mathbf{x}). \end{aligned}$$

□

Lemma 1.3.1 states that in case of the uniform background measure, samples whose images are contained in the same hypercube, even though $T_{(j)}(\mathbf{x}') \neq T_{(j)}(\mathbf{x})$ for some or all j ($j = 1, \dots, n$), receive the same probability.

The implications of Lemma 1.3.1 to PI problem (1.29) are for the objective function if for some $\mathbf{x}^* \in M$, $T_k(\mathbf{x}^*) = T_k(\mathbf{x})$ then by Lemma 1.3.1, the objective function of PI problem (1.29) can be written as

$$\sum_{\mathbf{x} \in M \setminus \mathbf{x}^*} f_k^*(\mathbf{x}) \log f_k^*(\mathbf{x}) + 2 f_k^*(\mathbf{x}') \log f_k^*(\mathbf{x}') \quad (1.30)$$

The reason f_k^* is written in Expression 1.30 instead of f_k is because f_k^* is not a probability mass function as opposed to f_k . A similar analysis can be conducted for the constraints of PI problem (1.29), which will result in the entries of the column of $A_{\mathbf{x}'}$ describing hypercube $T_k(\mathbf{x}')$ being equal to 2. Since propagated samples in the same hypercube receive the same probability, it is more convenient to write PI problem (1.29) in terms of hypercubes: let $M := M \cup \mathbf{x}'$ then, by Corollary 1.1.7, solving PI problem (1.29) is equivalent to determining \tilde{g}_k^* which is the solution of:

$$\begin{aligned} \min_{g_k^*} \quad & \sum_{l \in T_k(M)} a(l) g_k^*(l) \log g_k^*(l) & (1.31) \\ \text{s.t.} \quad & B_{M_k} g_k^* = \mathbf{b} \\ & g_k^* \geq 0 \end{aligned}$$

where $a(l) := \sum_{\mathbf{x} \in M} 1_{T_k(\mathbf{x})=l}(\mathbf{x})$. Intuitively, $a(l)$ represents the number of samples in hypercube l and $B_{M_k} \in \mathbb{R}^{n \cdot (K+1) \times \#(T_k(M))}$ represents the constraint matrix for this formulation, with $\#(T_k(M))$ representing the number

of hypercubes covered by $T_k(M)$. The entries of B_{M_k} are positive integers and not restricted to 0 and 1 like for A_M . The probability mass function \tilde{f}_k on M is defined as, for each $\mathbf{x} \in M$

$$\tilde{f}_k(\mathbf{x}) := \tilde{g}_k^*(l) 1_{T_k(\mathbf{x}) \in l}(\mathbf{x}). \quad (1.32)$$

Note that the solution of PI problem (1.31) $\tilde{g}_k^* \in \mathbb{R}^{\#(T_k(M))}$, whereas the solution of PI problem (1.28) $\tilde{f}_k \in \mathbb{R}^{n'}$. In case sample size n' is large, PI problem (1.31) may involve far less optimization variables than PI problem (1.28) which is interesting from a computational point of view.

1.3.2 Results

The model inversions, on which the measure $\tilde{\mu}_k$ of Table 1.1 is based, are given in Figure 1.15. From Figure 1.15 it can be seen that the model inversions of A_y cover several orders of magnitude.

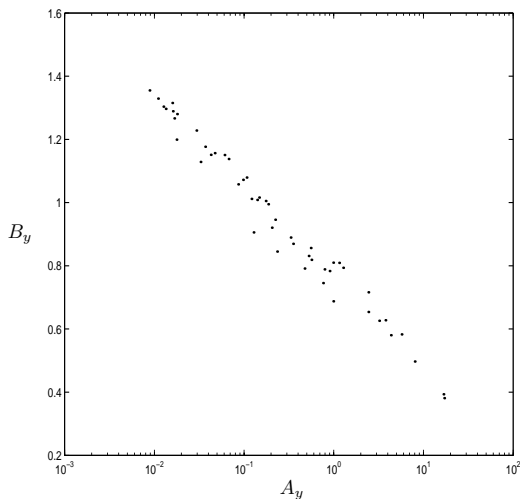


Figure 1.15: *Dispersion coefficient example (stability class C): model inversions for 50 scenarios.*

According to the recommendations for choosing a background measure, it is ‘reasonable’ to assume that the orders of magnitude of A_y are uniformly distributed. Table 1.3 compares quantile information on the target variables using a uniform distribution and a log-uniform background distribution for A_y . The background distribution for B_y is uniform. In both cases, the samples taken from domain M are identical.

The quantiles of the push-forward distributions of the target variables comply with the quantiles of the DM distributions, see Table 1. Figure 1.16 compares the marginal distributions of A_y and B_y , graphically. On the

	Quantile	A_y uniform	A_y log-uniform
A_y	5%	2.24e-2	9.85e-3
	50%	4.19e-1	1.65e-1
	95%	3.48	2.35
B_y	5%	5.87e-1	6.86e-1
	50%	8.63e-1	9.73e-1
	95%	1.19	1.33
ρ_{A_y, B_y}		-8.99e-1	-8.03e-1

Table 1.3: *Dispersion coefficient example (stability class C): quantile information on marginal distributions of A_y and B_y , using a uniform and log-uniform background distribution for A_y and a uniform distribution for B_y .*

basis of Table 1.3 and Figure 1.16 it is concluded that there is a difference between the two cases, which, just by eye-balling, is considered significant.

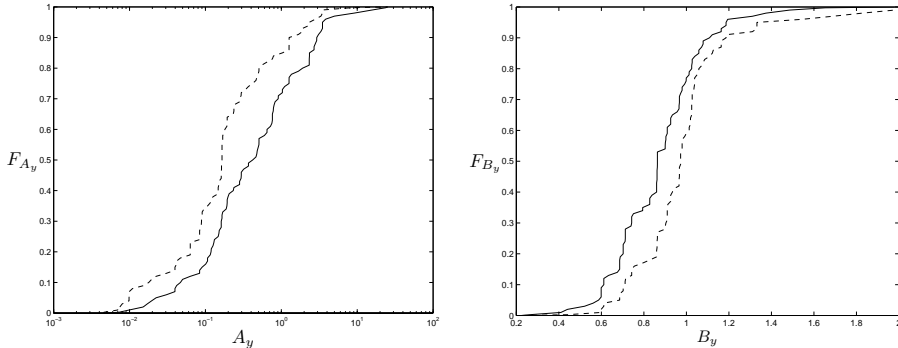


Figure 1.16: *Dispersion coefficient example (stability class C): graphical comparison of distributions of A_y (left) and B_y (right) as determined using a log-uniform background distribution (- -) and uniform background distribution (-) for A_y .*

Unless stated otherwise, the target variable A_y is assigned the log-uniform background distribution in the remainder of this thesis.

Even more important than determining the background measure on M is the determination of M itself. The most crucial element in the implementation is the determination of M , such that it is observationally complete. Since **Step 1** of the implementation is conducted for a selection of potentially observable scenarios, it is not clear if the resulting M is observationally complete. In the next section a heuristic is developed which is capable of checking if M is observationally complete.

1.4 Iterative PREJUDICE

Recall that the domain M is observationally complete if every potentially observable scenario of N can be reproduced using T and M , in notation $N \subseteq T(M)$. Many times a domain M , as determined in **Step 1** of the PREJUDICE implementation, is based on model inversions which are obtained from a selection of potentially observable scenarios. Because M is based on a selection, it is not clear if M is observationally complete. To illustrate this, consider Table 1.4 where results are presented for the dispersion coefficient example. For each run, 50 potentially observable scenarios were randomly selected, which determined the different domains on which the measure $\tilde{\mu}_k$ are specified. The total number of samples for each run equaled 900.

		Run 1	Run 2	Run 3	Run 4	Run 5
A_y	5%	9.85e-3	1.64e-2	1.13e-3	1.11e-2	1.97e-2
	50%	1.65e-1	1.71e-1	2.46e-1	1.68e-1	2.04e-1
	95%	2.35	3.88	3.80	1.82	4.63
B_y	5%	6.86e-1	6.38e-1	6.16e-1	7.08e-1	5.70e-1
	50%	9.73e-1	9.34e-1	9.62e-1	9.57e-1	9.12e-1
	95%	1.33	1.27	1.30	1.23	1.23
ρ_{A_y, B_y}		-8.03e-1	-8.96e-1	-8.65e-1	-7.74e-1	-8.22e-1
Relative Information		3.31	3.62	3.34	3.18	2.70

Table 1.4: *Dispersion coefficient example (stability class C): 5 independent runs.*

Based on the results presented in Table 1.4 it is concluded that the marginal distributions among runs are not similar, especially for A_y . Additionally, the spread in the relative information values suggests that the measures on the respective domains are not similar. In case the domains are considered to be observationally complete, the marginal distributions and relative information values between runs are expected to be more similar.

The domains obtained from the 5 runs can be regarded as the result of the first iteration of the probabilistic inversion implementation; in notation, let $M^{(1)}$ denote the domain obtained after the first iteration. The key question addressed in this section is:

If domain $M^{(1)}$ were to be extended with a certain set M^* , would the extension $M^{(1)} \cup M^*$ be observationally complete? And would the marginal distributions and relative information values between runs be similar?

Starting from scratch, one of the following situations will occur after propagation of $M^{(1)}$ through T :

Situation A : no distribution on $M^{(1)}$ can be determined.

Situation B : a distribution on $M^{(1)}$ can be determined, although $M^{(1)}$ may *not* be observationally complete.

Before discussing **Situation A** in terms of extending the domain $M^{(1)}$ and thus keeping the dimension of the observable space constant, another strategy is available which reduces the dimension of the observable space. This technique is known as Reduction of Dimension and is described in Appendix C.

Dealing with **Situation A** and **Situation B** in terms of extending domain $M^{(1)}$, let $M^{(\infty)}$ denote the observationally complete domain. Let the primal formulation of the observationally complete problem be given by:

$$\begin{aligned} \min_{f_k^{(\infty)}} \quad & \sum_{\mathbf{x} \in M^{(\infty)}} f_k^{(\infty)}(\mathbf{x}) \log \frac{f_k^{(\infty)}(\mathbf{x})}{f_\lambda^{(\infty)}(\mathbf{x})} \\ \text{s.t.} \quad & A_{M^{(\infty)}} f_k^{(\infty)} = \mathbf{b} \\ & f_k^{(\infty)} \geq 0. \end{aligned} \quad (1.33)$$

Without specifying the dimension of the vectors and matrices, it is assumed that the multiplications and equations of PI problem (1.33) are well defined. The dual formulation of PI problem (1.33) is

$$\max_{\mathbf{y}_{M^{(\infty)}}} \quad \mathbf{y}_{M^{(\infty)}}^T \mathbf{b} - \left(f_\lambda^{(\infty)} \right)^T \exp(A_{M^{(\infty)}}^T \mathbf{y}_{M^{(\infty)}} - \mathbf{e}) \quad (1.34)$$

where $\mathbf{y}_{M^{(\infty)}}$ is the dual variable for the observationally complete problem.

Note that the observationally complete problem is an artifact. It is introduced here for the purpose of illustration only; in performing probabilistic inversion the observationally complete problem is unknown, however it is desired to design criteria which are able to check if a PI problem formulation is close to the observationally complete problem.

As mentioned before, since domain $M^{(1)}$ is based on model inversions determined for a selection of potentially observable scenarios N , it is not unreasonable to assume that $M^{(1)}$ is not observationally complete. Let the primal formulation of the PI problem formulated on samples after propagation of $M^{(1)}$ be given by

$$\begin{aligned} \min_{f_k^{(1)}} \quad & \sum_{\mathbf{x} \in M^{(1)}} f_k^{(1)}(\mathbf{x}) \log \frac{f_k^{(1)}(\mathbf{x})}{f_\lambda^{(1)}(\mathbf{x})} \\ \text{s.t.} \quad & A_{M^{(1)}} f_k^{(1)} = \mathbf{b} \\ & f_k^{(1)} \geq 0 \end{aligned} \quad (1.35)$$

where $f_k^{(1)}$ and $f_\lambda^{(1)}$ represent the probability mass function and background probability mass function on the samples in $M^{(1)}$, respectively.

The dual formulation of PI problem (1.35) is

$$\max_{\mathbf{y}_{M^{(1)}}} \mathbf{y}_{M^{(1)}}^T \mathbf{b} - \left(f_\lambda^{(1)}\right)^T \exp(A_{M^{(1)}}^T \mathbf{y}_{M^{(1)}} - \mathbf{e}). \quad (1.36)$$

Note that the constraint matrix $A_M^{(\infty)}$ of the observationally complete problem can be represented as

$$A_{M^{(\infty)}} = A_{M^{(1)} \cup (M^{(\infty)} \setminus M^{(1)})}. \quad (1.37)$$

Using Expression 1.37, the dual formulation of the observationally complete problem (1.34) can be written as

$$\begin{aligned} \max_{\mathbf{y}_{M^{(\infty)}}} \quad & \mathbf{y}_{M^{(\infty)}}^T \mathbf{b} - \left(f_\lambda^{(1)}\right)^T \exp(A_{M^{(1)}}^T \mathbf{y}_{M^{(\infty)}} - \mathbf{e}) + \\ & - \left(f_\lambda^{(\infty) \setminus (1)}\right)^T \exp(A_{M^{(\infty)} \setminus M^{(1)}}^T \mathbf{y}_{M^{(\infty)}} - \mathbf{e}) \end{aligned} \quad (1.38)$$

where $f_\lambda^{(\infty) \setminus (1)}$ represents the probability mass function of the background measure on $M^{(\infty)} \setminus M^{(1)}$. In case $M^{(1)}$ is *not* observationally complete, the contribution of the third term of dual formulation 1.38

$$\left(f_\lambda^{(\infty) \setminus (1)}\right)^T \exp(A_{M^{(\infty)} \setminus M^{(1)}}^T \mathbf{y}_{M^{(\infty)}} - \mathbf{e}) \quad (1.39)$$

is greater than zero, because $M^{(\infty)} \setminus M^{(1)}$ is non-empty. The intuitive understanding of the criteria to check for observational completeness, is to extend domain $M^{(1)}$ to $M^{(2)}$ such that Expression 1.39 reduces to zero. In that case, the dual solution $\mathbf{y}_{M^{(2)}}$ will equal the dual solution $\mathbf{y}_{M^{(\infty)}}$.

Situation A : Given that PI problem (1.35) is infeasible, its dual formulation (1.36) is unbounded. If the dual formulation (1.36) is unbounded, then the existence of a \mathbf{y}^* can be shown such that

$$(\mathbf{y}^*)^T \mathbf{b} > 0 \quad (1.40)$$

$$A_{M^{(1)}}^T \mathbf{y}^* \leq 0. \quad (1.41)$$

The unboundedness of the dual formulation (1.36) is easily seen, since $\alpha \mathbf{y}^*$ ($\alpha > 0$) satisfies Inequalities 1.40 and 1.41 also. In order to make dual formulation (1.36) bounded, domain $M^{(1)}$ has to be extended with a set $M_d \subseteq M^{(\infty)} \setminus M^{(1)}$, such that the third term in dual formulation (1.38) becomes larger than zero. In order to determine M_d efficiently the following heuristic is proposed: find an observable hypercube, indexed $l_d \in T_k(M \setminus M^{(1)})$, which maximizes $\mathbf{e}_l^T \mathbf{y}^*$ for all $l \in T_k(M \setminus M^{(1)})$, where \mathbf{e}_l is the vector with entries 1 on positions which correspond to the observable hypercube indexed by l . Briefly,

determine model inversions $\mathbf{x}_{\mathbf{y}_{l_d}}$ based on potentially observable scenarios \mathbf{y}_{l_d} sampled from l_d , and generate a domain $M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$ for each $\mathbf{x}_{\mathbf{y}_{l_d}}$. The domain $M^{(1)}$ is extended with the union of $M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$. The stopping criteria for the heuristic would be if no \mathbf{y}^* can be found which satisfies Inequalities 1.40 and 1.41.

The construction of a feasible PI problem will result most likely in a PI problem of the type as described under **Situation B**.

Situation B : As mentioned earlier, note that the objective function of dual formulation (1.38) contains the objective function of dual formulation (1.36), in case $\mathbf{y}_{M^{(\infty)}} = \mathbf{y}_{M^{(1)}}$. Based on this observation it is concluded that if $M^{(1)}$ is close to being observationally complete then the third term of dual formulation (1.38) should be close to zero. The rough idea for the heuristic is to extend $M^{(1)}$ by $M_d \subset M^{(\infty)}$ such that the contribution of the third term of dual formulation (1.38) reduces to zero. In order to determine the extension M_d efficiently the following heuristic is proposed: find observable hypercube, indexed by l_d , which maximizes $\mathbf{e}_l^T \mathbf{y}_{M^{(1)}}$ for all $l \in T_k(M^{(\infty)})$, where \mathbf{e}_l is the vector with entries 1 on positions which correspond to hypercube indexed by l . Briefly, determine model inversions $\mathbf{x}_{\mathbf{y}_{l_d}}$ based on potentially observable scenarios \mathbf{y}_{l_d} sampled from l_d , and generate a domain $M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$ for each $\mathbf{x}_{\mathbf{y}_{l_d}}$. The domain $M^{(1)}$ is extended with the union of $M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$. Note that the heuristic suggested is similar to the heuristic suggested for **Situation A**. A stopping criteria in this situation would be to stop if the value of $\left(f_\lambda^{(i+1) \setminus (i)}\right)^T \exp(A_{M^{(i+1)} \setminus M^{(i)}}^T \mathbf{y}_{M^{(i+1)}} - \mathbf{e})$ at the $i+1$ iteration is close to zero.

In conclusion, the heuristic suggested will extend the domain $M^{(1)}$ based on the dual solution $\mathbf{y}_{M^{(1)}}$ and can be applied both in **Situation A** and **Situation B**.

Since the heuristic suggested can be applied both in **Situation A** and **Situation B**, the derivation of the iterative version of PREJUDICE will be illustrated for **Situation B** only. Hence, PI problem (1.35) is considered to be feasible although $M^{(1)}$ is not observationally complete. The extension of $M^{(1)}$ is modeled as an iterative process. The extension is illustrated using one sample but can be extended easily to sets of samples.

Suppose $M^{(1)}$ is extended with sample $\mathbf{x}_d \in M_{\mathbf{x}_{f_{j_d}}}$. PI problem (1.35) becomes:

$$\begin{aligned} \min_{f_k^{(1)}, f_k} \quad & \sum_{\mathbf{x} \in M^{(1)}} f_k^{(1)}(\mathbf{x}) \log \frac{f_k^{(1)}(\mathbf{x})}{f_k^{(1)}(\mathbf{x})} + f_k(\mathbf{x}_d) \log \frac{f_k(\mathbf{x}_d)}{f_k(\mathbf{x}_d)} \quad (1.42) \\ \text{s.t.} \quad & A_{M^{(1)}} f_k^{(1)} + A_{\mathbf{x}_d} f_k(\mathbf{x}_d) = \mathbf{b} \\ & f_k^{(1)}, f_k(\mathbf{x}_d) \geq 0 \end{aligned}$$

Depending on the ratio $\frac{f_k(\mathbf{x}_d)}{f_\lambda(\mathbf{x}_d)}$, the sign of the last term in the objective function of PI problem (1.42) will be positive or negative. Hence, it is unclear if the value of the objective function in the next iteration will increase or decrease. However, much more important is that repetitive application of the heuristic will lead to an observationally complete domain. Obtaining an observationally complete domain is much more important than the value of the objective function.

Furthermore, note that the solution of PI problem (1.35) can be used to construct a solution for PI problem (1.42); assign probability zero to \mathbf{x}_d and leave $f_k^{(1)}$ untouched⁷. This way of constructing a new measure in the next iteration should be satisfied by the implementation of the iterative version of PREJUDICE.

The implementation of the iterative version of PREJUDICE involves the following steps:

Step 1 Determination of $M^{(1)}$; see Section 1.2.2 and let $i = 1$.

Step 2 Determination of distribution: Solve

$$\max_{\mathbf{y}_{M^{(i)}}} \mathbf{y}_{M^{(i)}}^T \mathbf{b} - \left(f_\lambda^{(i)} \right)^T \exp(A_{M^{(i)}}^T \mathbf{y}_{M^{(i)}} - \mathbf{e}). \quad (1.43)$$

Step 3 $N \not\subseteq \mathbf{T}(M^{(i)})$: for the i -th iteration, the dual solution $\mathbf{y}_{M^{(i)}}$ of the PI problem is used to determine observable hypercube, indexed by $l_d \in \{1, \dots, k\}$, which maximizes $\mathbf{e}_l^T \mathbf{y}_{M^{(i)}}$ for all $l \in \{1, \dots, k\}$. If $\mathbf{e}_{l_d}^T \mathbf{y}_{M^{(i)}}$ is less than a certain cutoff, then it is assumed that the i -th PI problem formulation resembles the observationally complete problem closely and the algorithm is stopped; if $\mathbf{e}_{l_d}^T \mathbf{y}_{M^{(i)}}$ is larger than the cutoff, potentially observable scenarios \mathbf{y}_{l_d} are sampled randomly from l_d and model inversions $\mathbf{x}_{\mathbf{y}_{l_d}}$ are determined, which in turn determine $M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$. Domain $M^{(i)}$ is extended with the union of $M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$ to obtain $M^{(i+1)}$. Next, $M^{(i+1)}$ is propagated through T and the PI problem for the $i + 1$ -th iteration is obtained. Go to **Step 2**.

In case the iterative version of PREJUDICE ends at the i -th iteration, the probability mass function $\tilde{f}_k^{(i)}$ is obtained by element-wise multiplication:

$$\tilde{f}_k^{(i)} = f_\lambda^{(i)} \cdot \exp(A_{M^{(i)}}^T \tilde{\mathbf{y}}_{M^{(i)}} - \mathbf{e}) \quad (1.44)$$

where $\tilde{\mathbf{y}}_{M^{(i)}}$ maximizes PI problem (1.43).

⁷Note that this solution is a solution of PI problem (1.42), however it is likely not to be the solution which minimizes relative information value.

1.4.1 Iterative PREJUDICE: Uniform background measure

Section 1.3.1 discussed the implications of selecting the uniform background distribution on the target variables. In case of the uniform background distribution, the number of optimization variables is equal to the number of hypercubes covered. This observation allowed the PI problem to be formulated in terms of hypercubes covered. The question put forward here is: does the heuristic of Section 1.4 to determine if $M^{(i)}$ is observationally complete, hold for the formulation in terms of hypercubes as well? The answer to this question is negative.

Dealing with a PI problem formulation depending on hypercubes covered, domain M will be termed *hypercube-complete* if $\phi_k(N) \subseteq \phi_k(T(M)) = T_k(M)$. In case K quantile points are queried and uniform background distribution is considered, let the primal formulation of the hypercube-complete problem be given by

$$\begin{aligned} \min_{g_k} \quad & \sum_{l \in T_k(M)} a(l) g_k(l) \log g_k(l) \\ \text{s.t.} \quad & B_{M_k} g_k = \mathbf{b} \\ & g_k \geq 0. \end{aligned} \tag{1.45}$$

For notational convenience the star superscript, as introduced in the formulation of PI problem (1.31), has been dropped, however it must be stressed that g_k is still *not* a probability mass function in this section.

The dual formulation of the hypercube-complete problem is given by

$$\max_{\mathbf{y}_{M_k}} \quad \mathbf{y}_{M_k}^T \mathbf{b} - \mathbf{a}^T \exp(D_{M_k}^T \mathbf{y}_{M_k} - \mathbf{e})$$

where $\mathbf{a} \in \mathbb{R}^{\#(T_k(M))}$ with elements $a(l)$ for $l \in T_k(M)$ and D_{M_k} is of the same dimension as B_{M_k} , but has entries 1 where B_{M_k} has non-zero entries..

At the i -th iteration, let the dual formulation of a PI problem be given by

$$\max_{\mathbf{y}_{M_k^{(i)}}} \quad \mathbf{y}_{M_k^{(i)}}^T \mathbf{b} - (\mathbf{a}^{(i)})^T \exp(D_{M_k^{(i)}}^T \mathbf{y}_{M_k^{(i)}} - \mathbf{e}) \tag{1.46}$$

Application of the heuristic suggested will lead to the determination of an observable hypercube, indexed by l_d , which maximizes $\mathbf{e}_{l_d}^T \mathbf{y}_{M_k^{(i)}}$ for all $l \in \{1, \dots, k\}$. Potential observable scenarios \mathbf{y}_{l_d} are sampled for which model inversions $\mathbf{x}_{\mathbf{y}_{l_d}}$ will be determined. It is not unreasonable to assume that propagation of $M_{\text{ext}} := \cup_{\mathbf{x}_{\mathbf{y}_{l_d}}} M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$ will be such that some propagated samples are contained in l_d and some are included in hypercubes already contained in $T_k(M^{(i)})$; instead of extending $\mathbf{a}^{(i)}$ and $B_{M_k^{(i)}}$, its entries are likely to change as well. The changing of entries of $\mathbf{a}^{(i)}$ and $B_{M_k^{(i)}}$ in the $i + 1$ -th iteration has implications.

A requirement the iterative version of PREJUDICE had to satisfy is that the solution of the i -th iteration could be used to construct a solution for the $i + 1$ -th iteration, simply by setting the new optimization variables equal to 0. Looking at the constraints for the primal formulation of PI problem (1.46) at the $i + 1$ -th iteration, the following equation should hold element-wise

$$B_{M_k^{(i+1)}} \begin{pmatrix} g_k^{(i)} \\ \mathbf{0} \end{pmatrix} = \mathbf{b} \quad (1.47)$$

where $\mathbf{0} \in \mathbb{R}^{n^*}$ with $n^* = \#(T_k(M^{(i+1)})) - \#(T_k(M^{(i)}))$, i.e. the extra number of hypercubes covered by $T_k(M^{(i+1)})$ compared to $T_k(M^{(i)})$. Equation 1.47 cannot hold in general since the entries of $B_{M_k^{(i+1)}}$ are likely to be different than of $B_{M_k^{(i)}}$.

PI problem (1.45) can be looked at as a weighted entropy maximization problem. Roughly, the proposed solution to the problem described above is to focus first on the un-weighted version of the PI problem (1.45) and secondly weight the solution to obtain the measure on the domain.

The primal formulation of the un-weighted version of PI problem (1.45) at the i -th iteration is

$$\begin{aligned} \min_{g_k^{(i)}} \quad & \sum_{l \in T_k(M^{(i)})} g_k^{(i)}(l) \log g_k^{(i)}(l) \\ \text{s.t.} \quad & D_{M_k^{(i)}} g_k^{(i)} = \mathbf{b} \\ & g_k^{(i)} \geq 0. \end{aligned} \quad (1.48)$$

Next probability mass function $\tilde{f}_k^{(i)}(x)$ on $M^{(i)}$ can be determined via

$$\tilde{f}_k^{(i)}(\mathbf{x}) := \frac{\tilde{g}_k^{(i)}(l)}{a^{(i)}(l)} 1_{\{\mathbf{x} | T_k(\mathbf{x}) \in l\}}(\mathbf{x}) \quad (1.49)$$

PI problem formulation (1.48) in combination with Expression 1.49 bears strong resemblance with the Hora/Young method⁸ as described in [27].

The dual formulation of PI problem (1.48) for the i -th iteration is:

$$\max_{\mathbf{z}_{M_k^{(i)}}} \mathbf{z}_{M_k^{(i)}}^T \mathbf{b} - \mathbf{e}^T \exp(D_{M_k^{(i)}}^T \mathbf{z}_{M_k^{(i)}} - \mathbf{e}) \quad (1.50)$$

⁸Briefly, the main difference concerns the criteria to assign probability to the samples in case the assignment is not unique. The Hora/Young method chooses to minimize the average probability. The probability of an observable hypercube l is just $\frac{g_k^{(i)}(l)}{V_l}$, where V_l is the volume of observable hypercube l . Hence, in determining the minimal average probability the objective function of PI problem (1.48) is replaced by

$$\min_{g_k^{(i)}} \sum_{l \in T_k(M)} g_k^{(i)}(l) \frac{g_k^{(i)}(l)}{V_l}.$$

where $\mathbf{z}_{M_k^{(i)}} \in \mathbb{R}^{n \cdot (K+1)}$ is the dual variable. Note, if $\mathbf{a}^{(i)} = \mathbf{e}$ at the i -th iteration, PI problem (1.45) and PI problem (1.48) are the same; PI problem (1.48) can be viewed as a special case of PI problem (1.45). Furthermore, at each iteration PI problem (1.48) can be extracted from PI problem (1.45). The iterative version of PREJUDICE in case of the uniform background distribution will make use of the relationship between the two PI problem formulations.

Like the dual formulation for the observationally complete problem (1.38), the dual formulation for the hypercube-complete problem of PI problem (1.48) can be written as

$$\begin{aligned} \max_{\mathbf{z}_{M_k^{(\infty)}}} \quad & \left(\mathbf{z}_{M_k^{(\infty)}} \right)^T \mathbf{b} - \mathbf{e}^T \exp(D_{M_k^{(i)}}^T \mathbf{z}_{M_k^{(\infty)}} - \mathbf{e}) + \\ & - \mathbf{e}^T \exp(D_{M_k^{(\infty)} \setminus M_k^{(i)}}^T \mathbf{z}_{M_k^{(\infty)}} - \mathbf{e}). \end{aligned} \quad (1.51)$$

Application of the heuristic suggested results in the identification of hypercube l_d . From hypercube l_d , potentially observable scenarios \mathbf{y}_{l_d} are sampled randomly. Based on \mathbf{y}_{l_d} the model inversions $\mathbf{x}_{\mathbf{y}_{l_d}}$ are determined, which are used to determine $M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$. Again, it is reasonable to assume that propagation of $\cup_{\mathbf{x}_{\mathbf{y}_{l_d}}} M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$ will lead to propagated samples in observable hypercube l_d and hypercubes $T_k(M^{(i)})$. Since, the vector $\mathbf{a}^{(i)}$ is not contained in dual formulation (1.51), the value of the objective function will increase only if hypercube $l_d \in T_k(M^{(\infty)} \setminus M^{(i)})$ is covered.

These insights lead to a special iterative version of PREJUDICE in case of uniform background distribution.

Step 1 Determination of $M^{(1)}$; see Section 1.2.2. For $i = 1$, derive the dual formulation (1.46).

Step 2 Determination of distribution: extract dual formulation (1.50) from dual formulation (1.46). Solve dual formulation (1.50) to obtain $\mathbf{z}_{M_k^{(i)}}$.

Step 3 $\phi_K(N) \not\subseteq T_k(M^{(i)})$: in case of the i -th iteration, take dual solution $\mathbf{z}_{M_k^{(\infty)}}$ and determine an observable hypercube, indexed by l_d which maximizes $\mathbf{e}_l^T \mathbf{z}_{M_k^{(\infty)}}$ for all observable hypercubes, indexed by l . If $\mathbf{e}_{l_d}^T \mathbf{z}_{M_k^{(\infty)}}$ is less than a certain cutoff, then it is assumed that the i -th iteration of PI problem formulation resembles the hypercube-complete problem closely, go to **Step 4**; if l_d can be identified, potentially observable scenarios \mathbf{y}_{l_d} are sampled randomly from l_d and model inversions $\mathbf{x}_{\mathbf{y}_{l_d}}$ are obtained, which determine $M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$. Domain $M^{(i)}$ is extended with $\cup_{\mathbf{x}_{\mathbf{y}_{l_d}}} M_{\mathbf{x}_{\mathbf{y}_{l_d}}}$ to obtain $M^{(i+1)}$. Next, $M^{(i+1)}$ is propagated through T and the dual formulation (1.46) for the $i + 1$ -th iteration is obtained. Go to **Step 2**.

Step 4 : Solve dual formulation (1.46) to obtain $\mathbf{z}_{M_k^{(i)}}$ from which probability mass function $\tilde{f}_k^{(i)}$ on $M^{(i)}$ can be determined, via Expression 1.49.

Domain $M^{(i)}$ of **Step 4** will be close to hypercube-complete, however the aim of probabilistic inversion is to determine a domain such that $M^{(i)}$ is observationally complete. Given the uniform background distribution, it is has not been investigated at this stage how to extend a hypercube complete domain to an observationally complete domain.

1.4.2 Results

In this section results of the iterative version of PREJUDICE are given for the dispersion coefficient example, the systemic retention of Sr in the human body. No computation times are given as the software is research software, which at this stage is not optimized.

Dispersion coefficient example (stability class C)

The settings of the iterative version of PREJUDICE for the dispersion coefficient example were $\epsilon = 0.5$, the number of samples taken from the ϵ -neighborhood of the model inversions was 18, the number of scenarios to determine $M^{(1)}$ was 50, the number of scenarios sampled randomly from an observable hypercube equaled 10. Finally, it was stipulated that an observable hypercube could not be visited more than once.

In order to investigate the sensitivity of the choice of the cutoff value, the iterative version of PREJUDICE has been applied to the dispersion coefficient example when no cutoff was specified; the iterative version will stop when all observable hypercubes have been visited, which was obtained after 348 iterations at a relative information value of 4.78. At each iteration the cutoff value was determined, see Figure 1.17 (left).

Starting from the left and following the graph, the first circle indicates at how many iterations the iterative version would have stopped if the cutoff value would have been set to -1, the next circle gives information on how many iterations if the cutoff value would have been set to -2, etc. The same holds for the relative information graph of Figure 1.17 (right).

The behavior of relative information tends to increase with iterations, see Figure 1.17 (right). This is explained as follows: the iterative version of PREJUDICE searches for a domain M which is close to observationally complete. At the beginning (small number of iterations) it is likely that M is not observationally complete (or even such that the corresponding PI problem is infeasible). The dual solution is used to identify areas in the observable space which are not covered, but which need to be covered for M to be observationally complete. Thus $M^{(i)}$ at the i -th iteration may be ‘bigger’ than $M^{(i-1)}$, i.e. $M^{(i)}$ may include $M^{(i-1)}$. Since $M^{(i)}$

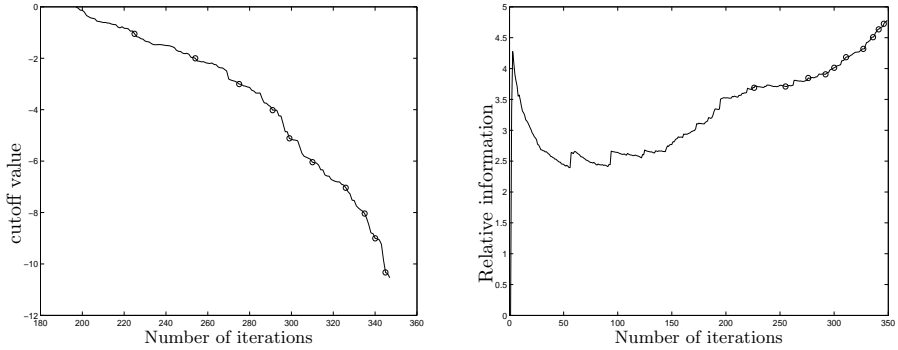


Figure 1.17: *Dispersion coefficient example (stability class C): cutoff value vs. number of iterations (left) and relative information value vs. number of iterations (right), with \circ representing different cutoff levels.*

is determined by the values of A_y and B_y , it is interesting to look at the minimum and maximum values of A_y and B_y over iterations, see Figure 1.18.

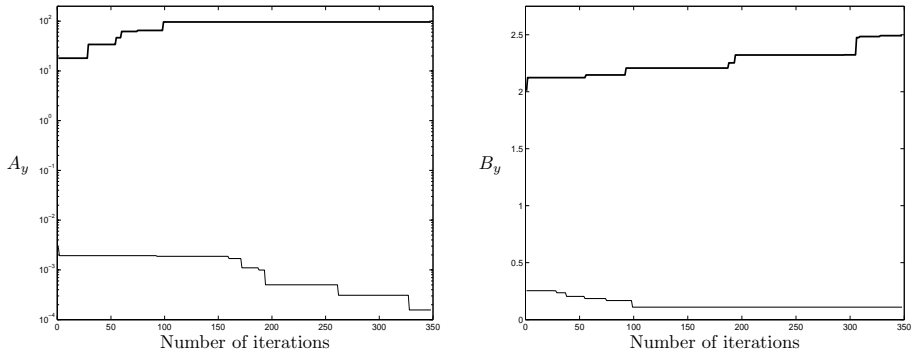


Figure 1.18: *Dispersion coefficient example (stability class C): The maximum (—) and minimum (---) values of A_y and B_y versus number of iterations.*

It is clear from Figure 1.18 that the domain $M^{(i)}$ ‘grows’ with the number of iterations, because the difference between the minimum and maximum values for both A_y and B_y grows with the number of iterations. The original domain $M^{(1)}$ and the final domain $M^{(348)}$ are displayed in Figure 1.19, clearly $M^{(348)}$ contains $M^{(1)}$.

It appears that the growing difference between the minimum and maximum values of A_y and B_y with iterations is the reason why the relative information value increases with the number of iterations. Intuitively, the domain $M^{(348)}$ contains domain $M^{(1)}$, which results in a higher relative in-

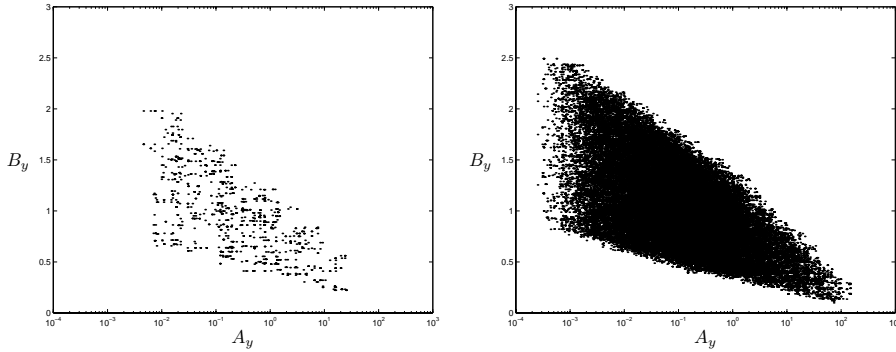


Figure 1.19: *Dispersion coefficient example (stability class C): graphical comparison between original domain $M^{(1)}$ (left) and final domain $M^{(348)}$ (right) for cutoff=-9.*

formation value for the distribution on $M^{(348)}$ compared to $M^{(1)}$ since the distribution on $M^{(348)}$ is more ‘dispersed’ than the distribution on $M^{(1)}$.

Besides the growing difference between minimum and maximum values of A_y and B_y with the number of iterations, one may wonder to what extent the sample size influences the relative information value; at each iteration a number of samples is added to the problem. Four situations were distinguished; for each model inversion 1, 4, 9 and 18 samples were drawn using the ϵ -neighborhood sampling scheme. Note that the situation involving 18 samples is equal to the situation considered previously. The relative information values for the four situations are listed in Table 1.4.2.

Number of samples in ϵ -neighborhood	Total number of samples	Relative information value
1	3520	4.77
4	14080	4.80
9	31680	4.77
18	63360	4.78

Table 1.5: *Dispersion coefficient example (stability class C): relative information value for different number of samples taken from ϵ -neighborhood.*

From Table 1.4.2 it is observed that sample size has very little influence on the relative information value.

At this moment no stopping criteria has been formulated; the user specifies a cutoff value based on the size of the PI problem. A large negative cutoff value will take longer than a small negative cutoff value. Research is conducted if a stopping criteria can be formulated which is based on the difference between the relative information value at cutoff c ($c < 0$) and $c - 1$. Based on Figure 1.17, a cutoff of -9 was chosen for the iterative version of

PREJUDICE for the dispersion coefficient example.

Next, the iterative version of PREJUDICE was applied 5 times to the dispersion coefficient example using a cutoff of -9. The setup for these 5 runs was similar to the setup of the 5 runs presented Section 1.4. For each run, $M^{(1)}$ was determined based on model inversion resulting from a random selection of 50 potentially observable scenarios. Hence, it is unlikely that $M^{(1)}$ for the 5 runs is the same. Table 1.6 lists the quantile information of A_y and B_y together with information on ρ_{A_y, B_y} , the number of iterations/samples and relative information values.

		Run 1	Run 2	Run 3	Run 4	Run 5
A_y	5%	2.44e-3	3.02e-3	2.52e-3	2.60e-3	2.68e-3
	50%	2.07e-1	2.01e-1	2.13e-1	2.09e-1	2.04e-1
	95%	3.54	3.13	3.00	3.17	3.44
B_y	5%	6.15e-1	6.18e-1	6.28e-1	6.23e-1	6.07e-1
	50%	9.51e-1	9.44e-1	9.43e-1	9.44e-1	9.36e-1
	95%	1.40	1.39	1.41	1.43	1.42
ρ_{A_y, B_y}		-9.07e-1	-8.97e-1	-9.15e-1	-9.10e-1	-9.19e-1
Number of iterations		323	301	303	339	329
Number of samples		57672	54180	54378	60894	58770
Relative Information		4.72	4.69	4.77	4.69	4.71

Table 1.6: *Dispersion coefficient example (stability class C): iterative PREJUDICE for 5 runs with a cutoff=-9.*

Due to the choice of cutoff and the stipulation of no more than ‘one visit’ to a hypercube, there is some variation in the relative information values and quantile points of the target variables in Table 1.6, but far less variation than the relative information values and quantile points as presented in Table 1.4. Figure 1.20 shows the relative information value as a function of number of iterations (left) and a close-up of the ‘tail’ of this relationship (right).

Based on the results of Table 1.6 and Figure 1.20 it is observed that the iterative version of PREJUDICE tends to ‘converge’ to a similar result. This observation is supported when the dispersion coefficient example of Section 1.3.2 is extended to look if it ‘converges’ to a similar distribution as presented in Table 1.6. Table 1.7 presents information of the distribution if the Original problem is extended to Extension, from which is concluded that the results under Extension are similar to the results presented in Table 1.6.

Figure 1.21 shows the distribution functions of the target variables for Original and Extension graphically. Note the difference at the lower and upper quantile points for both target variables.

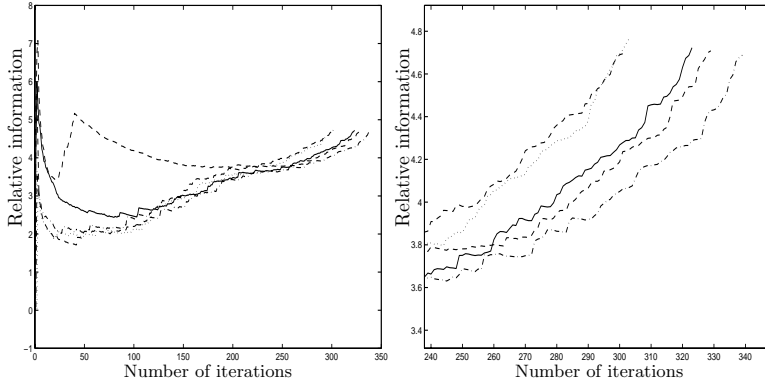


Figure 1.20: *Dispersion coefficient example (stability class C): graphical comparison of relative information values among 5 independent runs for cutoff=-9. Remark: Figure on the right is a close-up of the 'tail' of the Figure on the left.*

		Original	Extension
A_y	5%	9.85e-3	2.74e-3
	50%	1.65e-1	2.10e-1
	95%	2.35	3.39
B_y	5%	6.86e-1	6.12e-1
	50%	9.73e-1	9.46e-1
	95%	1.33	1.40
ρ_{A_y, B_y}		-8.03e-1	-9.08e-1
Number of iterations		N/A	340
Number of samples		900	62280
Relative Information		3.31	4.61

Table 1.7: *Dispersion coefficient example (stability class C): comparison Original and Extension for cutoff=-9.*

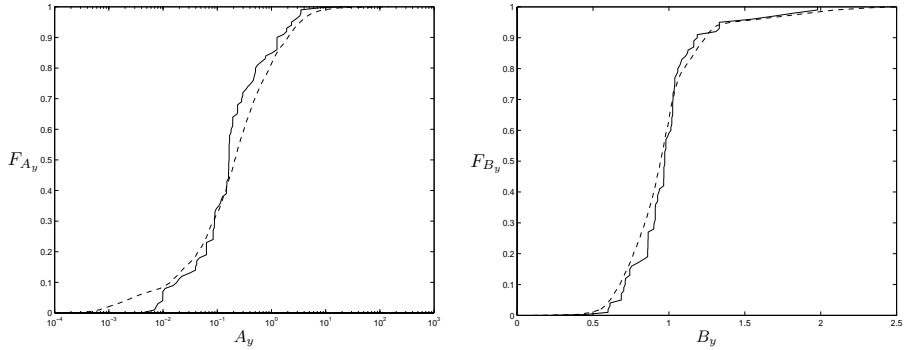


Figure 1.21: Dispersion coefficient example (stability class C): graphical comparison between marginal distributions of A_y (left) and B_y (right) from Original (-) and Extension (- -) for cutoff=-9.

Systemic retention of Sr in the human body

Figure 1.22 shows the acyclic compartmental model which is used to determine the retention of Sr in various parts of the human body.

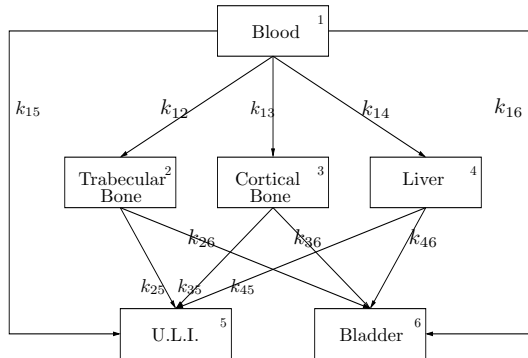


Figure 1.22: Systemic retention of Sr: Compartmental model.

The aim is to determine a distribution on the transfer coefficients k_{ij} . The transfer coefficients are considered to be non-measurable and hence probabilistic inversion was used to determine the distribution on the transfer coefficients. For a more detailed description of the Systemic retention of Sr in the human body see Appendix B.4. This PI problem was considered to be complex, because the number of target variables and information on the elicitation variables was large. In order to keep the PI problem of systemic retention of Sr in the human body computational tractable the background measure on the target variables k_{ij} was set to uniform.

Given the uniform background measure, determining the minimum relative information distribution is equivalent to determining the maximum entropy distribution, see Section 1.3.1. Therefore, in this section the value of the objective function is expressed in terms of entropy.

The following situations were considered:

Original : The information is based on the PI problem for Sr as described in Appendix B Section B.4 and is the same as presented in the column headed by `lsqcurvefit` of Table B.9. Entropy value 5.03.

Run A : The PI problem for Sr is solved using the iterative version of PREJUDICE. The number of potentially observable scenarios in **Step 1** equaled 1000. In order to keep the problem computationally tractable, it was stipulated that an observable hypercube cannot be visited more than once and the cutoff was set to -2. Under these settings, the number of iterations equaled 179. The total number of samples equaled $1.395e+5$. Entropy value 5.49.

Run B : Similar to **Run A**, however the number of iterations equaled 285. The reason why **Run B** takes more iterations than **Run A** is that the set of potentially observable scenarios at the first iteration is not the same for both runs. The total number of samples equaled $1.925e+5$ samples. Entropy value 5.47.

Extension : Extension of the PI problem as described under **Original**. The results listed under **Extension** are based on applying the iterative solution scheme to the PI problem as described under **Original**. Like for **Run A** and **Run B** it was stipulated that a particular hypercubes could not be visited more than once and the cutoff was set to -2; 488 iterations were needed. The total number of samples equaled $4.40e+5$ samples. Entropy value 5.42.

Figure 1.23 presents the relationship between entropy and the number of hypercubes covered. Based on the horizontal lines the different entropy values can be compared graphically easily.

Based on the information contained in Figure 1.23 it is observed that the entropy values of Run A, Run B and Extension are close to each other, however the number of hypercubes covered by Extension is far more than the number of hypercubes covered by Run A and Run B. It seems that Run A and Run B are most efficient; starting with a small number of potentially observable scenarios and extending it, will result in comparable entropy values and with fewer hypercubes covered.

Table 1.8 presents quantile information on the target variables for the situations considered, which is graphically presented in Figure 1.24. Tables 1.9 through 1.11 give the Spearman's rank correlations matrices.

Comparing the quantile information of marginal distributions headed by Original and Extension in Table 1.8, it is observed that the 95%-iles of

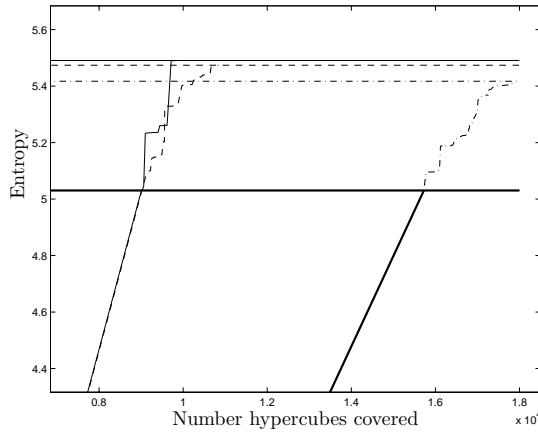


Figure 1.23: *Systemic retention of Sr: graphical display of entropy versus the number of hypercubes covered. Original (—), Run A (---), Run B (- · -) and Extension (- · - · -).*

	Quantile	Original	Run A	Run B	Extension
k_{13}	5%	1.57e-1	1.76e-1	1.59e-1	1.96e-1
	50%	4.23e-1	6.80e-1	6.77e-1	5.12e-1
	95%	1.21	2.25	2.58	4.58
k_{14}	5%	1.21e-3	3.60e-3	3.70e-3	1.70e-3
	50%	1.19e-2	1.29e-2	1.67e-2	1.72e-2
	95%	3.32e-1	1.52e-1	1.19e-1	2.81e-1
k_{15}	5%	2.44e-1	3.07e-1	2.40e-1	2.22e-1
	50%	5.59e-1	7.35e-1	7.17e-1	6.13e-1
	95%	9.07e-1	2.16	1.76	2.49
k_{25}	5%	5.83e-5	3.40e-4	3.30e-5	7.45e-5
	50%	2.49e-3	4.89e-3	6.50e-3	5.79e-3
	95%	1.17e-1	7.66e+1	6.60e+1	6.21e+1
k_{35}	5%	4.28e-6	1.06e-5	6.64e-6	4.17e-6
	50%	3.08e-5	3.14e-5	3.12e-5	3.38e-5
	95%	9.13e-5	8.08e-5	9.00e-5	9.90e-5
k_{45}	5%	1.02e-5	4.43e-12	1.26e-9	5.13e-6
	50%	4.47e-5	4.57e-5	4.46e-5	4.47e-5
	95%	9.68e-5	9.74e-5	1.05e-4	1.00e-4

Table 1.8: *Systemic retention of Sr: quantile information on target variables for Original, Run A, Run B and Extension (1/d.).*

marginal distributions of k_{13} , k_{15} and k_{25} have changed. The 5% and 50%-iles of these marginal distributions hardly changed. The difference among Original and Extension for the marginal distributions k_{14} , k_{35} and k_{45} is minor.

Comparing the quantile information of marginal distributions headed by Run A, Run B and Extension in Table 1.8, it is concluded that the

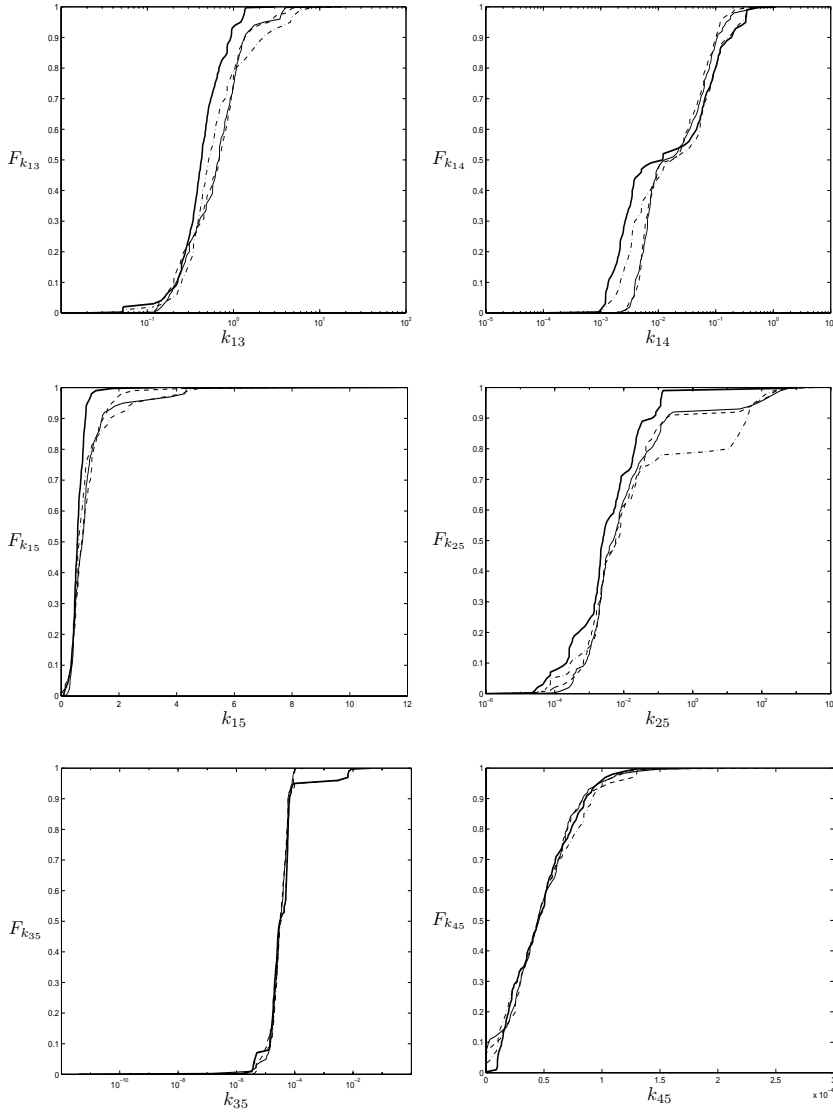


Figure 1.24: *Systemic retention of Sr: graphical display of distribution of k_{13} (top-left), k_{14} (top-right), k_{15} (middle-left), k_{25} (middle-right), k_{35} (below-left) and k_{45} (below-right). Original (—), Run A (---), Run B (· · ·) and Extension (- · -) .*

marginal distributions obtained from Run A and Run B are similar, even though $M^{(1)}$ for Run A is likely to be different from $M^{(1)}$ for Run B. In comparing the marginal distributions of Run A and Run B to Extension it

$$\begin{array}{l}
 k_{13} \\
 k_{14} \\
 k_{15} \\
 k_{25} \\
 k_{35} \\
 k_{45} \\
 tc
 \end{array}
 \left(
 \begin{array}{ccccccc}
 1 & -5.21e-1 & 5.00e-1 & 4.23e-1 & -1.30e-1 & -4.08e-1 & -2.94e-1 \\
 -5.21e-1 & 1 & -2.61e-1 & -1.12e-1 & 9.29e-2 & 3.41e-1 & 2.79e-1 \\
 5.00e-1 & -2.61e-1 & 1 & 3.76e-2 & -8.30e-2 & -1.62e-1 & -3.86e-2 \\
 4.23e-1 & -1.12e-1 & 3.76e-2 & 1 & 5.00e-3 & -2.43e-1 & -2.09e-1 \\
 -1.30e-1 & 9.29e-2 & -8.30e-2 & 5.00e-3 & 1 & 4.58e-1 & 7.18e-2 \\
 -4.08e-1 & 3.41e-1 & -1.62e-1 & -2.43e-1 & 4.58e-1 & 1 & 1.65e-1 \\
 -2.94e-1 & 2.79e-1 & -3.86e-2 & -2.09e-1 & 7.18e-2 & 1.65e-1 & 1
 \end{array}
 \right)$$

Table 1.9: *Systemic retention of Sr: Spearman's rank correlation matrix among target variables from Run A.*

$$\begin{array}{l}
 k_{13} \\
 k_{14} \\
 k_{15} \\
 k_{25} \\
 k_{35} \\
 k_{45} \\
 tc
 \end{array}
 \left(
 \begin{array}{ccccccc}
 1 & -5.42e-1 & 4.75e-1 & 4.11e-1 & -8.50e-2 & -3.35e-1 & -3.04e-1 \\
 -5.42e-1 & 1 & -3.51e-1 & -2.07e-1 & 2.03e-2 & 3.60e-1 & 9.01e-2 \\
 4.75e-1 & -3.51e-1 & 1 & -3.33e-2 & -1.87e-1 & -2.93e-1 & -1.00e-1 \\
 4.11e-1 & -2.07e-1 & -3.33e-2 & 1 & 1.71e-2 & -9.39e-2 & -4.29e-2 \\
 -8.50e-2 & 2.03e-2 & -1.87e-1 & 1.71e-2 & 1 & 4.79e-1 & -1.01e-1 \\
 -3.35e-1 & 3.60e-1 & -2.93e-1 & -9.39e-2 & 4.79e-1 & 1 & 1.20e-1 \\
 -3.04e-1 & 9.01e-2 & -1.00e-1 & -4.29e-2 & -1.01e-1 & 1.20e-1 & 1
 \end{array}
 \right)$$

Table 1.10: *Systemic retention of Sr: Spearman's rank correlation matrix among target variables from Run B.*

is concluded that the marginal distributions of Extension are quite similar to the marginal distributions obtained from Run A and Run B.

Except for a few instances the rank correlations between the marginal distribution of Run A and Run B are similar. Between the rank correlations of Run A, Run B and Extension the main difference are the rank correlations among the transfer coefficients (k_{13}, k_{14}) and (k_{14}, k_{15}) ; the sign of these rank correlations is different for Extension then for Run A and Run B. The remaining rank correlations are quite similar.

A number of rank correlations have changed when comparing the rank correlation matrices of Original (Table B.10) and Extension (Table 1.11); the rank correlations between (k_{13}, k_{45}) , (k_{13}, tc) , (k_{14}, tc) , (k_{15}, k_{25}) , (k_{15}, tc) , (k_{15}, k_{45}) , (k_{25}, tc) and (k_{45}, tc) are considered to have changed significant.

$$\begin{array}{l}
k_{13} \\
k_{14} \\
k_{15} \\
k_{25} \\
k_{35} \\
k_{45} \\
tc
\end{array}
\begin{pmatrix}
1 & 3.21e-1 & 5.92e-1 & 5.78e-1 & -9.09e-2 & -2.81e-1 & -3.77e-1 \\
3.21e-1 & 1 & 2.32e-1 & -1.63e-1 & -6.82e-2 & 3.16e-1 & 1.10e-1 \\
5.92e-1 & 2.32e-1 & 1 & 3.39e-1 & -2.80e-1 & -3.33e-1 & -3.02e-1 \\
5.78e-1 & -1.63e-1 & 3.39e-1 & 1 & -1.75e-1 & -4.69e-1 & -2.71e-1 \\
-9.09e-2 & -6.82e-2 & -2.80e-1 & -1.75e-1 & 1 & 3.80e-1 & -1.85e-1 \\
-2.81e-1 & 3.16e-1 & -3.33e-1 & -4.69e-1 & 3.80e-1 & 1 & 1.96e-1 \\
-3.77e-1 & 1.10e-1 & -3.02e-1 & -2.71e-1 & -1.85e-1 & 1.96e-1 & 1
\end{pmatrix}$$

Table 1.11: *Systemic retention of Sr: Spearman's rank correlation matrix among target variables from Extension .*

1.5 CP solvers

In this section different optimization solvers are used to solve either the primal or dual formulations of a PI problem. As stated before, both the primal and dual formulation are convex optimization problems. A wide range of optimization solvers are available which are capable of dealing with convex optimization problems.

Just by looking at any primal formulation of a PI problem, it is obvious that the solution is near the origin, because the optimization variables have to be greater than 0 and have to sum up to 1. Searching for the optimal solution along the feasible region will lead to numerical problems, since the gradient of the relative information function is not defined whenever f_k equals 0. One might argue that this problem could be solved by specifying a lower bound for f_k , but by specifying a lower bound on f_k the primal formulation of the PI problem might become infeasible, although the problem may be feasible. Searching for the optimal solution from within the feasible region will avoid such problems. Therefore it is recommended to use interior point methods whenever solving the primal formulation of a PI problem. For a detailed description on interior point methods, see [45].

In case of the dual formulation of the PI problem, the number of optimization variables is far less than the optimization variables in the primal formulation. Even though the number of optimization variables for the dual formulation is not large, the number of columns contained in the constraint matrix may be large. Hence, solving the dual formulation may be a time consuming task after all. Unlike for the optimization variables of the primal formulation, it is impossible to put restrictions on the range of the optimization variables of the dual formulation. Interior point methods may be used to solve the dual formulation, but non-interior point methods can be used as well.

The interior point methods used for solving the primal formulations are:

MOSEK : Commercial software package for large scale convex programming problems using interior point methods developed by Dr E.D. Andersen. (for more information www.mosek.com).

AFSNCP : Non-commercial interior point solver based on a polynomial primal-dual affine scaling algorithm for non-complementarity problems developed by Dr B. Jansen and described in detail in [32].

NUOPT : Extension for the commercial S-Plus statistical and data analysis software, which includes a primal-dual interior point method based on line search for general CP problems, for details see [48].

The non-interior point method used for solving the dual formulations is:

Minos5 : Activate set based algorithm for general NLP problems in GAMS and suitable for dealing with unconstrained NLP problems.

No computation times are given as some of the interior point methods are qualified research software and for others some external processing had to be done to derive the information as listed in Table 1.12. At this point, the optimization solvers MOSEK and Minos5 can be applied to PI problem with a large number of optimization variables, whereas S-Plus and AFSNCP are capable of dealing with a small number of optimization variables; S-plus and AFSNCP remain computationally tractable for problems involving around 1000 and 3000 optimization variables respectively.

The optimization routines will be applied to the dispersion coefficient example (A_y log-uniform), Lung morbidity and Systemic retention of Sr in the human body. Lung morbidity and Systemic retention of Sr in the human body are described in more detail in Appendix B.

For the dispersion coefficient example, the PI problem as introduced in Section 1.3.2 was used. The number of optimization variables equaled 900.

Before presenting the results of the different optimization routines, it must be stressed that the optimization solver are applied to the same PI problem, i.e. the samples taken from domain M , which make up the PI problem, are the same.

The results listed in Table 1.12 indicate that the AFSNCP, S-Plus and Minos5 routine are very similar and that the MOSEK results are somewhat out of line with them. Looking at the relative information values for the Lung morbidity and Systemic retention of Sr in Table 1.13 then MOSEK and Minos5 perform better than AFSNCP and S-Plus. Furthermore, AFSNCP and S-plus are suitable for small problems only, i.e. the constraint matrix consists of a relatively small number of columns. For larger problems, like the Systemic retention of Sr, these interior point methods are far more time consuming than MOSEK and Minos5.

In conclusion, for small problems like the dispersion coefficient example the difference between the optimization routines is considered to be minor.

	Quantile	Optimization Routines			
		MOSEK	AFSNCP	S-Plus	Minos5
A_y	5%	9.85e-3	1.01e-2	9.91e-3	9.91e-3
	50%	1.65e-1	1.69e-1	1.69e-1	1.69e-1
	95%	2.35	3.33	3.33	3.33
B_y	5%	6.86e-1	6.11e-1	6.11e-1	6.11e-1
	50%	9.73e-1	9.34e-1	9.34e-1	9.34e-1
	95%	1.33	1.33	1.33	1.33
ρ_{A_y, B_y}		-8.03e-1	-8.18e-1	-8.16e-1	-8.13e-1
Relative Information		3.31	3.21	3.20	3.20

Table 1.12: *Dispersion coefficient example (stability class C): comparison of quantile information on marginal distributions of A_y and B_y for different optimization solvers.*

	Number of opt. variables	Relative information			
		MOSEK	AFSNCP	S-Plus	Minos5
Lung morbidity	964	3.98	4.30	4.17	3.97
System Retention of Sr	15725	5.66	N/A	N/A	5.67

Table 1.13: *Relative information values for Lung morbidity and Systemic retention of Sr for different NLP solvers.*

For PI problems involving more optimization variables, like Lung morbidity, the difference in relative information value between MOSEK and Minos5 on the one hand and AFSNCP and S-plus on the other hand, tends to become apparent. For large PI problems, like the Systemic retention of Sr in the human body, the relative information values of MOSEK and Minos5 are still very close.

1.6 Efficient version of PREJUDICE

The main problem in solving the primal formulation of a PI problem concerns the number of optimization variables being equal to number of samples taken from the domain. As stated previously, the number of optimization variables in the dual formulation equals the number of interquantile intervals $(K + 1)$ multiplied by the number of elicitation variables n . In general the number of optimization variables for the dual formulation is far less than the number of optimization variables for the primal formulation, however, the number of columns contained in the constraint matrix will be equal to the number of samples taken from M . Hence, solving the dual formulation may be a time consuming task after all.

Section 1.3.1 discussed the implications of the uniform background measure and it turned out that the number of optimization variables was equal to the number of observable hypercubes covered rather than the total num-

ber of samples drawn. In this section a PI problem formulation is given which depend on the number of observable hypercubes covered for *any* background measure on the target variables. Corollary 1.1.10 is used to deal with PI problems involving a large number of samples.

Recall from Section 1.1.2 that ν_k and γ_k denote the push-forward measures of ν and γ onto $\{1, \dots, k\}$, respectively, with $k = (K + 1)^n$. Based on Corollary 1.1.10, the strategy to solve a PI problem is to determine measure ν_k which minimizes $I(\nu_k|\gamma_k)$ and determine $\tilde{\mu}_k$ based on ν_k .

The PI problem formulation based on ν_k and γ_k is

$$\begin{aligned} \min_{g_k} \quad & \sum_l g_k(l) \log \frac{g_k(l)}{g_{\gamma,k}(l)} \\ \text{s.t.} \quad & \sum_{\mathbf{x} \in M} \sum_l g_k(l) 1_{A_1}(\mathbf{x}) = \nu(I_{j,l_j}) \quad j = 1, \dots, n \\ & g_k \geq 0. \end{aligned} \quad (1.52)$$

where $A_1 = \{T_{(j)}(\mathbf{x}) \in B_l, B_l \cap I_{j,l_j} \neq \emptyset\}$, $g_{\gamma,k}(l) := \sum_{\mathbf{x} \in M} f_\lambda(\mathbf{x}) 1_{\{T_k(\mathbf{x})=l\}}(\mathbf{x})$ and I_{j,l_j} the l_j -th interquantile interval of elicitation variable Y_j , with $j = 1, \dots, n$ and $l_j = 1, \dots, K + 1$.

Let probability mass function \tilde{f}_k on sample $\mathbf{x} \in M$ be defined as

$$\tilde{f}_k(\mathbf{x}) := f_\lambda(\mathbf{x}) \sum_{l=1}^{(K+1)^n} \frac{g_k(l)}{g_{\gamma,k}(l)} 1_{\{T_k(\mathbf{x})=l\}}(\mathbf{x}) \quad (1.53)$$

Corollary 1.1.10 states in case of minimum relative information measure $\tilde{\mu}_k$, the relative information value $I(\tilde{\mu}_k|\lambda)$ in the target variable space equals the relative information value $I(\nu_k|\gamma_k)$ in the observable space. Briefly, the efficient version of PREJUDICE determines a distribution on the hypercubes covered which has minimum relative informative with respect to the push-forward of the background measure in the target variable space. This push-forward background measure is the main difference with PI problem formulation 1.48; for PI problem formulation 1.48 the background distribution (i.e. the uniform background distribution) is specified in the observable space rather than the target variable space.

The difference between PREJUDICE and the efficient version of PREJUDICE is illustrated using the dispersion coefficient example from Section 1.4.2 referred to as Extension. Both PI problem formulations are based on the same set of samples taken from domain M and the background distribution for A_y is log-uniform and for B_y uniform. Table 1.14 lists information of the distribution on the target variables.

As expected the results are the same, except for the number of optimization variables which is reduced dramatically from 62280 to 78. Since the number of optimization variables has dropped, the optimization routines as introduced in Section 1.5 (which were qualified as suitable for small problems only) can be used to determine the distribution as well. There is some variation in the results, but it is considered to be little.

	Quantile	Extension	Efficient
A_y	5%	2.74e-3	2.74e-3
	50%	2.10e-1	2.10e-1
	95%	3.39	3.39
B_y	5%	6.12e-1	6.12e-1
	50%	9.46e-1	9.46e-1
	95%	1.40	1.40
ρ_{A_y, B_y}		-9.08e-1	-9.08e-1
Number of optimization variables		62280	78
Relative Information		4.61	4.61

Table 1.14: *Dispersion coefficient example (stability class C): comparison of quantile information on marginal distributions of A_y and B_y between PREJUDICE (Extension) and the efficient version of PREJUDICE.*

	Quantile	Optimization Routines			
		MOSEK	AFSNCP	S-Plus	Minos5
A_y	5%	2.74e-3	2.90e-3	2.74e-3	2.77e-3
	50%	2.10e-1	2.16e-1	2.10e-1	2.10e-1
	95%	3.39	3.60	3.39	3.39
B_y	5%	6.12e-1	6.02e-1	6.12e-1	6.12e-1
	50%	9.46e-1	9.42e-1	9.46e-1	9.46e-1
	95%	1.40	1.40	1.40	1.40
ρ_{A_y, B_y}		-9.08e-1	-9.19e-1	-9.08e-1	-9.15e-1
Relative Information		4.61	4.63	4.61	4.61

Table 1.15: *Dispersion coefficient example (stability class C): comparison of quantile information on marginal distributions of A_y and B_y for different optimization solvers for the Efficient version of PREJUDICE.*

Clearly, the benefits of the efficient version of PREJUDICE are apparent; solving a PI problem with a ‘suitable’ background distribution on the target variables and a manageable number of optimization variables is both from conceptual and computational point of view very attractive. Furthermore the iterative version of PREJUDICE as described in Section 1.4 applies to the efficient version as well.

Finally, note that the uncertainty analyst may decide in which space the background measure is specified; either in the target variable space or observable space. Here it is recommended to specify the background measure in the target variable space, since the resulting measure will be minimum relative informative with respect to the specified background measure. It is not obvious what the background measure in the target variable space is when a background measure in the observable space is specified; to which background measure is the resulting measure minimal relative informative?

1.7 Conclusions

Due to the Structured Expert Judgment Elicitation Methodology or experimental data it may happen that the information available on variable \mathbf{Y} has to be translated/inverted into information on the variable \mathbf{X} using the mapping T . This translation can be seen as to invert T with respect to \mathbf{Y} to obtain \mathbf{X} . Since \mathbf{Y} is uncertain, this translation is called probabilistic inversion.

In cases where partial information is available, the concept of relative information was used to determine the unspecified information. The unspecified information was determined in such a way that the resulting information differed minimally compared to the partial information when measured using relative information.

It has been shown that probabilistic inversion has a sound theoretical foundation and that it is practical to implement; two implementations PARFUM and PREJUDICE were discussed. Based on results and conceptual reasons, PREJUDICE was qualified better than PARFUM (Section 1.2.3).

Roughly, probabilistic inversion involves two steps:

Step 1 Determination of the domain M in the target variable space.

Step 2 Determination of a measure on M .

Step 1 is of crucial importance for the success of probabilistic inversion. The determination of M is driven largely by heuristics which include the information on the elicitation variables, mapping T and the physics underlying the problem. These three elements meet when minimizing Expression 1.14 or 1.19. The resulting values for the target variables are termed model inversions. It is recommended that the number of elicitation variables is larger than the number of target variables. Since more target variables than elicitation variables will most likely result in multiple combinations of model inversions which minimize Expression 1.14 or 1.19.

The measure on M will be determined such that it has minimum relative information with respect to a certain background measure and that the marginal distributions of its push-forward comply with the marginal distributions of the DM; a constrained convex optimization problem. Since the determination of the distribution on M is done with respect to a background distribution, the specification of this background distribution is very important. Section 1.3 provides guidelines on how to construct an appropriate background measure and reports on the computational advantages of the uniform background measure.

The aim of probabilistic inversion is to determine a domain M which is observationally complete, meaning every potentially observable scenario N can be reproduced by $T(M)$. However, in the implementation a selection of potentially observable scenarios are used in **Step 1** to determine M , which in turn will make it unlikely that M is observationally complete. An iterative version of PREJUDICE is introduced in Section 1.4, which concentrates

on checking observable hypercubes which have not been covered so far. The results between PREJUDICE and the iterative version of PREJUDICE are compared for two examples, which clearly show the beneficial effect of the iterative version of PREJUDICE.

The primal and dual formulation of a PI problem are convex optimization problems. In Section 1.5 the effect of different optimization solvers capable of dealing with constrained convex optimization problems are investigated. When solving the primal formulation, it is strongly recommended to use interior point methods, since it is clear that the solution is near the origin and non-interior point methods will search along the feasible region which most likely will lead to numerical problems since the gradient of the relative information function is not defined whenever an optimization variable has a value equal or very close to 0. However, non-interior point methods may be used in solving the dual formulation of the PI problem. For problems involving a small number of optimization variables the solution and computation time of the different solvers are similar. For problems involving a large number of optimization variables it is recommended to use MOSEK. MOSEK is capable of solving large problems very quickly and delivers, besides the primal solution, the dual solution as well, which can be used in the iterative version of PREJUDICE.

When solving the primal formulation using a non-uniform background measure, the number of optimization variables is equal to the number of samples. As indicated in Section 1.5, the dual formulation involves a small number of optimization variables but still it may be a time consuming task after all. In Section 1.3.1 it was observed that the number of optimization variables in case of the uniform background measure was equal to the number of observable hypercubes covered. However the uniform background measure may be unsuitable for many problems. Based on Corollary 1.1.10 an efficient version of PREJUDICE has been formulated in Section 1.6 for which the number of optimization variables is equal to the number of hypercubes covered for any background measure. For the dispersion coefficient example the efficient version resulted in the same results as the iterative version, only it used far less optimization variables, which is attractive from a computational point of view.

Chapter 2

Dependencies

It has long been known that significant errors in uncertainty analysis can be caused by ignoring dependencies between uncertainties [1],[35]. In this chapter the use of dependencies in uncertainty analysis is addressed. Section 2.1 discusses different dependence elicitation techniques and considerations which have to be taken into account when applied in expert judgment studies. The use of dependencies in combining expert judgments is treated in Section 2.2. Finally, Section 2.3 investigates the use of dependencies in probabilistic inversion.

2.1 Dependence elicitation techniques

In the framework of the Structured Expert Judgment Elicitation Methodology, the best source of information about dependencies is often the experts themselves. The most thorough approach would be to elicit directly the expert's joint distributions. As stated in [34], one obvious strategy would be to ask experts directly to assess a (rank) correlation coefficient. Even trained statisticians have difficulty with this type of assessment [22], however in [5] it is argued that the most accurate way to obtain subjective dependence is simply to query experts on the rank correlation between two variables; the experts in [5] were MBA students who had taken a course in statistics.

2.1.1 Overview literature

It is stated in [6] that we should not ask ourselves whether experts can assess dependencies accurately, but to question if experts can assess dependencies well enough to be useful in an analysis. Although the results as presented in [6] are very preliminary, an affirmative to the latter question is suggested.

Roughly, the dependence elicitation techniques can be subdivided into three categories:

Statistical Approaches: two approaches are suggested, see [5], [6];

- **Strength of relationship:** the strength of relationship S_{Y_1, Y_2} between elicitation variables Y_1 and Y_2 is measured on a scale from 1 to 7, where 1 represents ‘very strong negative relationship’, 4 represents ‘no relationship’ and 7 represents ‘very strong positive relationship’. This method to query dependence is the least rigorous, but may serve as a starting point to think about the dependence between two variables. In [5] the value of S_{Y_1, Y_2} is transformed to a Spearman’s ρ_{Y_1, Y_2} using: $\rho_{Y_1, Y_2} = \frac{S_{Y_1, Y_2} - 4}{3}$
- **Spearman’s correlation:** the Spearman’s ρ_{Y_1, Y_2} is defined on the $[-1, 1]$ interval, with -1 expressing Y_1 and Y_2 ‘very strong negatively correlated’, 0 expressing Y_1 and Y_2 ‘uncorrelated’ and 1 expressing ‘very strong positively correlated’. Assessing ρ_{Y_1, Y_2} requires thorough knowledge of statistics.

Probability: in [5], [6] 3 types of probabilities are introduced which may be used to query dependence.

- **Probability of concordance:** The probability of concordance P_C between Y_1 and Y_2 is defined, based on n samples from (Y_1, Y_2) , as

$$P_C := \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n 1_{C^*}((y_{1,i}, y_{2,i}), (y_{1,j}, y_{2,j}))}{\binom{n}{2}}$$

where $C^* = \{(y_{1,i} - y_{1,j})(y_{2,i} - y_{2,j}) > 0\}$. A value for P_C close to 0.5 indicates ‘no relationship’, whereas a value of 0 expresses ‘very strong negative relationship’ and 1 expresses ‘very strong positive relationship’. P_C relates to Kendall’s τ by

$$\tau = 2 P_C - 1. \quad (2.1)$$

In [5] it is assumed that the distribution on Y_1 and Y_2 can be well approximated by a bivariate normal distribution, from which the following relationships between the product moment (Pearson) correlation ρ_{Y_1, Y_2}^* , Spearman’s ρ_{Y_1, Y_2} and Kendall’s τ can be established, [39] :

$$\rho_{Y_1, Y_2}^* = 2 \sin \left(\frac{\pi \rho_{Y_1, Y_2}}{6} \right) \quad (2.2)$$

$$\rho_{Y_1, Y_2}^* = \sin \left(\frac{\pi \tau}{2} \right) \quad (2.3)$$

The relationship between ρ_{Y_1, Y_2}^* and Spearman’s ρ_{Y_1, Y_2} as given in Equation 2.2 is studied in some more detail in [8]; given a Spearman’s rank correlation matrix it is not guaranteed that the product moment correlation matrix obtained from Equation

2.2 is positive definite. It must be stressed that there may be instances that the assumption that the distribution on (Y_1, Y_2) is well approximated using a bivariate normal distribution, is *not* valid. Another way to construct a joint distribution on (Y_1, Y_2) is to determine the minimum relative information distribution with respect to the product distribution of Y_1 and Y_2 and Spearman's ρ_{Y_1, Y_2} , see [11], [43]. Using this representation of the distribution on (Y_1, Y_2) , a numerical relationship between Kendall's τ and Spearman's ρ_{Y_1, Y_2} can be derived as well.

- **Joint probability:** The joint probability distribution is defined as $P_{JP}(y_1, y_2) := P(Y_1 \leq y_1, Y_2 \leq y_2)$. In case the expert regards Y_1 and Y_2 as independent, he/she should assign a probability of $P_{JP} = F_{Y_1}(y_1) F_{Y_2}(y_2)$, where F_{Y_i} represents the continuous cumulative distribution function of elicitation variable Y_i , $i = 1, 2$. A 'very strong negative relationship' is expressed by P_{JP} close to zero, whereas a 'very strong positive relationship' is expressed by $P_{JP} = F_{Y_1}(y_1)$, in the special case $F_{Y_1}(y_1) = F_{Y_2}(y_2)$. Expressing dependence by means of P_{JP} is considered to be very difficult for experts. It requires from the experts reasonable knowledge on probability theory, which in many situations is not available. By making assumptions on the distribution of (Y_1, Y_2) , it is possible to transform P_{JP} into a Spearman's ρ_{Y_1, Y_2} .
- **Conditional probability:** The conditional probability distribution of Y_1 given Y_2 is defined as $P_{CP}(y_1, y_2) := P(Y_1 \leq y_1 | Y_2 \leq y_2)$. In case the expert regards Y_1 and Y_2 as independent $P_{CP} = P(Y_1 \leq y_1)$, whereas a 'negative relationship' between Y_1 and Y_2 is expressed through $P_{CP} \in [0, P(Y_1 \leq y_1))$ and a 'positive relationship' through $P_{CP} \in (P(Y_1 \leq y_1), 1]$. Conditioning on an interval instead of a fractile or realization, is considered in [6] to be a potential difficulty in providing assessments for P_{CP} . Like for the joint probability, it is possible to transform P_{CP} into a Spearman's ρ_{Y_1, Y_2} by making assumptions on the distribution of (Y_1, Y_2) .

Conditional quantile estimates: for this type of dependence elicitation the experts is given the information that Y_2 takes the value y_2 which corresponds to a certain quantile of the distribution of Y_2 , next the expert is queried on his/her expected quantile of Y_1 which corresponds to the situation where $Y_2 = y_2$; since there may be many realizations y_1 of Y_1 possible, the conditional expected quantile estimate required $E(F_{Y_1}(y_1) | Y_2 = y_2)$. The conditional quantile estimate is related to Spearman's ρ_{Y_1, Y_2} by the standard nonparametric regression representation

$$E(F_{Y_1}(y_1) | Y_2 = y_2) = \rho_{Y_1, Y_2}(F_{Y_2}(y_2) - 0.5) + 0.5. \quad (2.4)$$

Based on this representation it can be easily seen that the conditional quantile estimate should satisfy

$$\mu_{\min} \leq E(F_{Y_1}(y_1)|Y_2 = y_2) \leq \mu_{\max} \quad (2.5)$$

with $\mu_{\min} = \min\{F_{Y_2}(y_2), 1 - F_{Y_2}(y_2)\}$ and $\mu_{\max} = \max\{F_{Y_2}(y_2), 1 - F_{Y_2}(y_2)\}$; the min and max operations take account of the position of the percentile of y_2 with respect to the median of Y_2 and guarantee the validity of Inequality 2.5. In case $F_{Y_2}(y_2) > 0.5$, values of $E(F_{Y_1}(y_1)|Y_2 = y_2)$ close to the left-hand side of Inequality 2.5 expresses a 'very strong negative relationship' and values close to the right-hand side of Inequality 2.5 expresses a 'very strong positive relationship' between Y_1 and Y_2 . In case of 'no relationship' the value of $E(F_{Y_1}(y_1)|Y_2 = y_2)$ should be equal to 0.5.

For all techniques, the experts require some training; for strength of relationship, correlation coefficient and conditional quantile estimates, the experts need to have a good understanding of statistical concepts; weak vs. strong relationship, correlation and quantile, respectively. In case of querying probabilities the training seems to be less technical; understanding of the events and getting a feel of quantifying the dependence in terms of probability. Finally, probabilities can easily be interpreted in terms of frequencies in contrast of strength of relationship, correlation coefficient and conditional quantile estimates. It is suggested that the frequency interpretation would make the expert less susceptible to cognitive biases.

Based on these considerations the probability dependence elicitation technique was chosen for the Structured Expert Judgment Elicitation Methodology.

2.1.2 Dependence Elicitation and Structured Expert Judgment Elicitation Methodology

Under the Structured Expert Judgment Elicitation Methodology, experts have to quantify their knowledge for events which are potentially measurable and with which they are familiar, hence it seems natural to quantify dependence among similar events as well. Therefore dependence questions have to be formulated among relevant pairs of events and in line with the Structured Expert Judgment Elicitation Methodology and, additionally, the dependence elicitation technique should be easy to understand for the experts. In performing an expert judgment study, experts will have only a limited amount of time to quantify their knowledge. They should focus firstly on the determination of the distributions for the elicitation variables and secondly on quantifying the dependence among elicitation variables; ignoring dependence may lead to significant errors in doing uncertainty analysis, but even greater errors will arise in case the distributions over the model input parameters are inaccurate.

The elicitation of dependence in the Joint CEC/USNRC Uncertainty Analysis was done via conditional probability. Querying dependence via joint probability was considered to be too difficult and specification of the probability of concordance for among independent samples too technical. The conditional probability dependence elicitation technique, however, is flexible, easy to communicate to the experts and the type of dependence elicitation questions can be easily formulated to satisfy the Structured Expert Judgment Elicitation Methodology.

Experts assess the marginal distributions for (continuous) elicitation variables Y_1 and Y_2 first. Next they are asked the question:

Consider an experiment involving both Y_1 and Y_2 . Suppose Y_2 were observed and its value was found to lie above your median value for Y_2 ; what's your probability that, in this same experiment, the value for Y_1 would also lie above your median value for Y_1 ?

Experts in the joint CEC/USNRC uncertainty analysis quickly became comfortable with this assessment technique, acknowledged its importance and dealt with the questions in terms of possible outcomes of experiments. Most experts addressed the question in terms of frequency which, as stated earlier, makes them less susceptible to cognitive biases.

If F_{Y_1} and F_{Y_2} are the (continuous invertible) cumulative distribution functions of Y_1 and Y_2 respectively, the experts thus assess,

$$\pi_{\frac{1}{2}, \frac{1}{2}}(Y_1, Y_2) := P(F_{Y_1}(Y_1) > \frac{1}{2} \mid F_{Y_2}(Y_2) > \frac{1}{2}). \quad (2.6)$$

In case there are a large number of potential dependencies among elicitation variables, not all dependencies will be queried; firstly there would be too many questions and secondly, in eliciting and combining all dependencies, it is almost impossible to ensure that the resulting Spearman's rank correlation matrix is positive definite. Therefore it is better to query the dependence for a selection of all possible dependencies. If this selection is such that the resulting dependency graph is acyclic, it is shown in [43] that a joint distribution can be found for which:

- (i) the marginal distributions complies the expert's distributions,
- (ii) Spearman's rank correlation matrix is positive definite and satisfies the expert's information as specified in the dependency structure.

From the set of distributions, which share Properties (i) and (ii), the distribution is selected which has minimum relative information with respect to the product distribution¹, for details see [11].

¹When referring to minimum relative information distribution in this chapter, it is understood to be minimum with respect to the product distribution.

2.1.3 $\pi_{\frac{1}{2},\frac{1}{2}}(Y_1, Y_2)$ and Spearman's ρ_{Y_1, Y_2}

The dependence information over a selection of elicitation variables queried from the experts is available via conditional probabilities $\pi_{\frac{1}{2},\frac{1}{2}}(Y_1, Y_2)$. However, a relationship between $\pi_{\frac{1}{2},\frac{1}{2}}(Y_1, Y_2)$ and Spearman's ρ_{Y_1, Y_2} must be derived in order to ensure that the Spearman rank correlation matrix of the minimum relative information distribution satisfies the dependency information as specified by the experts.

Consider all distributions for (Y_1, Y_2) having marginals F_{Y_1}, F_{Y_2} , having minimum relative information with respect to the distribution with independent marginals F_{Y_1}, F_{Y_2} and having Spearman rank correlation $\rho_{Y_1, Y_2} \in [-1, 1]$. Based on simulation results the unique relationship between Spearman's ρ_{Y_1, Y_2} and $\pi_{\frac{1}{2},\frac{1}{2}}(Y_1, Y_2)$ is determined, see Figure 2.1. Hence, the conditional probability $\pi_{\frac{1}{2},\frac{1}{2}}(Y_1, Y_2)$ may be considered as a char-

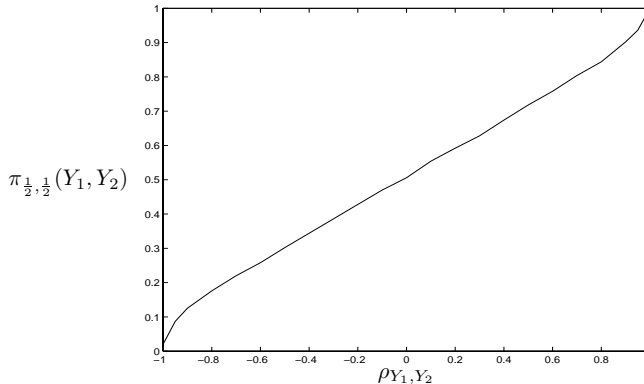


Figure 2.1: Relation between Spearman's rank correlation ρ_{Y_1, Y_2} and conditional probability $\pi_{\frac{1}{2},\frac{1}{2}}(Y_1, Y_2)$.

acterization of the minimum relative information distribution with Spearman's ρ_{Y_1, Y_2} .

This characterization is used to determine the unique relationship between Spearman's ρ_{Y_1, Y_2} and $\pi_{r_1, r_2}(Y_1, Y_2)$, $r_1, r_2 \in (0, 1)$. The simulation program UNICORN [11] is used to determine the numerical relationship between Spearman's ρ_{Y_1, Y_2} and conditional probability $\pi_{r_1, r_2}(Y_1, Y_2)$. The resulting table will be referred to as the $\rho\pi$ -table ².

²The values of r_1 and r_2 were taken from the set $\{0.0025, 0.05, 0.1, 0.2, \dots, 0.8, 0.9, 0.95, 0.9975\}$. The Spearman correlations ρ_{Y_1, Y_2} were taken from the set $\{-1, -0.9, \dots, 0.9, 1\}$. The information is stored as $[r_1, r_2, \rho_{Y_1, Y_2}, \pi_{r_1, r_2}(Y_1, Y_2)]$

2.2 Combining Expert Judgments

The current strategy for combining experts marginal distribution for elicitation variable Y , as described in [27], may be summarized as; let w_e represent the weight of the e -th expert ($e = 1, \dots, N_{\text{exp.}}$) such that $w_e \geq 0$ and $\sum_e w_e = 1$. The density of elicitation variable Y of the DM is obtained from linear pooling the densities of the experts for Y ;

$$f_{Y,\text{DM}} = \sum_{e=1}^{N_{\text{exp.}}} w_e f_{Y,e} \quad (2.7)$$

where $f_{Y,e}$ is the density associated with the e -th expert assessments for elicitation variable Y . The combination of the experts' densities is done using EXCALIBUR [14].

Two strategies will be introduced in this section which combine experts' marginal distributions and dependence information via linear pooling to obtain the information for the DM;

Strategy 1 : Information on the marginal distributions and conditional probabilities for the DM are obtained by combining firstly the experts' marginal distributions and secondly conditional probabilities using linear pooling on both occasions. Based on the information for the DM a minimum relative information distribution is determined satisfying Properties (i) and (ii) of Section 2.1.2, see [34].

Strategy 2 : For each expert, determine a minimum relative information distribution satisfying Properties (i) and (ii) of Section 2.1.2. The distribution for the DM is obtained by combining the minimum relative information distributions of experts using linear pooling.

The difference between the two strategies may be characterized as follows: the dependence information under Strategy 1 is combined after the marginal distributions are combined, whereas under Strategy 2 the marginal distributions and dependence information are combined at the same time.

2.2.1 Strategy 1

Before discussing Strategy 1 in more detail, consider the following; having queried the conditional probabilities from the experts, it is tempting to pool the conditional probabilities linearly to determine the conditional probability of the DM. However, in general the medians of the experts will be different, for this reason one *cannot* combine the conditional probabilities $\pi_{\frac{1}{2}, \frac{1}{2}}(Y_1, Y_2)$ via linear pooling; the pooling will not be over the same events.

Let $y_{1,\text{DM},50}$ and $y_{2,\text{DM},50}$ denote the medians for the DM's distribution for Y_1 and Y_2 . With each expert a minimum relative information distribution on Y_1 and Y_2 is associated; for each such distribution the conditional probabilities $P(Y_1 > y_{1,\text{DM},50} \mid Y_2 > y_{2,\text{DM},50})$ are computed. Since

these conditional probabilities are defined over the same events for all experts, they can be combined using linear pooling. This yields a value for $\pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_1, Y_2)$ for DM, for which the corresponding Spearman's $\rho_{Y_1, Y_2}^{\text{DM}}$ for the DM can be found.

Solution scheme

The various steps involved in Strategy 1 are summarized in the solution scheme below:

Step 1 : For each expert e query $\pi_{\frac{1}{2}, \frac{1}{2}}^e(Y_1, Y_2) = P(F_{Y_1, e}(Y_1) > \frac{1}{2} \mid F_{Y_2, e}(Y_2) > \frac{1}{2})$.

Step 2 : For each expert e , find Spearman's ρ_{Y_1, Y_2}^e from the $\rho\pi$ -table which passes through $(\frac{1}{2}, \frac{1}{2}, \pi_{\frac{1}{2}, \frac{1}{2}}^e(Y_1, Y_2))$.

Step 3 : Take linear pooling of experts' marginals to determine $F_{Y_1, \text{DM}}$ and $F_{Y_2, \text{DM}}$ and $y_{1, \text{DM}, 50}$, $y_{2, \text{DM}, 50}$.

Step 4 : For each expert e ;

- Determine $P(F_{Y_1, e}(Y_1) > F_{Y_1, e}(y_{1, \text{DM}, 50})) = 1 - F_{Y_1, e}(y_{1, \text{DM}, 50})$ and $P(F_{Y_2, e}(Y_2) > F_{Y_2, e}(y_{2, \text{DM}, 50})) = 1 - F_{Y_2, e}(y_{2, \text{DM}, 50})$.
- Based on $P(F_{Y_l, e}(Y_l) > F_{Y_l, e}(y_{l, \text{DM}, 50}))$ with $l = 1, 2$, determine for each Spearman's $\rho_{Y_1, Y_2} \in \{-1, -0.9, \dots, 0.9, 1\}$,

$$P_{\rho_{Y_1, Y_2}}(F_{Y_1, e}(Y_1) > F_{Y_1, e}(y_{1, \text{DM}, 50}) \mid F_{Y_2, e}(Y_2) > F_{Y_2, e}(y_{2, \text{DM}, 50}))$$

using linear interpolation, from the $\rho\pi$ -table.

- Determine

$$P_{\rho_{Y_1, Y_2}^e}(F_{Y_1, e}(Y_1) > F_{Y_1, e}(y_{1, \text{DM}, 50}) \mid F_{Y_2, e}(Y_2) > F_{Y_2, e}(y_{2, \text{DM}, 50}))$$

by linear interpolation from the

$$P_{\rho_{Y_1, Y_2}}(F_{Y_1, e}(Y_1) > F_{Y_1, e}(y_{1, \text{DM}, 50}) \mid F_{Y_2, e}(Y_2) > F_{Y_2, e}(y_{2, \text{DM}, 50}))$$

values.

Step 5 : Take linear pooling of

$$P_{\rho_{Y_1, Y_2}^e}(F_{Y_1, e}(Y_1) > F_{Y_1, e}(y_{1, \text{DM}, 50}) \mid F_{Y_2, e}(Y_2) > F_{Y_2, e}(y_{2, \text{DM}, 50}))$$

to find

$$\pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_1, Y_2) = P(F_{Y_1, \text{DM}}(Y_1) > \frac{1}{2} \mid F_{Y_2, \text{DM}}(Y_2) > \frac{1}{2}).$$

Step 6 : Find Spearman's $\rho_{Y_1, Y_2}^{\text{DM}}$ from the $\rho\pi$ table as the value which passes through $(\frac{1}{2}, \frac{1}{2}, \pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_1, Y_2))$.

Step 7 : Construct a minimum relative information distribution on the elicitation variables which satisfies Properties (i) and (ii) of Section 2.1.2 for the DM data obtained in **Step 3** and **Step 6**.

2.2.2 Strategy 2

Strategy 2 extends the current strategy of combining experts' marginal density (Expression 2.7) to combining experts' joint densities on the elicitation variables. Let $\mathbf{Y} = (Y_1, \dots, Y_n)$, then a minimum relative information distribution satisfying Properties (i) and (ii) of Section 2.1.2 is assigned to the expert's assessments of \mathbf{Y} . Hence the joint density for each expert is available. Let $f_{\mathbf{Y}, e}$ represent the joint density associated with expert e 's minimum relative information distribution on the elicitation variables \mathbf{Y} . The density of \mathbf{Y} of the DM is obtained from

$$f_{\mathbf{Y}, \text{DM}} = \sum_{e=1}^{N_{\text{exp.}}} w_e f_{\mathbf{Y}, e}. \quad (2.8)$$

The combination of the experts' joint densities is done using the simulation program UNICORN, [11].

It is assumed here that all experts provided assessments for elicitation variables \mathbf{Y} , the procedure can be adapted in case experts did not provided assessments for certain elicitation variables.

Solution scheme

The various steps involved in Strategy 2 are summarized in the solution scheme below:

Step 1 : For each expert e query

$$\pi_{\frac{1}{2}, \frac{1}{2}}^e(Y_1, Y_2) = P(F_{Y_1, e}(Y_1) > \frac{1}{2} \mid F_{Y_2, e}(Y_2) > \frac{1}{2}).$$

Step 2 : For each expert e , find Spearman's ρ_{Y_1, Y_2}^e from the $\rho\pi$ -table which passes through $(\frac{1}{2}, \frac{1}{2}, \pi_{\frac{1}{2}, \frac{1}{2}}^e(Y_1, Y_2))$.

Step 3 : For each expert, construct a minimum relative information distribution on the elicitation variables which satisfies Properties (i) and (ii) of Section 2.1.2.

Step 4 : Take linear pooling of experts' minimum relative information distributions to obtain the minimum relative information distribution for the DM.

2.2.3 Performance based weights

The question put forward in this section is ‘Do the strategies apply in case of performance based weights as well?’. For an introduction to performance based weights, see [9],[13]. The application of performance based weights (global and item) for both strategies, needs some care. In case of global weights, both Strategy 1 and 2 can be applied. Item weights differ from item to item, hence Strategy 2 can be applied but Strategy 1 cannot be applied, since the weights to determine $F_{Y_1,DM}$ will be different from the weights to determine $F_{Y_2,DM}$, hence it is unclear which set of weights to use in **Step 5** of the solution scheme of Strategy 1.

2.2.4 Examples

The similarities and differences between Strategies 1 and 2 will be illustrated by two examples. The elicitation variable is a target variable in the first example, whereas the elicitation variable is *not* a target variable in the second example.

Example 1 : Elicitation variable is target variable; the example is taken from the Foodchain panel [26] and deals with data on the daily intake of dairy cows (D) and beef cattle (B) for different feedstuff (Pasture (P), Silage (S) and Cereals (C)) both indoors (I) and outdoors (O); for example the abbreviation DPO represents the daily intake of pasture for a dairy cow grazing outside and BCI represents the daily intake of cereals for a beef cow eating indoors.

The individual expert assessments on conditional probabilities and corresponding Spearman’s rank correlations are given in Appendix E in Table E.1 and E.2, respectively. Note the difference between the conditional probability assessments among some experts, especially Expert 3 and Expert 4.

The marginal distribution for the DM as determined by Strategy 1 and Strategy 2 are the same. It is more interesting to look at the Spearman’s rank correlation matrices for the DM, since dependencies are combined at different stages under both strategies.

The difference among the Spearman’s rank correlation matrices resulting from both strategies is illustrated graphically using so-called ‘radar-plots’, see Figure 2.2. Each radar-plot presents graphically the Spearman’s rank correlation with respect to a certain variable resulting from Strategy 1 (—) and Strategy 2 (---); the radar-plot in the upper left corner presents the Spearman’s rank correlations for variables with respect to variable DPO.

The actual Spearman’s rank correlation matrices are given in Table E.3 and E.4 of Appendix E. It is clear from the radar-plots that com-

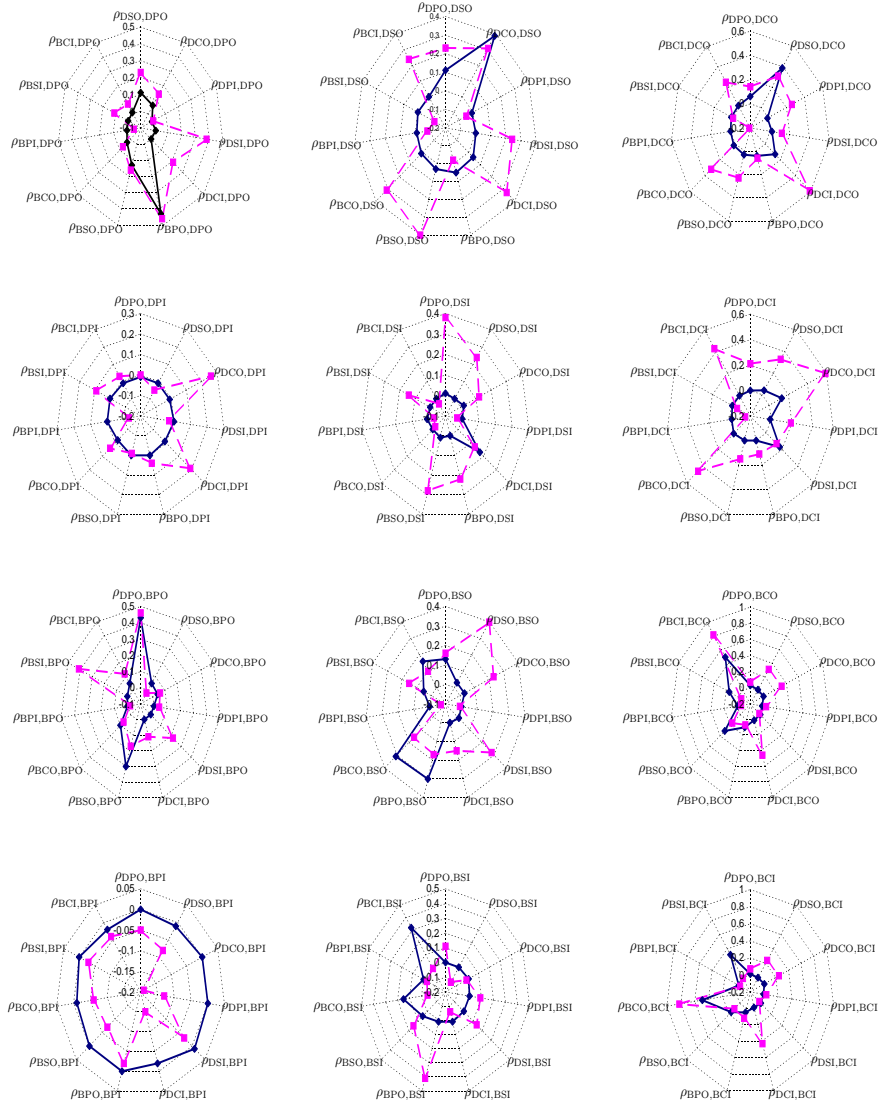


Figure 2.2: Example 1: radar plots of Spearman's rank correlations obtained from Strategy 1 (—) and Strategy 2 (---).

binning dependencies using Strategy 1 or Strategy 2 result in different dependencies for the DM.

Example 2 : Elicitation variable is not a target variable. No probabilistic inversion is required, but the distribution on the target vari-

ables is obtained by simulation. The example is taken from the Late Health effects panel [30] and deals with data concerning the number of radiation-induced cancer deaths fatalities following an exposure in a population of a hundred million persons ($5.0e+7$ male, $5.0e+7$ female) each receiving a whole body dose of 1 Gy low LET (=gamma) radiation at a uniform dose rate DR . The information on the elicitation variables is represented as:

$Y_{i,1}$: fraction of radiation induced cancer deaths up to 40 years for cancer site i following a whole body dose at *high dose rate*.

$Y_{i,2}$: fraction of radiation induced cancer deaths over a *lifetime* for cancer site i following a whole body dose at *high dose rate*.

$Y_{i,3}$: fraction of radiation induced cancer deaths up to 40 years for cancer site i following a whole body dose at *low dose rate*.

The cancer sites considered were Bone Marrow, Bone, Breast, Lung, Stomach, Colon, Liver, Pancreas, Thyroid, Skin and Other cancers, where Other cancers has to be interpreted as a set of cancer sites other than the ones mentioned³. The target variables X_i were defined as the fraction of radiation induced cancer deaths over a lifetime for cancer site i following a whole body dose at low dose rate. The distribution of X_i has been obtained from the 3 elicitation variables for cancer site i using the model:

$$X_i = \frac{Y_{i,2}}{Y_{i,1}} Y_{i,3} \quad (2.9)$$

The simulation was done conditional on $y_{i,3} \leq y_{i,1} \leq y_{i,2}$ and $\sum_i x_i \leq 1$, where $y_{i,1}$, $y_{i,2}$, $y_{i,3}$ and x_i represent samples from $Y_{i,1}$, $Y_{i,2}$, $Y_{i,3}$ and X_i , respectively.

The dependency was queried among cancer sites for $Y_{i,2}$, between cancer sites of $Y_{i,1}$ and $Y_{i,2}$, and between cancer sites for $Y_{i,1}$ and $Y_{i,3}$.

As in Example 1, Appendix E lists the information on the conditional probabilities and corresponding Spearman's rank correlations for Example 2 via Tables E.5 and E.6. For this example the assessed conditional probabilities show agreement between experts.

The difference between Example 1 and Example 2 is that the elicitation variable is *no* target variable in Example 2; model 2.9 is used to obtain the distribution on the target variables X_i . The implications of this extra feature to both strategies will be discussed now.

Let D_i be defined as

$$D_i := \{(y_{i,1}, y_{i,2}, y_{i,3}) | y_{i,3} \leq y_{i,1} \leq y_{i,2}\}. \quad (2.10)$$

³The index i runs over the set consisting of {Bone Marrow, Bone, ..., Other Cancers}.

In case of Strategy 1, the expert assessments and the conditional probability are aggregated using Strategy 1 which result in a distribution for the DM on *all* elicitation variables, denoted by \mathbf{Y}_{DM} . The distribution on the target variables \mathbf{X} is obtained through correlated combination of marginal distributions of \mathbf{Y}_{DM} ; for cancer site i , the density of X_i is obtained by

$$x_i = \frac{y_{\text{DM},i,2}}{y_{\text{DM},i,1}} y_{\text{DM},i,3} 1_{D_i}(y_{\text{DM},i,1}, y_{\text{DM},i,2}, y_{\text{DM},i,3}). \quad (2.11)$$

and conditional on $\sum_i x_i \leq 1$. The sample $(y_{\text{DM},i,1}, y_{\text{DM},i,2}, y_{\text{DM},i,3})$ in model 2.11 is taken from \mathbf{Y}_{DM} . In summary, the distribution for the DM on the target variables is obtained in two steps; firstly the individual expert assessments are combined into a distribution on the elicitation variables \mathbf{Y}_{DM} for the DM and secondly this distribution is used to determine the minimal relative information distribution on the target variables \mathbf{X} using model 2.11.

For this example, Strategy 2 is modified slightly in order to determine the distribution for the DM on the target variables in a single step. For expert e , determine the minimal relative information distribution on all elicitation variables \mathbf{Y}_e satisfying Properties (i) and (ii) of Section 2.1.2. The density on the target variables is obtained by linear pooling the densities resulting from model 2.9; for cancer site i , the density of X_i is obtained by

$$x_i = \sum_{e=1}^{N_{\text{exp.}}} w_e \frac{y_{e,i,2}}{y_{e,i,1}} y_{e,i,3} 1_{D_i}(y_{e,i,1}, y_{e,i,2}, y_{e,i,3}). \quad (2.12)$$

and conditional on $\sum_i x_i \leq 1$. The sample $(y_{e,i,1}, y_{e,i,2}, y_{e,i,3})$ is taken from \mathbf{Y}_e . Because model 2.9 is non-linear, the marginal distributions on X_i resulting from Strategy 1 and Strategy 2 will be different, since for Strategy 1 the experts' assessments were combined prior to applying model 2.9, whereas for Strategy 2 the experts' assessments and application of the model were done at the same time.

Based on the quantile information⁴ of Table 2.1 the difference between Strategy 1 and Strategy 2 is observed; most of the 5% quantiles are roughly the same, but the majority of the 50% and 95% quantiles are different. This difference is illustrated graphically, see Figure 2.3.

The difference between the Spearman's rank correlation matrices resulting from Strategy 1 and Strategy 2 is illustrated via 'radar plots' in Figure 2.4. The Spearman's rank correlation matrices based on Strategy 1 and Strategy 2 are given in Appendix E Tables E.7 and E.3, respectively.

⁴The results for Strategy 1 and Strategy 2 for Example 2 are based on 987 and 916 samples, respectively and were determined with respect to a log-uniform background distribution.

X_i	5%		50%		95%	
	Strat.1	Strat.2	Strat.1	Strat.2	Strat.1	Strat.2
Bone Marrow (BM)	8.39e-8	8.77e-8	1.41e-4	2.98e-3	1.61e-2	1.31e-2
Bone (Bo)	2.82e-8	1.93e-8	3.21e-5	1.47e-5	3.56e-3	1.08e-3
Breast (Br)	1.07e-7	8.05e-8	2.11e-4	2.48e-3	1.90e-2	8.01e-3
Lung (Lu)	1.96e-8	1.66e-8	2.26e-4	1.77e-6	4.17e-2	1.76e-2
Stomach (St)	3.62e-8	1.92e-8	2.04e-4	8.56e-6	1.74e-2	6.23e-3
Colon (Co)	4.47e-8	1.97e-8	3.09e-4	2.40e-6	1.96e-2	1.16e-2
Liver (Li)	3.84e-8	2.13e-8	9.37e-5	5.63e-6	1.06e-2	1.35e-3
Pancreas (Pa)	4.32e-8	2.19e-8	3.00e-4	5.24e-6	7.75e-3	2.48e-3
Thyroid (Th)	3.29e-8	1.78e-8	1.22e-4	8.57e-7	3.53e-3	1.04e-3
Skin (Sk)	4.35e-8	5.59e-7	1.16e-4	1.36e-4	2.13e-3	6.96e-4
Other Cancers (OC)	5.51e-8	2.04e-8	1.29e-3	2.29e-5	6.41e-2	2.29e-2

Table 2.1: *Example 2: quantile information comparison of marginal distributions of DM using Strategy 1 and Strategy 2 for the target variables.*

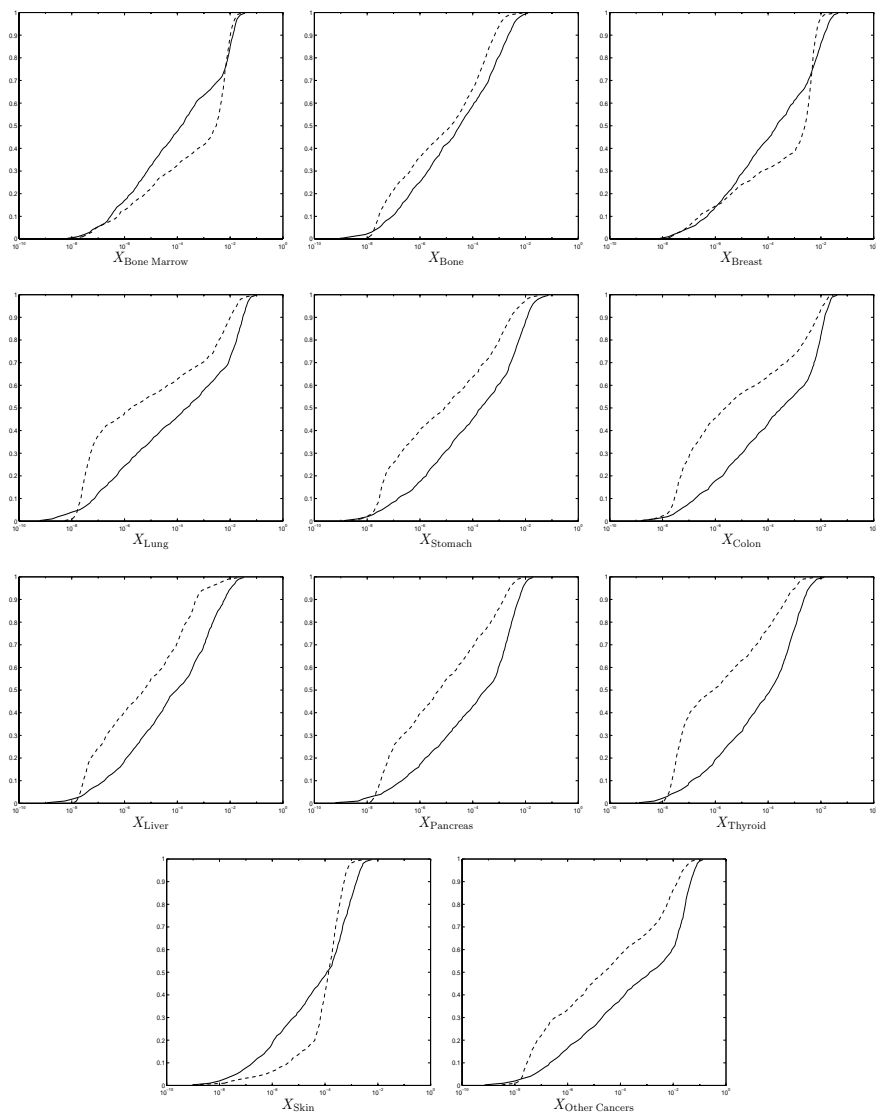


Figure 2.3: *Example 2: graphical display of marginal distributions of X_i for cancer site i resulting from Strategy 1 (--) and Strategy 2 (-).*

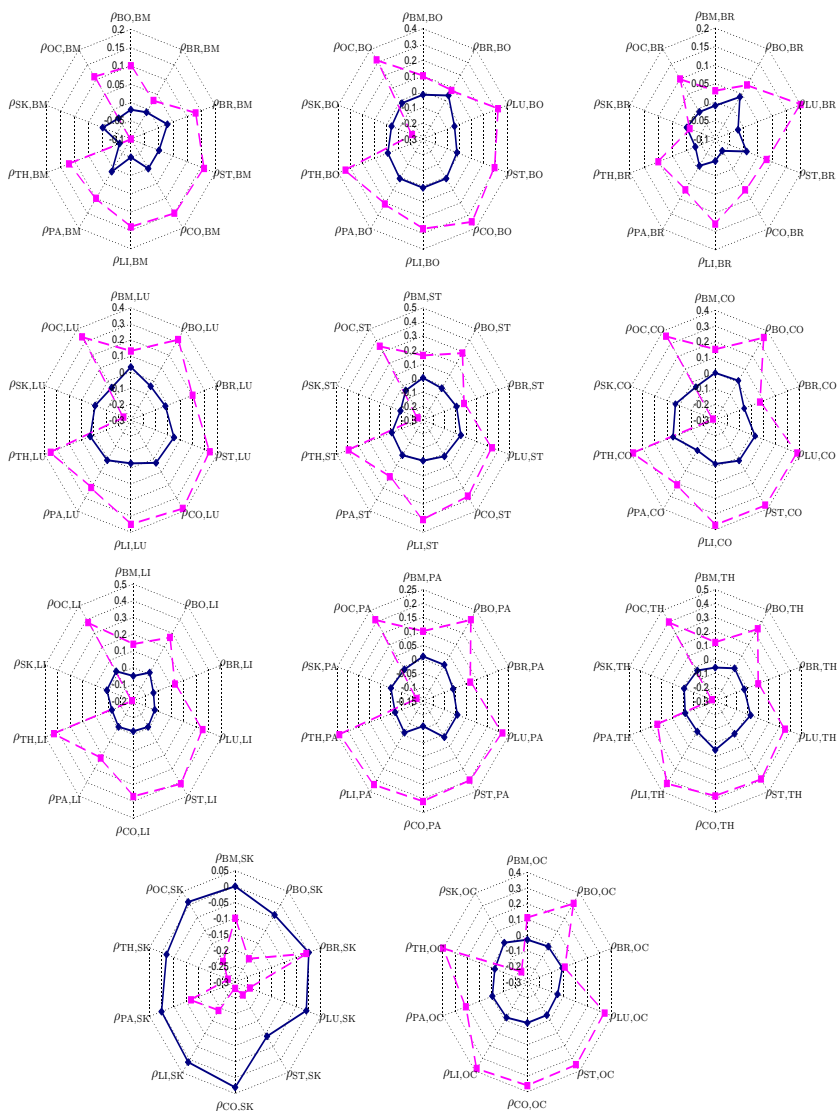


Figure 2.4: Example 2: radar plots of Spearman's rank correlations obtained from Strategy 1 (—) and Strategy 2 (---).

2.3 Dependencies and Probabilistic inversion

As a result of the Structured Expert Judgment Elicitation Methodology, dependencies may have been queried among elicitation variables which are not target variables. For example, the distribution on the target variables for Example 2 was determined by simulation. But what if the distribution on the target variables has to be determined by probabilistic inversion. Is it possible to use dependency information among the elicitation variables in the probabilistic inversion solution scheme? The answer to this question is affirmative.

Suppose the dependency between elicitation variables Y_1 and Y_2 is available for the DM as a conditional probability $\pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_1, Y_2)$. This conditional probability and measure ν_k (with density g_k), as determined using the efficient version of PREJUDICE (Section 1.6), are related by

$$\pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_1, Y_2) = 2 \sum_{\mathbf{x} \in M} \sum_l g_k(l) 1_{A_2}(\mathbf{x}) \quad (2.13)$$

where $A_2 = \{T_{(1)}(\mathbf{x}) > y_{1, \text{DM}, 50\%}, T_{(2)}(\mathbf{x}) > y_{2, \text{DM}, 50\%}, T_k(\mathbf{x}) = l\}$. The inclusion of Expression 2.13 in the efficient version of PREJUDICE (Section 1.6) is best explained as the addition of constraints to the PI problem; in notation

$$\begin{aligned} \min_{g_k} \quad & \sum_l g_k(l) \log \frac{g_k(l)}{g_{\gamma, k}(l)} \\ \text{s.t.} \quad & \sum_{\mathbf{x} \in M} \sum_l g_k(l) 1_{A_1}(\mathbf{x}) = \nu(I_{j, l_j}) \\ & \sum_{\mathbf{x} \in M} \sum_l g_k(l) 1_{A_2}(\mathbf{x}) = \frac{1}{2} \pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_1, Y_2). \\ & g_k \geq 0. \end{aligned} \quad (2.14)$$

with $A_1 = \{T_{(j)}(\mathbf{x}) \in B_l, B_l \cap I_{j, l_j} \neq \emptyset\}$, $j = 1, \dots, n$ and $l_j = 1, \dots, K + 1$. Including dependence information in a PI problem by the addition of constraints will most likely result in a smaller set of feasible solutions. Currently, the issue on what to do if the set of feasible solutions for PI problem (2.14) is empty has not been investigated. Consequently, the implications to the iterative version of PREJUDICE are unclear at this moment. At this moment the recommended strategy would be to determine a distribution $\tilde{\mu}^{(1)}$ on M using the iterative version of the efficient version of PREJUDICE and next determine a distribution $\tilde{\mu}^{(2)}$ on M which has minimum relative information with respect to $\tilde{\mu}^{(1)}$ and which also satisfies the dependence information.

2.3.1 Example

To illustrate the effect of taking dependencies into account in probabilistic inversion, the dispersion coefficient example of Section 1.6 is considered. No dependencies were elicited among the elicitation variables for the Dispersion

& Deposition panel, however, in order to illustrate the process, the conditional probabilities for the DM, as given in Table 2.2, are assumed to be representative: The conditional probabilities express a strong dependence

Elicitation variables	$\pi_{\frac{Y_i}{2}, \frac{Y_j}{2}}^{\text{DM}}(Y_i, Y_j)$
Y_1, Y_2	0.90
Y_2, Y_3	0.85
Y_3, Y_4	0.75
Y_4, Y_5	0.60

Table 2.2: *Dispersion coefficient example (stability class C): conditional probabilities between elicitation variables.*

between the elicitation variables Y_1 and Y_2 , the lateral plume spreads at 500 m. and 1 km. downwind, respectively, and decreases as the difference between the downwind distances of the elicitation variables increases. In Section 1.6, the distribution on the samples, taking no dependency information into account, was determined, i.e. $\tilde{\mu}^{(1)}$. Here, a distribution on the same samples will be determined taking the dependence information into account, i.e. $\tilde{\mu}^{(2)}$; the intuitive understanding is that the probabilities on the propagated samples in the ‘independent’ situation are ‘re-weighted’ such that the conditional probabilities are also satisfied⁵.

Table 2.3 compares quantile information of the target variables for the ‘independent’ case and ‘dependent’ case.

	Quantile	Extension	
		independent case	dependent case
A_y	5%	2.74e-3	2.31e-3
	50%	2.10e-1	2.36e-1
	95%	3.39	5.85
B_y	5%	6.12e-1	5.73e-1
	50%	9.46e-1	8.91e-1
	95%	1.40	1.47
ρ_{A_y, B_y}		-9.08e-1	-9.63e-1
Relative Information		4.61	5.62

Table 2.3: *Dispersion coefficient example (stability class C): quantile information and rank correlations for target variables for independent case and dependent case.*

The results of Table 2.3 show that the 5%-95% quantile intervals for dependent case are wider and the rank correlation has become more negative

⁵Here the background distribution for the ‘dependent’ case is taken the same as the background distribution in the ‘independent’ case. However, another interesting background distribution for the ‘dependent’ case would be the distribution on the samples of the ‘independent’ case.

compared to the results for independent case. A graphical comparison of the marginal distributions on the target variables is presented in Figure 2.5.

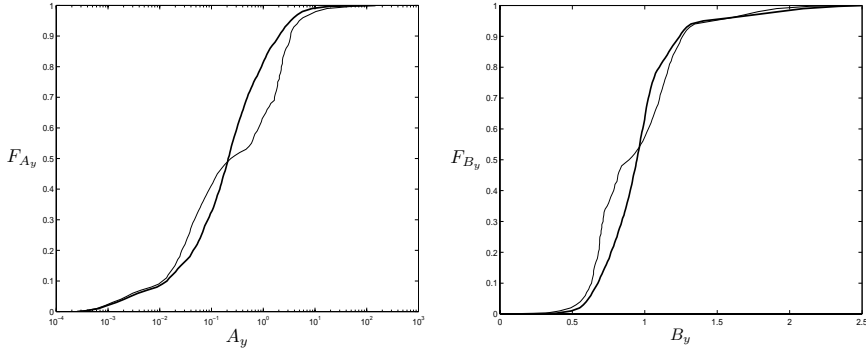


Figure 2.5: *Dispersion coefficient example (stability class C): graphical illustration of distribution of A_y (left), B_y (right) for the independent case (-) and dependent case (-).*

The 5%, 50% and 95%-iles of the marginal distributions of the push-forward distributions of the two cases are not given as they agree with the 5%, 50% and 95%-iles of the marginal distributions of the DM, see Table 1. For the dependent case, the conditional probabilities inherent in the push-forward distribution agree with the initial conditional probabilities $\pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_i, Y_j)$, see Table 2.2.

The Spearman's rank correlation matrices for target variables and elicitation variables for the independent and dependent case are given in Table 2.4 and Table 2.5, respectively. Based on comparing Spearman's rank correlations between the correlation matrices, the effect of taking dependencies into account in the probabilistic inversion becomes clear, see Figure 2.6.

Based on the $\rho\pi$ -table, it is possible to convert the rank correlations ρ_{Y_i, Y_j} into conditional probabilities $\pi_{\frac{1}{2}, \frac{1}{2}}(Y_i, Y_j)$ which can be compared to the conditional probabilities $\pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_i, Y_j)$ of Table 2.2. Table 2.6 shows that the 'converted' conditional probabilities $\pi_{\frac{1}{2}, \frac{1}{2}}(Y_i, Y_j)$ differ significantly from the 'input' conditional probabilities $\pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_i, Y_j)$. This observation arouses the suspicion that the dependence structure contained in the distribution on the target variables cannot be represented satisfactorily by a Spearman's rank correlation matrix.

Hence the push-forward results of the joint distribution based on marginal distribution and Spearman's rank correlation matrix may not represent the DM results. See Table 2.7 for a comparison between DM quantiles and quantiles from the push-forward distributions of the dependent and independent case

From Table 2.7 the effect of representing the distribution as marginal

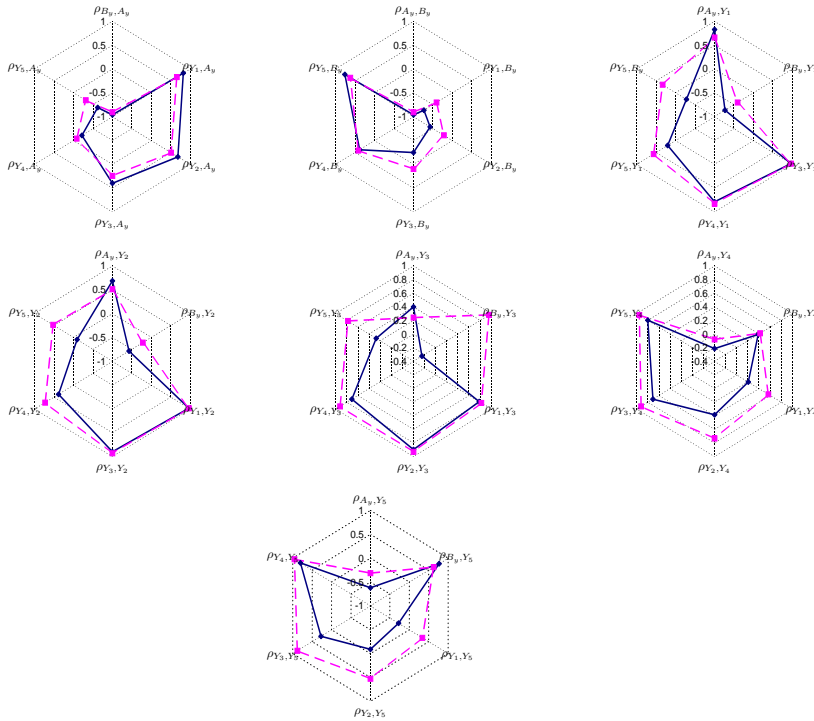


Figure 2.6: *Dispersion coefficient example (stability class C): graphical comparison of Spearman's rank correlations resulting from the independent case (—) and the dependent case (--).*

distributions and Spearman's rank correlation becomes evident; the results of the correlated propagation do not represent the DM well at all.

Based on the above it is questionable to represent the joint distribution resulting from PI as marginal distributions and a Spearman's rank correlation matrix. Isn't it possible to make use of a special sampling scheme which will generate a file of realizations based on the joint distribution on the target variables which can be fed into the computer code? In the future more research will be done to develop sampling techniques which do not violate the dependency structure as much as the marginal distributions and Spearman's rank correlation representation.

2.4 Conclusions

This section started with a review of the existing dependence elicitation techniques. It is concluded that the conditional probability dependence

$$\begin{array}{l}
 A_y \\
 B_y \\
 Y_1 \\
 Y_2 \\
 Y_3 \\
 Y_4 \\
 Y_5
 \end{array}
 \left(
 \begin{array}{cccccccc}
 1 & -9.08\text{e-}1 & 6.58\text{e-}1 & 5.09\text{e-}1 & 2.41\text{e-}1 & -7.96\text{e-}2 & -3.14\text{e-}1 & \\
 -9.08\text{e-}1 & 1 & -4.08\text{e-}1 & -2.18\text{e-}1 & 9.60\text{e-}2 & 4.21\text{e-}1 & 6.27\text{e-}1 & \\
 6.58\text{e-}1 & -4.08\text{e-}1 & 1 & 9.68\text{e-}1 & 8.22\text{e-}1 & 5.69\text{e-}1 & 3.37\text{e-}1 & \\
 5.09\text{e-}1 & -2.18\text{e-}1 & 9.68\text{e-}1 & 1 & 9.28\text{e-}1 & 7.32\text{e-}1 & 5.26\text{e-}1 & \\
 2.41\text{e-}1 & 9.60\text{e-}1 & 8.22\text{e-}1 & 9.28\text{e-}1 & 1 & 9.23\text{e-}1 & 7.81\text{e-}1 & \\
 -7.96\text{e-}2 & 4.21\text{e-}1 & 5.69\text{e-}1 & 7.32\text{e-}1 & 9.23\text{e-}1 & 1 & 9.56\text{e-}1 & \\
 -3.14\text{e-}1 & 6.27\text{e-}1 & 3.37\text{e-}1 & 5.26\text{e-}1 & 7.81\text{e-}1 & 9.56\text{e-}1 & 1 &
 \end{array}
 \right)$$

Table 2.4: *Dispersion coefficient example (stability class C): Spearman's rank correlations for target variables and elicitation variables for the 'independent' case.*

$$\begin{array}{l}
 A_y \\
 B_y \\
 Y_1 \\
 Y_2 \\
 Y_3 \\
 Y_4 \\
 Y_5
 \end{array}
 \left(
 \begin{array}{cccccccc}
 1 & -9.63\text{e-}1 & 8.25\text{e-}1 & 6.84\text{e-}1 & 3.97\text{e-}1 & -2.14\text{e-}1 & -6.21\text{e-}1 & \\
 -9.63\text{e-}1 & 1 & -7.32\text{e-}1 & -5.73\text{e-}1 & -2.48\text{e-}1 & 3.86\text{e-}1 & 7.67\text{e-}1 & \\
 8.25\text{e-}1 & -7.32\text{e-}1 & 1 & 9.66\text{e-}1 & 7.84\text{e-}1 & 2.07\text{e-}1 & -2.75\text{e-}1 & \\
 6.84\text{e-}1 & -5.73\text{e-}1 & 9.66\text{e-}1 & 1 & 8.95\text{e-}1 & 3.85\text{e-}1 & -9.10\text{e-}2 & \\
 3.97\text{e-}1 & -2.48\text{e-}1 & 7.84\text{e-}1 & 8.95\text{e-}1 & 1 & 7.11\text{e-}1 & 2.72\text{e-}1 & \\
 -2.14\text{e-}1 & 3.86\text{e-}1 & 2.07\text{e-}1 & 3.85\text{e-}1 & 7.11\text{e-}1 & 1 & 8.03\text{e-}1 & \\
 -6.21\text{e-}1 & 7.67\text{e-}1 & -2.75\text{e-}1 & -9.10\text{e-}2 & 2.72\text{e-}1 & 8.03\text{e-}1 & 1 &
 \end{array}
 \right)$$

Table 2.5: *Dispersion coefficient example (stability class C): Spearman's rank correlations for target variables and elicitation variables for the 'dependent' case.*

elicitation technique is most suitable in the context of expert judgment. This type of dependence elicitation has a sound foundation in probability theory and is in line with the Structured Expert Judgment Elicitation Methodology. It is *not* claimed to be the best way to elicit dependence, but experts with different backgrounds understood the technique and type of questioning fairly easily and acknowledged its importance. Furthermore, it was pointed out that it is more important to get the marginal distributions 'right' compared to getting the dependence between the elicitation variables 'right'.

Next the attention was focused on the aggregation experts' assessments on marginal distributions and dependence into a joint distribution on the

Elicitation variables	Cond. prob. from experts	Independent case		Dependent case	
	$\pi_{\frac{1}{2}, \frac{1}{2}}^{\text{DM}}(Y_i, Y_j)$	ρ_{Y_i, Y_j}	$\pi_{\frac{1}{2}, \frac{1}{2}}(Y_i, Y_j)$	ρ_{Y_i, Y_j}	$\pi_{\frac{1}{2}, \frac{1}{2}}(Y_i, Y_j)$
Y_1, Y_2	9.00e-1	9.68e-1	9.94e-1	9.66e-1	9.93e-1
Y_2, Y_3	8.50e-1	9.27e-1	9.51e-1	8.95e-1	9.03e-1
Y_3, Y_4	7.50e-1	9.23e-1	9.45e-1	7.11e-1	8.09e-1
Y_4, Y_5	6.00e-1	9.56e-1	9.91e-1	8.03e-1	8.50e-1

Table 2.6: *Dispersion coefficient example (stability class C): comparison between conditional probabilities from experts and ‘converted’ conditional probabilities based on the $\rho\pi$ table.*

Distance (z_i)	Quantiles	$\sigma_y(z_i)$		
		DM	UNICORN	
	Dependent case		Independent case	
			$\rho_{A_y, B_y} = -0.963$	$\rho_{A_y, B_y} = -0.908$
500 m.	5%	3.30e+1	7.95	4.37
	50%	9.49e+1	6.52e+1	6.58e+1
	95%	1.95e+2	2.40e+2	3.18e+2
1 km.	5%	6.48e+1	1.89e+1	9.14
	50%	1.72e+2	1.25e+2	1.24e+2
	95%	3.46e+2	4.17e+2	6.03e+2
3 km.	5%	1.75e+2	5.24e+1	2.86e+1
	50%	4.46e+2	3.34e+2	3.38e+2
	95%	1.04e+3	1.11e+3	1.79e+3
10 km.	5%	4.48e+2	1.52e+2	9.34e+1
	50%	1.22e+3	9.93e+2	1.02e+3
	95%	3.37e+3	4.03e+3	6.42e+3
30 km.	5%	1.10e+3	3.75e+2	2.38e+2
	50%	2.82e+3	2.40e+3	2.78e+3
	95%	8.25e+3	1.42e+4	2.30e+4

Table 2.7: *Dispersion coefficient example (stability class C): comparison of quantile information of DM and UNICORN for the independent and dependent case.*

elicitation variables. Two strategies based on linear pooling were introduced. Roughly, Strategy 1 combines the experts’ assessments on marginal distributions and dependence separately, whereas Strategy 2 combines the experts’ assessments on marginal distributions and dependence together. Looking at how the assessments of the expert on marginal distributions and dependence are treated by both strategies, it is concluded that Strategy 2 treats the experts’ assessments in a more ‘natural’ way than Strategy 1; currently the marginal distributions of the experts are combined, in case dependency information is available it would be ‘natural’ to combine the joint distributions of the experts. Additionally, Strategy 2 is able to deal with situations where the structure of the acyclic dependency trees of experts are not identical. In this situation Strategy 1 may result in a cyclic dependence

tree for the DM, which may result in a Spearman's rank correlation matrix which is not positive definite. In case of Strategy 2, a minimal relative information distribution satisfying Properties (i) and (ii) of Section 2.1.2 is fitted to the experts' assessments, hence the structure of the experts acyclic dependency trees do not have to be identical.

Both strategies were applied to two examples: Example 1 (elicitation variable is target variable) and Example 2 (elicitation variable is not a target variable). For Example 1 the marginal distributions resulting from the strategies were the same but the Spearman's rank correlation matrices differed significantly. Due to the non-linearity of the model, both the marginal distributions and Spearman's rank correlation matrices obtained from Strategy 1 and Strategy 2 differed.

Section 2.3 focused on using dependence information among the elicitation variables in probabilistic inversion. It was shown that it is very easy to extend the probabilistic inversion solution scheme with information on dependence between elicitation variables. For the dispersion coefficient example, the difference between taking and not taking dependence into account in determining the distribution on the target variables became apparent. At this moment it is unclear what the implications of taking account of dependence information are for the iterative version of PREJUDICE. Therefore it is suggested to apply first the iterative version of PREJUDICE taking no account of dependence information in order to determine a domain M such that it is observationally or hypercube complete. The final distribution on the target variables will be a re-weighted version of this distribution, such that the dependence structure of its push-forward will also comply with the elicited dependence information and such that the difference between the distribution obtained in the first step will be minimal, when measured using relative information.

In doing uncertainty analysis, the distribution on the target variables is often represented by marginal distributions and Spearman's rank correlation matrix, see [37], [38]. For many situations, this type of representation of a distribution on the target variables will result in push-forward results which will resemble the distribution of the DM poorly. It seems that the dependence structures inherent in the distribution on target variables, obtained using probabilistic inversion, are difficult to capture using a dependence measure, like the Spearman's rank correlation. Therefore, future research will focus on application/development of new sampling schemes which can deal with more complex dependence structures.

Chapter 3

Calibration with Uncertain Observations

Classical references for calibration of subjective probabilities ([41],[42] for an update), treat calibration as a measure of ‘correspondence with reality’ but do not deal with the problem of the stochastic nature of numerical measures of calibration. Asymptotic definitions of *well-calibrated* were proposed in [16],[17], and a definition in terms of statistical likelihood was introduced in [7],[9],[13],[24]. In [9], the statement that an expert is well calibrated is interpreted as the statistical hypothesis that the uncertain quantities are independent and identically distributed with the appropriate distribution. This hypothesis is tested against observations/experimental results using the distribution of an appropriate scoring variable. In the most common implementation, experts assess K quantiles of their subjective distributions for continuous quantities Y_j . Let probability vector $\mathbf{q} = (q_1, \dots, q_{K+1})$ contain the probabilities of the $K + 1$ interquantile intervals, and probability vector $\mathbf{p} = (p_1, \dots, p_{K+1})$ be defined as the proportion of observations contained in the interquantile intervals and n be the total number of observations. Calibration is scored as the upper tail probability (UTR) of the variable R

$$R = 2n I(\mathbf{p}|\mathbf{q}) \tag{3.1}$$

with $I(\mathbf{p}|\mathbf{q}) = \sum_{l=1}^{K+1} p_l \ln \frac{p_l}{q_l}$. If the observations are drawn independently from distributions with interquantile probabilities q_l , $l = 1, \dots, K + 1$, then R is asymptotically χ_K^2 distributed, as $n \rightarrow \infty$. Figure 3.1 shows the density of the χ_3^2 distribution, where the calibration score is equal to the shaded area and equals $P(R > r) = 1 - \chi_3^2(r)$.

Evaluation of probabilistic forecasters in terms of scoring rules was introduced in [47]. [18] studied quantities similar to R , though in a slightly different context. The evaluation of sets of assessments against sets of ob-

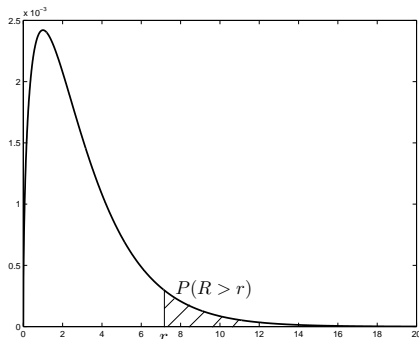


Figure 3.1: Density of χ_3^2 distribution.

servations was undertaken in [9], and R and UTR were shown to be strictly proper in an appropriately generalized sense.

All of the developments sketched above assume that ‘crisp’ values of uncertain quantities Y_j can be observed. In the Joint CEC/USNRC Uncertainty Analysis, the question was raised by project oversight authorities to what extent ‘measurement variability’ might affect the calibration scores of experts. For example, experts were asked to assess quantiles for lateral plume spread at various downwind distances under various atmospheric conditions. In some cases experimental values were available, and these were used to calibrate experts. Since lateral plume spread is a very complex physical process and relevant physical variables cannot all be controlled in any given experimental procedure, hence the reproducibility of experimental results is not high. Different experimental procedures for measuring lateral plume spread may have different strengths and weaknesses. Since the elicitation did not specify an experimental procedure, the values used to score calibration might be contaminated by a measurement variability which the experts could not take into account. It is assumed throughout that the distribution of measurement variability is known and is independent of other measurement variabilities and elicited quantities.

Let the observed variables be $\mathbf{Z} = (Z_1, \dots, Z_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ the elicitation variable, which are related via

$$Z_j = Y_j + \epsilon_j. \quad (3.2)$$

The measurement variability ϵ_j is independent of ϵ_{j^*} ($j \neq j^*$) and independent of the observed and elicitation variables. Because of physical considerations it could be that the observed variables may take positive values only. If the additive model 3.2 yield negative values, it will be replaced by the multiplicative version

$$Z_j = Y_j \epsilon_j. \quad (3.3)$$

Taking the logarithm of multiplicative model 3.3 results in an additive model among the logarithms of the random variables;

$$\log(Z_j) = \log(Y_j) + \log(\epsilon_j). \quad (3.4)$$

Roughly, scoring calibration taking account of measurement variability involves 2 steps

Step 1 : The distribution of Z_j is obtained by *folding* the measurement variability ϵ_j into the distribution of Y_j , via model 3.2 or 3.4.

Step 2 : Determination of the calibration score based on the distribution of Z_j .

3.1 Step 1: The distribution of Z_j

If the entire distribution function for Y_j is available and the distribution of ϵ_j is known and independent of Y_j , then the distribution for each expert for Z_j is computed by simply taking the convolution of the density functions f_{Y_j} and f_{ϵ_j} . However, it is assumed here that the entire distribution of Y_j , and consequently Z_j , is not known. For reasons explained in Section 1.1, the distribution with minimum relative information with respect to a certain background distribution¹ under the constraint that the 5%, 50% and 95% quantiles agree with the expert's assessments, is determined. The question which remains to be answered is: should the minimum relative information arguments be applied to elicitation variables Y_j or to observed variables Z_j ?

The following numerical example illustrates the differences between the application of minimum relative information arguments to either Y_j or Z_j . The 5%, 50% and 95% quantile assessments for Y are $y_5 = 11.5$, $y_{50} = 13$, $y_{95} = 14$. Additionally the 0% and 100% quantile assessments were chosen as $y_0 = 10$ and $y_{100} = 15$. The vector of interquantile probabilities is

$$\mathbf{b} = (0.05, 0.45, 0.45, 0.05)^T. \quad (3.5)$$

The measurement variability was chosen to be normal distributed on the interval $[-4, 4]$, such that $P(\epsilon \leq -4) = P(\epsilon > 4) = 10^{-3}$, see Figure 3.2 (the method is not restricted to the measurement variability being normal distributed).

In the next sections the methods of applying minimum relative information arguments on Y and Z , referred to as Relinf Y and Relinf Z respectively, are introduced and illustrated using an the above example. In the implementation of the methods, discrete approximations of the respective densities will be determined.

¹Consideration for the choice of the appropriate background distribution are given in Section 1.3.

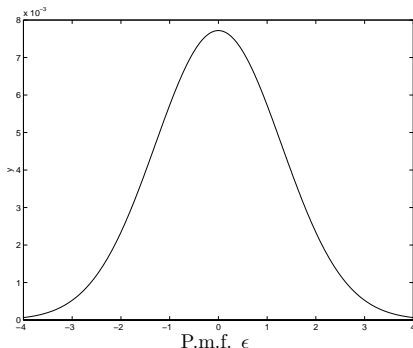


Figure 3.2: Probability density function of measurement variability ϵ on $[-4, 4]$.

3.1.1 Relinf Y : Minimum relative information for Y

Determining the minimum relative information density function f_Y with respect to background density f_λ , under the constraint that the sum of probabilities of realizations contained in the respective interquantile intervals equals \mathbf{b} , can be written as a CP problem. Let the primal formulation of the CP problem be given by

$$\begin{aligned} \min_{f_Y} \quad & \sum_{l=1}^{N_Y} f_Y(y_l) \ln \frac{f_Y(y_l)}{f_\lambda(y_l)} \\ \text{s.t.} \quad & Af_Y = \mathbf{b} \end{aligned} \quad (3.6)$$

The matrix $A \in \mathbb{R}^{(K+1) \times N_Y}$, with N_Y the number of discretization points for Y . The i -th column of A consists of zeros, except on the entry which corresponds to the interquantile interval in which y_i is contained

$$a_{l,i} := 1_{I_l}(y_i) \quad (3.7)$$

where I_l is the l -th interquantile interval of Y . Corollary 1.2.2 of Section 1.2.2 stated that if

$$f_Y = f_\lambda \cdot \exp(A^T \mathbf{y} - \mathbf{e}). \quad (3.8)$$

holds for some feasible primal solution f_Y and feasible dual solution $\mathbf{y} \in \mathbb{R}^{K+1}$, then f_Y is the optimal solution of primal formulation (3.6) and \mathbf{y} the optimal solution of the dual formulation of CP problem (3.6). The vector $\mathbf{e} \in \mathbb{R}^{K+1}$ represents the unit vector of dimension $K+1$.

In case the background density f_λ is uniform, it can be seen easily that $f_Y(y_1) = f_Y(y_2)$ for all $y_1, y_2 \in I_l$. Hence, f_Y is a piece-wise uniform density, see Figure 3.3 (left), with $N_Y = 500$.

The final step is to convolute the densities of Y and ϵ to obtain the density f_Z of Z , see Figure 3.3 (right). The relative information of f_Z with respect to its background density f_γ is equal to 6.94e-1.

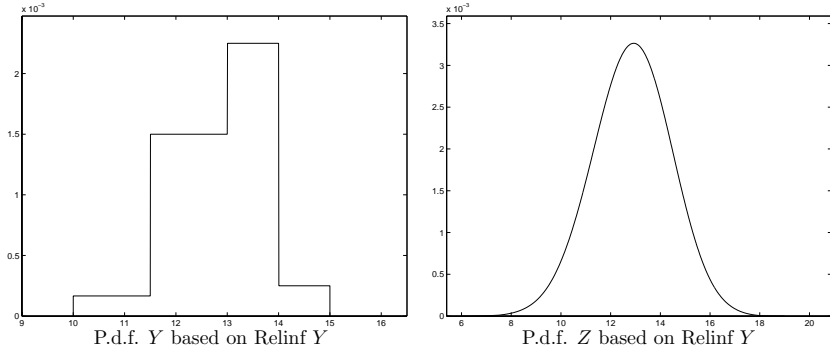


Figure 3.3: *Relinf Y*: probability density functions of Y and Z .

3.1.2 Relinf Z : Minimum relative information for Z_j

Instead of applying the minimum relative information argument to Y_j , it may also be applied to Z_j . Relinf Z determines the minimum relative information distribution of Z_j under the constraints that the sum of probabilities of realizations contained in the respective interquantile intervals of Y_j equals \mathbf{b} , and that the density of Z is the convolution of the densities of Y_j and ϵ ;

$$\begin{aligned} \min_{f_Z} \quad & \sum_{l=1}^{N_Z} f_Z(z_l) \ln \frac{f_Z(z_l)}{f_\gamma(z_l)} \\ \text{s.t.} \quad & Af_Y = \mathbf{b} \\ & f_Z(z_l) = \sum_{m=1}^{N_Y} f_Y(y_m) f_\epsilon(z_l - y_m) 1_{[-4,4]}(z_l - y_m) \end{aligned} \tag{3.9}$$

where N_Z represent the number of discretization points for Z . NLP problem 3.9 is also a convex programming problem², for the same reasons as described in Section 1.5 is optimization package MOSEK (www.mosek.com) has been used. The example was calculated based on Relinf Z ($N_Z = N_Y = 500$) and resulted in a density of Z with a relative information value of 5.08e-1, see Figure 3.4 (Right).

Note that the relative information value for the density of Z based on Relinf Y is greater than based on Relinf Z . But the density of Y based on Relinf Z is highly informative, see Figure 3.4 (Left). It seems that minimizing the relative information of Z given the constraints, results in a density of f_Y with probability on or close to the 0%, 5%, 95% and 100% percentiles only.

3.1.3 Limiting behavior

In case of no measurement variability it is easy to see that the Relinf Y and Relinf Z approaches are the same. Figure B.7 illustrates the effect

²The constraint describing the convolution may look non-linear, however f_Y is the optimization variable and f_ϵ is known.

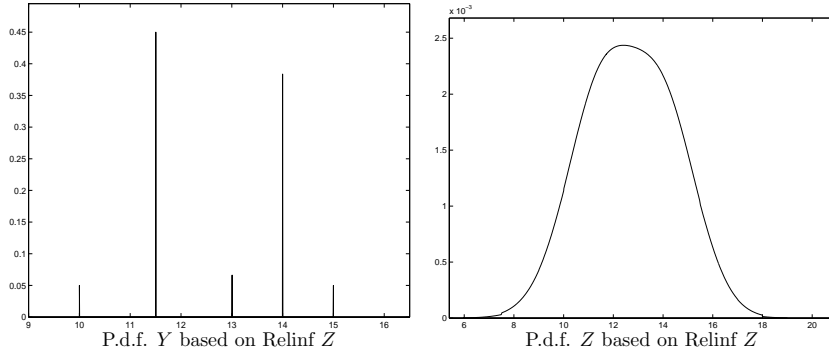


Figure 3.4: *Relinf Z*: probability density functions of Y and Z .

on the densities of Z based on Relinf Y (left) and densities of Y (middle) and Z (right) based on Relinf Z for different measurement variabilities. The density for Y based on Relinf Y remains the same in all instances, see Figure 3.3. The numbers of discretization points are $N_Y = N_Z = 1000$. The ‘wild’ behavior for $\epsilon \in [-0.01, 0.01]$ is due to the discretization.

3.2 Step 2: Scoring calibration based on Z_j

The calibration score introduced at the beginning of the Chapter is described in detail in [9], [13]. However, this calibration score is not capable of dealing with measurement variability. In this section the calibration score is generalized such that assessments can be scored with respect to sampling distributions of variables rather than with individual values of the variables.

This section starts off with a more mathematical description of scoring calibration with individual values and is followed by the description of the extension which scores calibration using sampling distributions over the values. Clearly, both scoring techniques can be used in combination with Relinf Y and Relinf Z .

3.2.1 Scoring with individual values

For the purpose of illustration, suppose that for each elicitation variable Y_j ($j = 1, \dots, n$), the 5%, 50% and 95% quantiles are assessed using expert judgment. The vector of interquantile probabilities \mathbf{b} is given by Expression 3.5. Let I_{j,l_j}^* represent the k -th interquantile interval of Z_j , such that the probability of I_{j,l_j}^* is equal to the l_j -th entry of \mathbf{b} .

When observations/experimental results $\mathbf{z}^* = (z_1^*, \dots, z_n^*)$ are available,

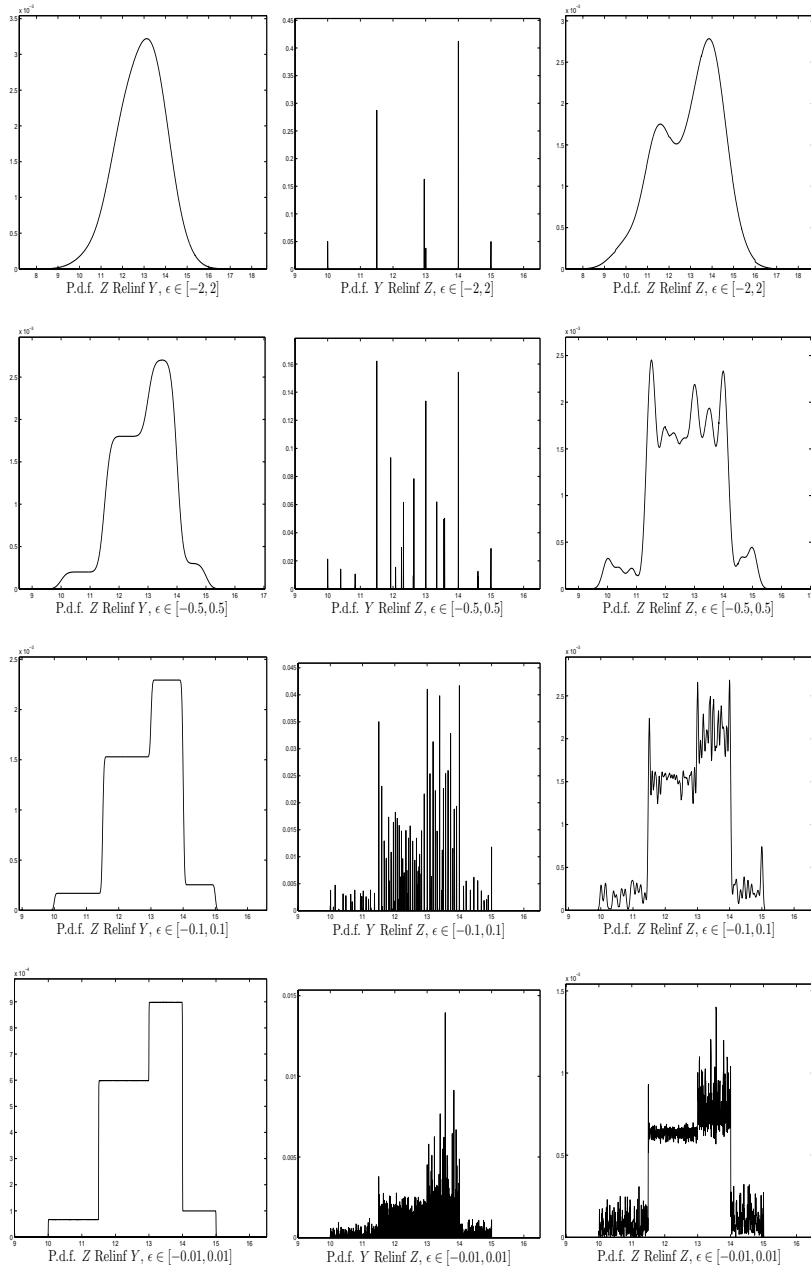


Figure 3.5: Probability density functions of Y and Z based on $Relinf Y$ and $Relinf Z$ for various measurement variabilities ϵ .

the sample distribution is

$$\begin{aligned} \mathbf{p}_{\mathbf{z}^*}(n) &:= (p_{\mathbf{z}^*,1}(n), \dots, p_{\mathbf{z}^*,4}(n)) \\ p_{\mathbf{z}^*,k}(n) &= \frac{\sum_{i=1}^n 1_{I_{j,l_j}^*}(z_j^*)}{n} \quad \text{for } l_j = 1, \dots, 4. \end{aligned}$$

Based on $\mathbf{p}_{\mathbf{z}^*}(n)$ and \mathbf{b} , the value of R is calculated using Expression 3.1, from which the calibration score is determined.

3.2.2 Scoring with sampling distributions

Given $\mathbf{z}^* = (z_1^*, \dots, z_n^*)$, let q_{j,l_j} denote the probability that the unobserved y_j falls in the k -th interquantile interval I_{j,l_j} of Y_j ; the conditional probability q_{j,l_j} can be written as

$$\begin{aligned} q_{j,l_j} &= P(y_j \in I_{j,l_j} \mid Z_j = z_j^*) = \\ &= \sum_{y_j \in I_{j,l_j}} \frac{p_\epsilon(z_j^* - y_j) p_{Y_j}(y_j)}{p_{Z_j}(z_j^*)} 1_{A_\epsilon}(z_j^* - y_j) \end{aligned} \quad (3.10)$$

where A_ϵ is the support of ϵ ; a measurable set such that $P(A_\epsilon) = 1$. The set of possible interquantile intervals for y_j is a subset of $\{1, \dots, K + 1\}$. In words, there may be many unobserved realizations y_j from different interquantile intervals I_{j,l_j} which in combination with a possible value of ϵ_j result in observation z_j^* . The likelihood that unobserved realizations $y_j \in I_{j,l_j}$ in combination with the measurement variability result in z_j^* is given by q_{j,l_j} .

Let $\pi = (\pi_1, \dots, \pi_n)$ be a vector of length n whose entries take values in the set $\{1, \dots, K + 1\}$, i.e. $\pi_j := k$ if $y_j \in I_{j,l_j}$. Given (z_1^*, \dots, z_n^*) , let the probability mass function of the sample distributions \mathbf{p}_π , generated by all possible interquantile interval combinations π , be proportional to

$$P(\mathbf{p}_\pi) \propto \prod_{j=1}^n q_{j,\pi_j}. \quad (3.11)$$

Given the probability mass function of \mathbf{p}_π , we compute $\text{UTR}(\pi)$ from the χ^2 approximation as in the previous section, for n sufficiently large. The calibration is obtained by computing the expected value over all possible interquantile interval combinations π

$$E(\text{UTR}) = \sum_{\pi} \text{UTR}(\pi) P(\mathbf{p}_\pi). \quad (3.12)$$

Theoretically, such computations may give rise to computational problems on current platforms for problems involving a large number of observations/experimental results. However, the next section will list some results.

3.3 Examples

In this section the methods for determining the distribution on Z_j (i.e. Relinf Y and Relinf Z) and to score calibration are tested on two examples. The first example is taken from the Joint CEC/USNRC Uncertainty Analysis [27] and the second example is taken from an expert judgment study for an uncertainty analysis for the inundation probability conducted for the Ministry of Transport, Public Works and Water Management [21].

3.3.1 Dispersion example

For this example the quantile assessments for the DM are given in Appendix F; the density of the DM has been constructed based on an equal weighted combination of experts densities. In total 10 seed variable questions, i.e. questions for which the answers are known to the project staff but not to the experts, are considered. The reader is referred to [27] for descriptions of the experiments which measured the seed variables. Considering no measurement variability, the calibration score based on the DM quantile assessments and measurements equaled 6.83e-1. This calibration score will be compared to the calibration scores resulting from the possible combinations of determining the distribution on Z_j and scoring calibration.

Information on the measurement variability was specified by the author, in terms of medians and error factor EF_{95} ³. Instead of considering one measurement variability, three measurement variabilities were considered; the median values were fixed but the error factor EF_{95} was multiplied by c , the value of c considered are 1, 1.5 and 2. The multiplicative model 3.3 was used to avoid negative realizations of Z_j , except for question B-5-600 sig-y for which the additive version was used. The number of discretization points are $N_Y = N_Z = 1000$. In general, the distributions for the Z_j 's obtained from Relinf Y are 'less informative/more dispersed' than the distributions obtained from Relinf Z , see Table 3.1 for the relative information values and Table 3.2 for the assessments for $c = 1$, where Y represents the DM assessments.

Figure 3.6 presents the densities of Y and Z for seed variable question B-1-220 chi/Q for $c = 1$. Note that the probability density function for Y was determined relative to a log-uniform background distribution.

Looking closer at Table 3.2, some counter intuitive results surface; by 'folding' the measurement variability into the assessments intuitively one would think that the resulting 5%-95% quantile interval of Relinf Y and Relinf Z would contain the 5%-95% quantile interval of the original assessments Y . This does not hold for most questions, especially the resulting 95% quantiles are smaller than the original 95% quantiles. At this moment

³The error factor EF_{95} is defined as $EF_{95} := \frac{z_{95\%}}{z_{50\%}}$, where z_{50} and z_{95} are the median and 95% quantiles of the distribution of Z , respectively.

Seed variable		Relinf Y	Relinf Z
B-1-220 chi/Q	$c = 1$	2.84	2.52
	$c = 1.5$	3.48	3.17
	$c = 2$	3.99	3.36
B-1-315 chi/Q	$c = 1$	1.34	1.28
	$c = 1.5$	1.88	1.57
	$c = 2$	2.29	1.82
B-2-220 chi/Q	$c = 1$	2.58	2.48
	$c = 1.5$	3.29	2.99
	$c = 2$	3.66	3.17
B-2-315 chi/Q	$c = 1$	1.53	1.47
	$c = 1.5$	2.23	1.95
	$c = 2$	2.58	2.15
B-3-300 chi/Q	$c = 1$	1.13	1.09
	$c = 1.5$	1.85	1.60
	$c = 2$	2.13	1.87
B-3-600 chi/Q	$c = 1$	1.02	9.84e-1
	$c = 1.5$	1.75	1.45
	$c = 2$	2.07	1.72
B-4-300 chi/Q	$c = 1$	1.68	1.55
	$c = 1.5$	2.19	1.93
	$c = 2$	2.73	2.44
B-4-600 chi/Q	$c = 1$	9.85e-1	9.41e-1
	$c = 1.5$	1.63	1.32
	$c = 2$	2.18	1.78
B-5-600 sig-y	$c = 1$	7.73e-1	6.83e-1
	$c = 1.5$	7.55e-1	6.76e-1
	$c = 2$	7.40e-1	6.52e-1
B-5-600 sig-z	$c = 1$	5.56e-1	4.55e-1
	$c = 1.5$	1.02	7.09e-1
	$c = 2$	1.30	9.83e-1

Table 3.1: *Dispersion example: overview of relative information values of the p.d.f. of Z as determined using Relinf Y and Relinf Z.*

no explanation has been found for this, but it may be due to the way of discretization.

Based on the seed variable questions the calibration scores, based on scoring with individual values (for this example the median values) and sampling distributions, are given in Table 3.3 for B-1-220 chi/Q in case $c = 1$.

The calibration scores in Table 3.3 for $c = 1$ are similar, which is not surprising since the measurement variability in this case is not large (see Table F.1); this will result in similar distributions for Y and Z (see Section 3.1.3), which in turn will result in similar calibration scores. Next the measurement variability is changed; keeping the median values fixed, the error factors EF_{95} are multiplied by $c = 1.5$ and $c = 2$. The results presented in Table 3.3 show that the calibration scores with sampling distributions are decreasing both for Relinf Y and Relinf Z . The rate is different for Relinf Y than for Relinf Z ; the difference between the calibration scores for $c = 1.5$

Seed variable	Method	Quantile assessments for DM		
		$y_{5\%}$	$y_{50\%}$	$y_{95\%}$
B-1-220 chi/Q	Y	5.28e-7	3.46e-5	2.79e-4
	Relinf Y	4.19e-7	2.20e-5	2.77e-4
	Relinf Z	3.14e-7	1.43e-5	2.00e-4
B-1-315 chi/Q	Y	3.31e-6	3.57e-5	1.60e-4
	Relinf Y	3.19e-6	3.43e-5	1.61e-4
	Relinf Z	2.97e-6	3.21e-5	1.69e-4
B-2-220 chi/Q	Y	3.56e-7	4.31e-5	2.39e-4
	Relinf Y	3.33e-7	3.73e-5	2.40e-4
	Relinf Z	2.94e-7	3.39e-5	2.32e-4
B-2-315 chi/Q	Y	2.10e-6	3.44e-5	1.55e-4
	Relinf Y	2.02e-6	3.29e-5	1.56e-4
	Relinf Z	1.89e-6	3.09e-5	1.64e-4
B-3-300 chi/Q	Y	2.15e-6	1.62e-5	9.44e-5
	Relinf Y	2.10e-6	1.59e-5	9.41e-5
	Relinf Z	2.01e-6	1.52e-5	9.17e-5
B-3-600 chi/Q	Y	1.05e-6	6.22e-6	4.95e-5
	Relinf Y	1.02e-6	6.11e-6	4.92e-5
	Relinf Z	9.77e-7	5.82e-6	4.73e-5
B-4-300 chi/Q	Y	3.05e-6	5.15e-5	3.71e-4
	Relinf Y	2.80e-6	4.60e-5	3.67e-4
	Relinf Z	2.37e-6	3.89e-5	3.25e-4
B-4-600 chi/Q	Y	5.30e-6	2.78e-5	1.85e-4
	Relinf Y	5.14e-6	2.72e-5	1.84e-4
	Relinf Z	4.89e-6	2.58e-5	1.77e-4
B-5-600 sig-y	Y	1.51e+1	3.85e+1	1.33e+2
	Relinf Y	1.43e+1	3.86e+1	1.33e+2
	Relinf Z	1.30e+1	4.07e+1	1.35e+2
B-5-600 sig-z	Y	4.04	1.22e+1	2.91e+1
	Relinf Y	5.46	1.93e+1	4.83e+1
	Relinf Z	5.03	2.00e+1	5.66e+1

Table 3.2: *Dispersion example: overview of quantiles for DM for the different methods ($c=1$).*

$c * EF_{95}$	Calibration score	Y	Relinf Y	Relinf Z
$c = 1$	Scoring with median values	6.83e-1	6.83e-1	6.83e-1
	Scoring with sampling distributions	N/A	6.76e-1	6.83e-1
$c = 1.5$	Scoring with median values	6.83e-1	6.83e-1	6.83e-1
	Scoring with sampling distributions	N/A	4.17e-1	2.96e-1
$c = 2$	Scoring with median values	6.83e-1	6.83e-1	6.83e-1
	Scoring with sampling distributions	N/A	3.98e-1	1.56e-1

Table 3.3: *Dispersion example: overview of calibration scores for B-1-220 chi/Q for different error factors EF_{95} .*

and $c = 2$ is 5% for Relinf Y and 47% for Relinf Z. A possible explanation is that the support for unobserved variable Y remains the same for $c = 1.5$ and $c = 2$, however, the support of Z changes since it depends on the measurement variability ϵ , see model 3.2. Additionally, the distributions

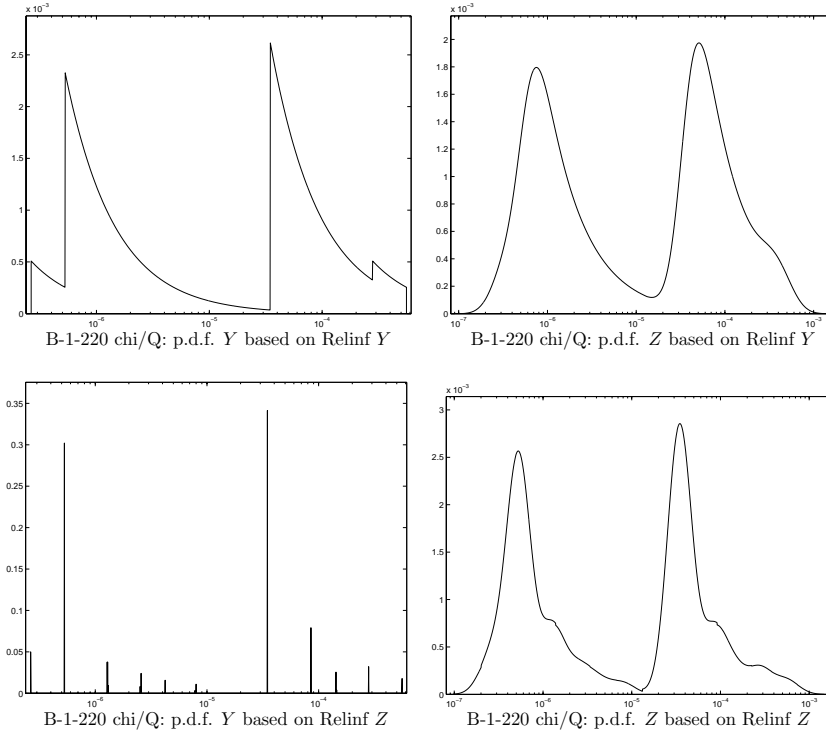


Figure 3.6: *Dispersion example: probability density functions of Y and Z based on Relinf Y and Relinf Z for B-1-220 χ/Q ($c=1$).*

for Z obtained via Relinf Z are more informative than Relinf Y , see Table 3.1. These factors will influence the calibration score.

At this moment it is difficult to draw general conclusions, however it seems that the distributions obtained from Relinf Y are less sensitive to scoring calibration with sampling distributions compared to Relinf Z . This may be because the distributions obtained from Relinf Y are more ‘disperse’ than the distributions obtained from Relinf Z .

Table 3.4 gives an overview of interquartile intervals and conditional probabilities q_{j,l_j} for $c = 1, 1.5$ and 2 .

Since the measurement variability for $c = 1$ was small, it is observed from Table 3.4 that the majority of the uncertain observations are contained in a single interquartile interval, except B-1-220 χ/Q and B-3-600 χ/Q . However, it is observed for $c = 1.5$ and $c = 2$ that the uncertain observations may be contained in different interquartile intervals. Note the change in conditional probabilities between $c = 1.5$ and $c = 2$; the q_{j,l_j} ’s for $c = 2$ are not as ‘peaked’ as for $c = 1.5$. This is in compliance with the intuition; a small q_{j,l_j} for $c = 1.5$ gets larger for $c = 2$ since it is likely that more

Seed variable	Interquantile intervals and q_{j,l_i} for Relif Y and Relif Z		
	Relif Y	Relif Z	
B-1-220 chi/Q	$c = 1$	$\begin{matrix} [I_{1,2}, & I_{1,3}] \\ [1.60e-1, & 8.40e-1] \end{matrix}$	$\begin{matrix} I_{1,3} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{1,2}, & I_{1,3}] \\ [8.84e-2, & 9.12e-1] \end{matrix}$	$\begin{matrix} I_{1,3} \\ 1 \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{1,2}, & I_{1,3}] \\ [9.52e-2, & 9.05e-1] \end{matrix}$	$\begin{matrix} I_{1,3} \\ 1 \end{matrix}$
B-1-315 chi/Q	$c = 1$	$\begin{matrix} I_{2,3} \\ 1 \end{matrix}$	$\begin{matrix} I_{2,3} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{2,2}, & I_{2,3}] \\ [4.45e-2, & 9.56e-1] \end{matrix}$	$\begin{matrix} I_{2,3} \\ 1 \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{2,2}, & I_{2,3}, & I_{2,4}] \\ [7.18e-2, & 9.28e-1, & 3.50e-4] \end{matrix}$	$\begin{matrix} [I_{2,3}, & I_{2,4}] \\ [9.99e-1, & 1.00e-3] \end{matrix}$
B-2-220 chi/Q	$c = 1$	$\begin{matrix} I_{3,2} \\ 1 \end{matrix}$	$\begin{matrix} I_{3,2} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{3,2}, & I_{3,3}] \\ [6.00e-2, & 9.40e-1] \end{matrix}$	$\begin{matrix} I_{3,3} \\ 1 \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{3,2}, & I_{3,3}] \\ [4.70e-2, & 9.53e-1] \end{matrix}$	$\begin{matrix} I_{3,3} \\ 1 \end{matrix}$
B-2-315 chi/Q	$c = 1$	$\begin{matrix} I_{4,3} \\ 1 \end{matrix}$	$\begin{matrix} I_{4,3} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{4,2}, & I_{4,3}] \\ [7.00e-3, & 9.93e-1] \end{matrix}$	$\begin{matrix} I_{4,3} \\ 1 \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{4,2}, & I_{4,3}, & I_{4,4}] \\ [2.12e-2, & 9.76e-1, & 2.91e-3] \end{matrix}$	$\begin{matrix} [I_{4,3}, & I_{4,4}] \\ [9.88e-1, & 1.20e-2] \end{matrix}$
B-3-300 chi/Q	$c = 1$	$\begin{matrix} I_{5,3} \\ 1 \end{matrix}$	$\begin{matrix} I_{5,3} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{5,3}, & I_{5,4}] \\ [9.92e-1, & 8.00e-3] \end{matrix}$	$\begin{matrix} [I_{5,3}, & I_{5,4}] \\ [9.74e-1, & 2.60e-2] \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{5,2}, & I_{5,3}, & I_{5,4}] \\ [1.39e-3, & 9.55e-1, & 4.40e-2] \end{matrix}$	$\begin{matrix} [I_{5,3}, & I_{5,4}] \\ [8.43e-1, & 1.57e-1] \end{matrix}$
B-3-600 chi/Q	$c = 1$	$\begin{matrix} [I_{6,3}, & I_{6,4}] \\ [3.40e-2, & 9.66e-1] \end{matrix}$	$\begin{matrix} I_{6,4} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{6,3}, & I_{6,4}] \\ [4.56e-1, & 5.44e-1] \end{matrix}$	$\begin{matrix} [I_{6,3}, & I_{6,4}] \\ [4.80e-2, & 9.52e-1] \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{6,3}, & I_{6,4}] \\ [5.49e-1, & 4.51e-1] \end{matrix}$	$\begin{matrix} [I_{6,3}, & I_{6,4}] \\ [7.23e-1, & 2.77e-1] \end{matrix}$
B-4-300 chi/Q	$c = 1$	$\begin{matrix} I_{7,2} \\ 1 \end{matrix}$	$\begin{matrix} I_{7,2} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{7,2}, & I_{7,3}] \\ [9.62e-1, & 3.80e-2] \end{matrix}$	$\begin{matrix} [I_{7,2}, & I_{7,3}] \\ [5.81e-1, & 4.19e-1] \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{7,2}, & I_{7,3}] \\ [7.94e-1, & 2.06e-1] \end{matrix}$	$\begin{matrix} [I_{7,2}, & I_{7,3}] \\ [3.31e-1, & 6.69e-1] \end{matrix}$
B-4-600 chi/Q	$c = 1$	$\begin{matrix} I_{8,3} \\ 1 \end{matrix}$	$\begin{matrix} I_{8,3} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{8,2}, & I_{8,3}] \\ [4.72e-2, & 9.53e-1] \end{matrix}$	$\begin{matrix} I_{8,3} \\ 1 \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{8,2}, & I_{8,3}] \\ [1.39e-1, & 8.61e-1] \end{matrix}$	$\begin{matrix} [I_{8,2}, & I_{8,3}] \\ [4.00e-3, & 9.96e-1] \end{matrix}$
B-5-600 sig-y	$c = 1$	$\begin{matrix} I_{9,2} \\ 1 \end{matrix}$	$\begin{matrix} I_{9,2} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} I_{9,2} \\ 1 \end{matrix}$	$\begin{matrix} I_{9,2} \\ 1 \end{matrix}$
	$c = 2$	$\begin{matrix} I_{9,2} \\ 1 \end{matrix}$	$\begin{matrix} I_{9,2} \\ 1 \end{matrix}$
B-5-600 sig-z	$c = 1$	$\begin{matrix} I_{10,2} \\ 1 \end{matrix}$	$\begin{matrix} I_{10,2} \\ 1 \end{matrix}$
	$c = 1.5$	$\begin{matrix} [I_{10,1}, & I_{10,2}, & I_{10,3}] \\ [1.00e-3, & 9.17e-1, & 8.20e-2] \end{matrix}$	$\begin{matrix} [I_{10,2}, & I_{10,3}] \\ [6.67e-1, & 3.26e-1] \end{matrix}$
	$c = 2$	$\begin{matrix} [I_{10,1}, & I_{10,2}, & I_{10,3}, & I_{10,4}] \\ [1.09e-2, & 7.70e-1, & 2.19e-1, & 3.80e-4] \end{matrix}$	$\begin{matrix} [I_{10,1}, & I_{10,2}, & I_{10,3}] \\ [1.84e-3, & 5.46e-1, & 4.53e-1] \end{matrix}$

Table 3.4: Dispersion example: overview of interquantile intervals information and conditional probabilities for $c = 1, 1.5$ and 2 .

realizations in I_{j,l_j} will satisfy the indicator function in Expression 3.10, likewise a large q_{j,l_j} is likely to get smaller.

3.3.2 Ministry of Transport, Public Works and Water Management

The Public Works and Water Management example involves 8 seed variable questions, for a detailed description see [21]. For this example multiple observations/experimental results were available for each seed variable questions, see Table F.2 for experimental results and information on measurement variability distributions.

In total 17 experts were involved in [21]. Results for only 3 experts will be presented: Expert 1, Expert 6 and Expert 10. The quantile assessments of these experts are listed in Table F.3. From this table it is observed that the assessments of Expert 1 express more confidence than the assessments of Expert 6, Expert 10. The effect of folding measurement variability into the distributions of Expert 1 is shown in Tables 3.5 and 3.6: the Y -rows give the quantile assessments of the expert, the Relinf Y and Relinf Z rows give the quantile assessments of Expert 1 when measurement variability is folded in using the Relinf Y and Relinf Z approach, respectively. It is observed that the majority of the 5%-95% quantile intervals of Relinf Y include the 5%-95% quantile intervals of Y (except for HS and TS) and the majority of the 5%-95% quantile intervals of Relinf Z include the 5%-95% quantile intervals of Relinf Y and Y (except for mod1 and mod2).

Seed variable	Expert 1	
	Relinf Y	Relinf Z
ZR	1.35	1.13
ZG	1.33	1.11
ZD	1.34	1.12
HS	9.77e-1	7.61e-1
TS	9.19e-1	7.75e-1
mod1	1.87	1.57
mod2	1.76	1.48
mod3	1.02	9.63e-1

Table 3.5: *Public Works and Water Management example: overview of relative information values for Relinf Y and Relinf Z for Expert 1.*

Next, the use of the two techniques to score calibration of Section 3.2 is explained.

Scoring with individual values

Suppose n_j observations/experimental results are available for the j -th seed variable question, then n_j independent realizations of the j -th seed variable

Seed variable	Method	Quantile assessments for Expert 1				
		$y_{5\%}$	$y_{25\%}$	$y_{50\%}$	$y_{75\%}$	$y_{95\%}$
ZR	Y	9.20e-1	1.03	1.10	1.17	1.28
	Relinf Y	8.38e-1	1.01	1.14	1.29	1.92
	Relinf Z	8.06e-1	9.86e-1	1.13	1.32	2.57
ZG	Y	9.20e-1	1.03	1.10	1.17	1.28
	Relinf Y	7.96e-1	9.99e-1	1.16	1.36	1.96
	Relinf Z	7.63e-1	9.80e-1	1.15	1.39	2.72
ZD	Y	9.20e-1	1.03	1.10	1.17	1.28
	Relinf Y	8.28e-1	1	1.14	1.31	1.92
	Relinf Z	7.95e-1	9.86e-1	1.14	1.34	2.60
HS	Y	6.50e-1	8.50e-1	1	1.20	1.50
	Relinf Y	6.90e-1	9.13e-1	1.10	1.34	2.41
	Relinf Z	6.29e-1	8.63e-1	1.14	1.49	2.91
TS	Y	6.50e-1	8.50e-1	1	1.20	1.50
	Relinf Y	6.94e-1	9.14e-1	1.10	1.34	2.41
	Relinf Z	6.48e-1	8.75e-1	1.12	1.44	2.74
mod1	Y	1.00e-1	3.00e-1	5.00e-1	1	3
	Relinf Y	8.46e-2	2.54e-1	5.02e-1	1.02	3.13
	Relinf Z	6.26e-2	1.71e-1	3.87e-1	7.51e-1	2.96
mod2	Y	1	3	5	1.00e+1	3.00e+1
	Relinf Y	8.62e-1	2.54	5	1.01e+1	3.10e+1
	Relinf Z	6.41e-1	1.72	3.90	7.48	2.90e+1
mod3	Y	1.00e+1	3.00e+1	5.00e+1	1.00e+2	3.00e+2
	Relinf Y	4.68	3.24e+1	5.12e+1	1.13e+2	3.20e+2
	Relinf Z	3.08	3.36e+1	6.56e+1	1.56e+2	4.32e+2

Table 3.6: *Public Works and Water Management example: overview of quantiles for Expert 1 for the different methods.*

will be available. In this way for each observation a copy of the seed variable question can be associated and the calibration is scored with the n_j observed values. For example, if 10 and 15 observations are available for seed variable question Y_1 and Y_2 respectively, then 10 copies of seed variable question Y_1 will be generated and 15 copies for Y_2 . The 10 observations will be scored against the 10 copies of Y_1 and likewise the 15 observations will be scored against the 15 copies of Y_2 . Hence, the calibration score can be considered to be determined on 25 seed variable questions.

Table 3.7 gives the calibration scores for the 3 experts for different situations: the column headed by Y presents the calibration scores as determined in [21], whereas the columns headed by Relinf Y and Relinf Z give the calibration scores when folding the measurement variability into the expert's assessment of elicitation variables Y . In total 47 seed variable questions were generated. Looking at the calibration scores it is observed that the ranking of experts under Relinf Y and Relinf Z has changed with respect to the original ranking under Y ; Expert 1 (the confident expert) is ranked first for Relinf Y and Relinf Z whereas he/she is ranked third for Y . Furthermore, note the difference between calibration scores based on Relinf Y and Relinf Z for Expert 6 and Expert 10.

Experts	Y	Relinf Y	Relinf Z
Expert 1	1.47e-5	2.69e-1	2.41e-1
Expert 6	3.41e-2	1.64e-6	4.37e-10
Expert 10	3.96e-1	2.06e-2	1.65e-5

Table 3.7: *Public Works and Water Management example: overview of calibration scores for 47 observations.*

Scoring with sampling distributions

Suppose n_j observations are available for the j -th seed variable question, then a distribution is determined based on the observations which will represent the measurement variability. For example, suppose 10, 15 observations are available for seed variable question Y_1, Y_2 respectively. Based on the 10 observations available for Y_1 a distribution will be determined which will represent the measurement variability, similarly based on the 15 observations for Y_2 a measurement variability distribution will be determined. Note that the number of seed variable questions remains the same.

Calibration scores are determined based on the original expert assessments Y and based on Relinf Y and Relinf Z in combination with scoring calibration with median values and with sampling distributions.

Experts	Calibration scores	Y	Relinf Y	Relinf Z
Expert 1	Scoring with median values	8.56e-2	8.56e-2	8.56e-2
	Scoring with sampling distributions	N/A	2.85e-1	2.42e-1
Expert 6	Scoring with median values	3.01e-1	3.01e-1	3.01e-1
	Scoring with sampling distributions	N/A	4.99e-1	3.93e-1
Expert 10	Scoring with median values	7.94e-1	7.94e-1	7.94e-1
	Scoring with sampling distributions	N/A	6.36e-1	3.98e-1

Table 3.8: *Public Works and Water Management example: overview of calibration scores for 8 seed variable questions.*

The results in Table 3.8 when scoring with median values are identical for Y , Relinf Y and Relinf Z . Hence, for this example no effect of ‘folding’ measurement variability into the expert assessments is observed when scoring calibration with median values. Looking at the results for scoring with sampling distributions, it is observed that the spread in calibration scores among experts based on Relinf Z is less than the spread based on Relinf Y ; based on Relinf Z the experts are almost equally weighted. The ranking among experts under Y remains the same under Relinf Y and Relinf Z . Furthermore the difference between the calibration scores for Expert 1 determined using the two calibration scores is noted; as mentioned the assessments of Expert 1 were regarded as confident and this resulted in a low calibration score (8.56e-2), however by folding in measurement variability and take it into account when scoring calibration his/her calibration score

has improved and is close to the calibration scores of the other experts. To some extent the reverse holds for Expert 6 and Expert 10. This shows the effect of taking measurement variability into account.

See Tables 3.9 for an overview of interquantile intervals and q_{j,l_j} when scoring calibration with sampling distributions. See Tables F.5 and F.6 for a similar overview of experts 6 and 10.

Scoring with individual values versus Scoring with sample distributions

In scoring calibration, both scoring rules use the information on multiple observations directly or indirectly. Scoring with individual values uses the multiple observations directly to score calibration, whereas scoring calibration with sampling distributions, the observations are used to determine the sampling distribution. Clearly the main difference between the two techniques is the number of seed variable questions; the number of seed variable questions was 47 for scoring with individual values and 8 for scoring with sampling distributions. In order to compare the two techniques, the power of the test ⁴ should be reduced such that the effective number of samples taken from the sampling distributions equals the total number of seed variable questions when scoring with individual values.

A disadvantage of scoring calibration with sampling distributions can be the determination of the sampling distributions. For this example, the number of observation/experimental results available for the seed variable questions ranged from 3 to 8. Hence, it is questionable if a sampling distribution should be determined based on 3 observations/experimental results.

3.4 Conclusion

At this moment it is very difficult to draw any final conclusions. However, it is clear that taking measurement variability into account can make a difference when scoring calibration.

In the dispersion example, the measurement variability ranged from small to large. In case of a small measurement variability, all calibration

⁴In [14]: the power of a statistical test is the ability to distinguish between rival hypotheses, and increases with the number of independent samples. Calibration power may be chosen from the interval $[0,1]$, and determines the effective number of samples. Choosing 50% power means reducing the resolution of the significance test to that of a test with half the number of samples.

Instead of calculating experts calibration with the formula

$$P(R > r) = 1 - \chi_K^2(2nI(\mathbf{p}|\mathbf{q}))$$

the following formula is used:

$$P(R > r) = 1 - \chi_K^2(2\text{Power}nI(\mathbf{p}|\mathbf{q}))$$

where $\text{Power} \in [0,1]$.

Expert 1	
Seed variable	Interquantile intervals which may contain observation and $q_{i,k}$ for Relinf Y and Relinf Z
ZR	Relinf Y
	$[I_{1,1}, I_{1,2}, I_{1,3}, I_{1,4}, I_{1,5}, I_{1,6}]$ [1.13e-3, 1.08e-1, 2.68e-1, 3.41e-1, 2.72e-1, 9.61e-3]
	Relinf Z
	$[I_{1,2}, I_{1,3}, I_{1,4}, I_{1,5}]$ [6.49e-2, 2.64e-1, 3.80e-1, 2.91e-1]
ZG	Relinf Y
	$[I_{2,1}, I_{2,2}, I_{2,3}, I_{2,4}, I_{2,5}, I_{2,6}]$ [1.46e-3, 9.54e-2, 2.29e-1, 3.19e-1, 3.23e-1, 3.17e-2]
	Relinf Z
	$[I_{2,2}, I_{2,3}, I_{2,4}, I_{2,5}]$ [6.42e-2, 2.20e-1, 3.30e-1, 3.86e-1]
ZD	Relinf Y
	$[I_{3,1}, I_{3,2}, I_{3,3}, I_{3,4}, I_{3,5}, I_{3,6}]$ [1.31e-4, 3.86e-2, 1.59e-1, 3.11e-1, 4.35e-1, 5.63e-2]
	Relinf Z
	$[I_{3,2}, I_{3,3}, I_{3,4}, I_{3,5}, I_{3,6}]$ [1.43e-2, 1.16e-1, 2.58e-1, 6.09e-1, 2.41e-3]
HS	Relinf Y
	$[I_{4,3}, I_{4,4}, I_{4,5}]$ [1.19e-1, 8.28e-1, 5.36e-2]
	Relinf Z
	$[I_{4,4}, I_{4,5}]$ [9.69e-1, 3.1e-2]
TS	Relinf Y
	$[I_{5,3}, I_{5,4}]$ [4.31e-1, 5.69e-1]
	Relinf Z
	$[I_{5,3}, I_{5,4}]$ [9.8e-2, 9.02e-1]
mod1	Relinf Y
	$[I_{6,2}, I_{6,3}, I_{6,4}]$ [1.50e-1, 7.66e-1, 8.41e-2]
	Relinf Z
	$[I_{6,3}, I_{6,4}]$ [7.53e-1, 2.47e-1]
mod2	Relinf Y
	$[I_{7,3}, I_{7,4}, I_{7,5}]$ [1.30e-1, 8.18e-1, 5.26e-2]
	Relinf Z
	$[I_{7,4}, I_{7,5}]$ [7.34e-1, 2.66e-1]
mod3	Relinf Y
	$I_{8,4}$ 1
	Relinf Z
	$I_{8,4}$ 1

Table 3.9: *Public Works and Water Management example: overview of interquantile information for Expert 1.*

scores were similar. For larger measurement variabilities the effect between scoring calibration with median values and sampling distributions was observed. In the Public Works and Water Management example the effect of measurement variability was also observed. The calibration score of the expert, regarded as confident, increased by a few orders of magnitude and the ranking of the expert calibration scores changed.

In conclusion, the effect of folding measurement variability into expert assessments is apparent. However it is difficult at this stage to recommend Relinf Y or Relinf Z in combination with the scoring calibration using individual values or sampling distributions. However, the difference between the distribution of the (unobserved) variable Y from Relinf Y and Relinf Z is very significant and in case of Relinf Z points a weakness of the relative information principle. Furthermore, the observations/experimental results available may be very scarce that it is questionable if an 'appropriate' sampling distribution, reflecting measurement variability, can be determined. These observations may be reasons to prefer Relinf Y above Relinf Z and to score calibration with individual values above scoring it with sampling distributions. However, more research needs to be done to give better recommendations.

Chapter 4

Modeling Uncertainty

Enough real data will eventually prove that any physical model, with any particular choice of model input parameters, will be incorrect and thus inconsistent with reality. Usage of models in this way does not deal with the uncertainty associated with the physical process. Hence, instead of developing physical models which produce predictions, it is better to focus on developing models which are able to capture the uncertainty. The use of mathematical models to capture uncertainty rather than to make predictions requires experts and decision makers to think about these models in new and different ways, see [38].

Modeling uncertainty is the topic of this chapter, with special attention to the capture of uncertainty in acyclic compartmental models (ACMs). The attention will be restricted to ACMs encountered in the Joint CEC/USNRC Uncertainty Analysis and therefore some remarks/conclusions will not hold in general. The type of ACMs addressed in this chapter may be characterized as ACMs with constant transfer coefficients, for which very little/imprecise data on the transfer coefficients are available and which can be represented by a set of first order linear differential equations. The ACM of the systemic retention of Sr in the human body is used to illustrate different concepts/features. Finally, since ACMs bear great resemblance to influence diagrams, the relationship between them is investigated.

4.1 How to quantify uncertainties?

The Sr model of Figure 4.1 is a typical example of an ACM used by international bodies charged with setting standards for radiation exposure for the general public and for radiological workers [31]. The reader is referred to [2] for a detailed description of this problem. This ACM is used to compute Sr dose-coefficients for certain organs which express risk from being exposed to Sr.

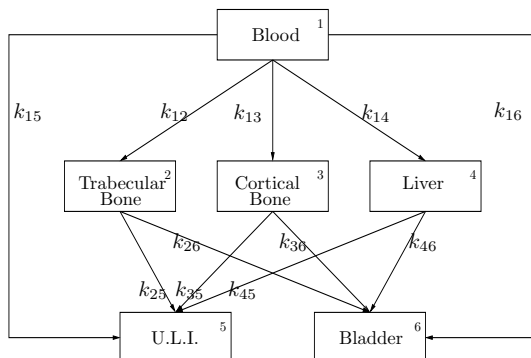


Figure 4.1: *Systemic retention of Sr: acyclic compartmental model.*

The compartmental models discussed here make strong assumptions, in particular,

1. The compartments and only these compartments are involved in the transfer of material.
2. The rates of transfer from ‘source’-compartment to a ‘sink’-compartment are proportional to the amount of material in the ‘source’-compartment, and independent of all other physical variables.

The compartmental model itself is not derived from underlying physical laws, nor can it be verified by direct observation. Most of the transfer coefficients cannot be measured directly by experiments. The uncertainty analyst is tasked with quantifying the uncertainty attending the use of such models in a traceable and defensible way. If compartmental models were derived from accepted physical laws, and if the transfer coefficients could be measured, subject to measurement variability, then the quantification of uncertainty would be straightforward. The transfer coefficients would be regarded as drawn from a sampling distribution reflecting measurement variability and the uncertainty attending the use of such models would be obtained by propagating the sampling distribution through the model.

The above remarks make it clear that this straightforward method of quantifying uncertainty is not available for compartmental models. The method by which these models are chosen and quantified cannot form the basis of a quantification of uncertainty. Indeed, no generic method was encountered for choosing and subsequently quantifying such models. The type of arguments leading to a choice of a given model are peculiar to the species in question. Once a compartmental model is chosen, the method for determining the values of the transfer coefficients is also highly specific to the problem at hand and involves a great deal of qualitative reasoning.

The absence of direct physical measurements of transfer coefficients means that the uncertainty cannot be determined by objective statistical

methods, rather the relevant uncertainty takes the form of subjective uncertainty of experts. The uncertainty must be quantified using Structured Expert Judgment Elicitation Methodology. At the same time, the lack of validation for the models themselves entails that it is not possible to simply ask the experts ‘express your uncertainty in transfer coefficient k_{25} of the Sr compartmental model’, as no assumption should be made regarding the model the experts use.

4.2 Uncertainty Capture

Physical models like ACMs are traditionally used with ‘best estimates’ of the transfer coefficients to predict phenomena. When models cannot be inferred from accepted laws and the values of the transfer coefficients cannot be measured, then the predictions of models are uncertain. Straightforward use of the model with ‘best estimates’ does not give any picture of the uncertainty attending model predictions. It is suggested here that these models can be employed legitimately to capture uncertainty. This employment differs in fundamental ways from straightforward prediction.

Capturing uncertainty in observable phenomena via a distribution over transfer coefficient involves: (i) using Structured Expert Judgment Elicitation Methodology to quantify uncertainty on measurable quantities predicted by the model; (ii) performing probabilistic inversion to pull this uncertainty back onto transfer coefficients of the model; (iii) comparing the uncertainty pushed through the model with the uncertainty expressed by the experts.

If there is an imperfect fit in Step (iii), the conclusion is not that the model is wrong, rather the conclusion is that the expert’s uncertainty cannot be captured via a distribution over its transfer coefficients. Here, three reasons for this will be discussed.

Firstly, although experts believe that the model is ‘roughly right’, their uncertainty may involve departures from the assumptions of the model. Thus with regard to Figure 4.1, experts may believe recirculation may occur from, for example compartments Cortical Bone to Blood: under certain circumstances, a portion of material in the compartment Cortical Bone may be transferred back to the compartment Blood. In this case the amount transferred to the compartment Cortical Bone in a unit time would not be proportional to the amount in the blood. It may be impossible to capture the experts’ uncertainty via a distribution over the transfer coefficients of Figure 4.1.

Secondly, although the experts each represent their uncertainties via distributions over the parameters of the ICRP-67 model for Sr [31], it may be impossible to represent their combined distribution in this way, taking account of physical constraints. For example, suppose each expert believes in the ACM of Figure 4.1 and believes that the assumption between random

variables k_{12} and k_{13} , for some constant tc ,

$$k_{12} = tc * k_{13} \quad (4.1)$$

is correct, but they do not agree on the value of tc . If Equation 4.1 is interpreted as a physical constraint with tc a constant, then it may be impossible to capture the combined expert distribution via a distribution over transfer coefficients satisfying Equation 4.1, even though this is possible for each expert individually.

Thirdly, the mathematical processing itself may impose simplifications which cause a discrepancy in Step (iii). From Table 4.1 it can be seen that representing the distribution over transfer coefficients as a minimum relative information distribution with respect to the product distribution under marginal distribution and Spearman's rank correlation constraints introduces significant discrepancies; the assessments in the column DM are the combined expert assessments based on equal weights and the information in the column PREJUDICE is the push-forward of the distribution on the transfer coefficients as determined by PREJUDICE. Finally the column headed by Marg. Rank contains information on the push-forward of the distribution over transfer coefficients represented as a minimum relative information distribution with respect to the product distribution under marginal distribution and Spearman's rank correlation constraints.

Assuming that the fit in Step (iii) is good, the use of ACMs to capture uncertainty may involve features which are unfamiliar to experts and decision makers alike, and which deserve special attention.

First of all, the distribution for the DM will not agree in general with the distribution of any one expert. Typically, the uncertainty in the distributions are obtained by linear pooling the uncertainties of several experts, will be larger than the uncertainties of each individual expert.

Secondly, the distribution on transfer coefficients may involve strong correlations, either positive or negative, which complicate the ways experts traditionally think about the models. Thus experts tend to think of transfer coefficients in terms of retention half-times; looking at Figure 4.1, if the compartments Cortical Bone and ULI are considered in isolation, then the time at which half of a unit deposit to Cortical Bone is transferred to ULI is equal to $\frac{\log 2}{k_{35}}$, and is called the retention half-time for Cortical Bone. Similarly, $\frac{\log 2}{k_{45}}$ is the retention half-time for Liver. These expressions suggest that k_{35} and k_{45} have a meaning independent of the model in which they are considered. This is not the case however as may become glaringly evident when k_{35} and k_{45} are assigned distributions as given in Table 1.8 under the columns Run A, Run B and Extension. Note that from Tables 1.9 through 1.11 these variables have a relatively strong positive correlation. A 'representative value' for k_{35} (e.g. the median) together with a 'representative value' for k_{45} may not yield representative values for simple functions of (k_{35}, k_{45}) . Consider the following simple example: X and Y are uniformly

Time		Skel+Liver		
		DM	PREJUDICE	Marg. Rank
1 day	5%	1.70e-1	1.30e-1	1.20e-1
	50%	3.24e-1	3.24e-1	3.10e-1
	95%	5.76e-1	5.76e-1	5.30e-1
1 week	5%	1.17e-1	1.17e-1	1.10e-1
	50%	2.29e-1	2.29e-1	2.60e-1
	95%	4.76e-1	4.76e-1	5.50e-1
1 month	5%	1.04e-1	1.04e-1	5.80e-2
	50%	2.11e-1	2.11e-1	2.00e-1
	95%	3.51e-1	3.51e-1	4.90e-1
1 year	5%	6.74e-2	6.65e-2	1.10e-2
	50%	1.38e-1	1.38e-1	1.80e-1
	95%	2.43e-1	2.43e-1	4.30e-1
10 years	5%	1.81e-2	1.80e-2	4.80e-3
	50%	6.45e-2	6.45e-2	1.10e-1
	95%	1.37e-1	1.37e-1	2.50e-1
50 years	5%	1.11e-3	1.10e-3	2.10e-4
	50%	1.85e-2	1.85e-2	1.30e-2
	95%	8.88e-2	8.87e-2	7.40e-2
Time		$\frac{\text{Skel}}{\text{Skel+Liver}}$		
		DM	Pred.	Marg. Rank
1 day	5%	8.46e-1	8.45e-1	7.70e-1
	50%	9.56e-1	9.56e-1	9.80e-1
	95%	9.98e-1	9.98e-1	9.99e-1
1 week	5%	8.22e-1	8.20e-1	6.70e-1
	50%	9.57e-1	9.56e-1	9.70e-1
	95%	9.98e-1	9.98e-1	9.98e-1
1 month	5%	8.51e-1	8.16e-1	8.50e-1
	50%	9.84e-1	9.56e-1	9.58e-1
	95%	9.99e-1	9.98e-1	9.99e-1
1 year	5%	7.70e-1	7.70e-1	7.70e-1
	50%	9.94e-1	9.60e-1	9.60e-1
	95%	9.99e-1	9.98e-1	9.99e-1
10 years	5%	6.79e-1	6.77e-1	6.89e-1
	50%	9.95e-1	9.87e-1	9.79e-1
	95%	9.99e-1	9.99e-1	9.99e-1
50 years	5%	6.39e-1	6.40e-1	6.40e-1
	50%	9.96e-1	9.90e-1	9.90e-1
	95%	9.99e-1	9.99e-1	9.99e-1

Table 4.1: *Systemic retention of Sr: Quantile information comparison of distributions of DM vs. push-forward distributions based on PREJUDICE and marginal distributions and Spearman’s rank correlation matrix.*

distributed on $[0,2]$ and completely negative correlated, so that $Y = 2 - X$. The median of X and Y is 1. Hence the product of the medians is 1, but 1 is also the *maximum* of XY ; in other words the product of the medians is not the median of the product.

And finally, if uncertainty over observable phenomena can be captured via a distribution over transfer coefficients, then this can in general be cap-

tured in more than one way. In other words, if the probabilistic inverse of a distribution over observables exists, then it is generally not unique; two uncertainty analysts using different search algorithms, different heuristics and a measure other than the relative information measure might come up with different distributions over the transfer coefficients in Figure 4.1, both of them adequately reproducing the uncertainty over observable phenomena.

4.3 ACMs and influence diagrams

The complexity of some compartmental models is sometimes formidable. This necessitated the development of new techniques to enable tractable probabilistic inversion. Influence diagrams are helpful in extracting conditional dependency structures embedded in ACMs. The main assumption linking ACMs to influence diagrams is that functional conditional independence entails statistical conditional independence. At this moment no numerical results are available, however the technique will be illustrated qualitatively.

4.3.1 Example

Consider acyclic compartmental model ACM-I as shown in Figure 4.2. ACM-I will be used to illustrate the different steps which led to the new solution scheme.

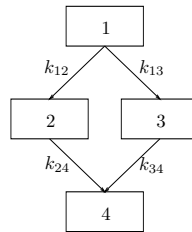


Figure 4.2: *Acyclic compartmental model ACM-I.*

For ACM-I, the target variables are the transfer coefficients k_{ij} which describe the movement of material from compartment i to compartment j , and which is considered to be constant within a short time period. Based on Figure 4.2, a set of first order linear differential equations can be constructed which, with the appropriate initial conditions, fully specifies the movement between compartments. Let $m_i(t)$ represent the amount of material contained in compartment i at time t , furthermore let $\mathbf{k} = (k_{12}, \dots, k_{34})$ and $F_{\mathbf{k}}$ represent the distribution on \mathbf{k} . The aim is to determine $F_{\mathbf{k}}$. Note that the uncertain quantities in Figure 4.2 are $m_1(t), \dots, m_4(t)$ and k_{12}, \dots, k_{34} . As explained before, almost all transfer coefficients do not satisfy the criteria

which an elicitation variable has to satisfy, and therefore in many cases the elicitation variables will be formulated on the amount of material retained at different times in various compartments.

Like influence diagrams, ACMs are directed acyclic graphs, but ACMs are not influence diagrams. For example, if ACM-I were an influence diagram the following statement would be true (\perp denotes statistical independence).

$$m_2(t) \perp m_3(t) \quad \text{given } m_1(t)$$

However, based on the equations describing ACM-I it is easy to see this statement is false. Briefly, the relevant equations for compartments 1, 2 and 3 are, starting at $t = 0$ with a unit deposit in compartment 1:

$$m_1(t) = e^{-(k_{12}+k_{13})t} \tag{4.2}$$

$$m_2(t) = k_{12} \frac{e^{-k_{24}t} - e^{-(k_{12}+k_{13})t}}{k_{12} + k_{13} - k_{24}} \tag{4.3}$$

$$m_3(t) = k_{13} \frac{e^{-k_{34}t} - e^{-(k_{12}+k_{13})t}}{k_{12} + k_{13} - k_{34}}. \tag{4.4}$$

Since the equations describing $m_2(t)$ and $m_3(t)$ are not functionally independent given $m_1(t)$, they cannot be statistically conditional independent; any choice of k_{12} fully specifies k_{13} , because the sum $k_{12} + k_{13}$ is given.

However, for fixed $t_0, t_1 > 0$ it follows from Equations 4.2, ..., 4.4 that $\frac{m_2(t)}{m_2(t_0)}$ and $\frac{m_3(t)}{m_3(t_1)}$ given $m_1(t)$ are functionally independent. Assuming that functional conditional independence entails statistical conditional independence, the conditional independence statement among elicitation variables can be formulated

$$\frac{m_2(t)}{m_2(t_0)} \perp \frac{m_3(t)}{m_3(t_1)} \quad \text{given } m_1(t).$$

Proposition 4.3.1 relates the conditional independence statement among elicitation variables to a conditional independence statement among target variables.

Proposition 4.3.1. *Considering compartmental model ACM-I. For all $t_0, t_1 > 0$ (though fixed), the conditional independence statement among elicitation variables*

$$\frac{m_2(t)}{m_2(t_0)} \perp \frac{m_3(t)}{m_3(t_1)} \quad \text{given } m_1(t) \text{ for all } t$$

implies the conditional independence statement among target variables

$$k_{24} \perp k_{34} \quad \text{given } k_{12} + k_{13}.$$

Proof: It suffices to show the existence of monotone mappings f and g , such that $k_{24} = f^{-1}\left(\frac{m_2(t)}{m_2(t_0)}\right)$ and $k_{34} = g^{-1}\left(\frac{m_3(t)}{m_3(t_1)}\right)$, respectively. Hence it is sufficient to prove that the derivatives of $\frac{m_2(t)}{m_2(t_0)}$ with respect to k_{24} and $\frac{m_3(t)}{m_3(t_1)}$ with respect to k_{34} are either strictly positive or negative.

The proof will focus on showing strict positivity for the derivative for $\frac{m_2(t)}{m_2(t_0)}$ with respect to k_{24} only, since the same line of reasoning can be applied for showing strict positivity or negativity for the derivative for $\frac{m_3(t)}{m_3(t_1)}$ with respect to k_{34} .

Since the denominator of $\frac{d}{dk_{24}}\left(\frac{m_2(t)}{m_2(t_0)}\right)$ is always positive, the attention is focused on the numerator. Let $\alpha = e^{-k_{24}}$ and $\beta = \frac{e^{-(k_{12}+k_{13})}}{\alpha}$, then the numerator of $\frac{d}{dk_{24}}\left(\frac{m_2(t)}{m_2(t_0)}\right)$ can be written as

$$\frac{d}{dk_{24}}\left(\frac{m_2(t)}{m_2(t_0)}\right)_{\text{numerator}} = \alpha^{t_0+t} \left((t_0 - t) - (t_0(\beta)^t - t(\beta)^{t_0}) \right) \quad (4.5)$$

Because of $\frac{d}{dk_{24}}\left(\frac{m_2(t)}{m_2(t_0)}\right)_{\text{numerator}} = -\frac{d}{dk_{24}}\left(\frac{m_2(t_0)}{m_2(t)}\right)_{\text{numerator}}$, the case $t_0 > t$ is considered only. Since $\beta = 1$ causes the numerator of Equation 4.3 to be zero, it will not be considered¹.

It will be shown that $\frac{d}{dk_{24}}\left(\frac{m_2(t)}{m_2(t_0)}\right) > 0$ for all $\beta > 0$, $\beta \neq 1$. Since $\alpha^{t_0+t} > 0$ always holds, it suffices to show

$$h(\beta) := (t_0 - t) - (t_0\beta^t - t\beta^{t_0}) > 0.$$

Write

$$h(\beta) = (t_0 - t) - (t_0 \exp(t \log \beta) - t \exp(t_0 \log \beta))$$

and take the derivative $\frac{dh(\beta)}{dk_{24}}$:

$$\begin{aligned} \frac{dh(\beta)}{dk_{24}} &= -\frac{t_0 t}{\beta} \beta^t + \frac{t_0 t}{\beta} \beta^{t_0} \\ &= \frac{t_0 t}{\beta} \beta^t (\beta^{t_0-t} - 1). \end{aligned} \quad (4.6)$$

Since $t_0 - t > 0$, it can be concluded from Equation 4.6 that if $\beta > 1$ then $\frac{dh(\beta)}{dk_{24}}(\beta) > 0$ and if $\beta < 1$ then $\frac{dh(\beta)}{dk_{24}}(\beta) < 0$. This result in combination with $h(1) = 0$, it is concluded that $h(\beta) > 0$ for all $\beta > 0$, $\beta \neq 1$.

Since the derivative $\frac{d}{dk_{24}}\left(\frac{m_2(t)}{m_2(t_0)}\right)$ attains strictly positive values in case $\beta \neq 1$, a mapping f exists such that $k_{24} = f^{-1}\left(\frac{m_2(t)}{m_2(t_0)}\right)$. The above

¹If $\beta = 1$ it means that the rate with which material is exiting box 1, $k_{12} + k_{13}$, is equal to the rate with which the material is leaving box 2, k_{24} . This situation is regarded as non-realistic.

analysis can be performed on $\frac{m_3(t)}{m_3(t_1)}$ as well in order to prove the existence of a mapping g such that $k_{34} = g^{-1}\left(\frac{m_3(t)}{m_3(t_1)}\right)$. The existence of f and g enable us to formulate the conditional independence statements among the target variables

$$k_{24} \perp k_{34} \quad \text{given} \quad k_{12} + k_{13}.$$

□

Proposition 4.3.1 leads to relating ACMs and influence diagrams. The influence diagram of compartmental model ACM-I is given in Figure 4.3.

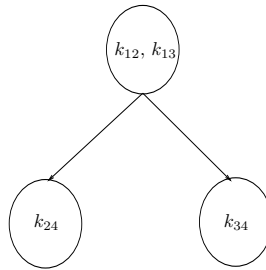


Figure 4.3: *The influence diagram of ACM-I.*

The influence diagram of ACM-I can be used to infer conditional independence statements among the target variables \mathbf{k} . From Figure 4.3, it is easy to see conditional independence between (sets of) transfer coefficients. Using these conditional independence relationships, $F_{\mathbf{k}}$ can be written as

$$F_{\mathbf{k}} = F_{k_{24}|k_{12},k_{13}} F_{k_{34}|k_{12},k_{13}} F_{k_{12},k_{13}} \tag{4.7}$$

By making use of the conditional independence embedded in the ACM the original problem of determining a four-dimensional distribution is reduced to determining 2 three-dimensional distributions and one two-dimensional problem. The conditional independence statements are helpful in reducing the dimension of the original probabilistic inversion problem.

In case of ACM-I, the following solution scheme is suggested for determining the distribution on \mathbf{k} :

Step 1 Construct the influence diagram of ACM-I.

Step 2 Determine distribution $F_{k_{12},k_{13}}$.

Step 3 Determine distribution $F_{k_{24},k_{12},k_{13}}$ with marginal $F_{k_{12},k_{13}}$.

Step 4 Determine distribution $F_{k_{34},k_{12},k_{13}}$ with marginal $F_{k_{12},k_{13}}$.

Step 5 Combine the distributions $F_{k_{12},k_{13}}$, $F_{k_{24},k_{12},k_{13}}$ and $F_{k_{34},k_{12},k_{13}}$ to obtain $F_{\mathbf{k}}$.

It is recommended to determine the distributions in **Step 2** through **Step 4** using the iterative version of PREJUDICE as described in Section 1.4.

4.3.2 Systemic retention of Pu in the human body

The ACM given in Figure 4.4 describes the systemic retention of Pu in the human body and was encountered in the Internal Dosimetry Panel of the Joint CEC/USNRC Uncertainty Analysis [28]. For a detailed description

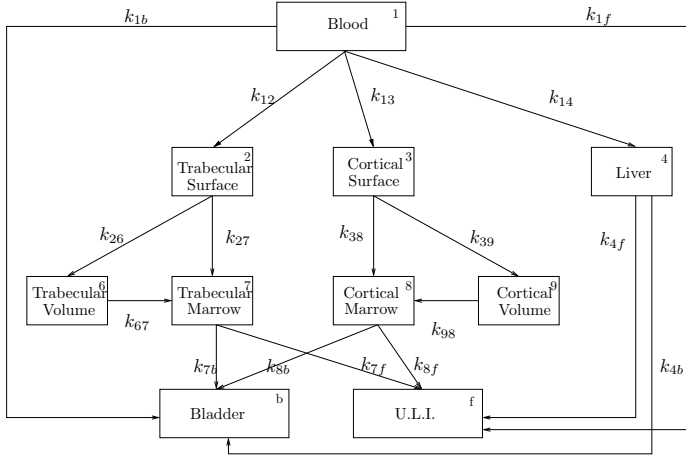


Figure 4.4: Systemic retention of Pu: acyclic compartmental model.

on this probabilistic inversion problem, see [2]. This ACM was simplified even more by assuming that the transfer coefficient to the Bladder can be written as

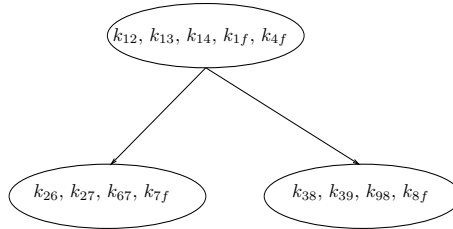
$$k_{ib} = UF_{pu} * k_{if} \quad \text{with } i \in \{1, 4, 7, 8\} \quad (4.8)$$

where UF_{pu} represents the ‘Bladder to ULI’-ratio for PU. Hence, the target variable space of the probabilistic inversion problem for the systemic retention in the Human Body consists of 13 dimensions.

Using Proposition 4.3.1 the influence diagram based on Figure 4.4 is given in Figure 4.5. Based on this influence diagram the distribution $F_{\mathbf{k}_{pu}}$ can be written as

$$F_{\mathbf{k}_{pu}} = F_{k_{26}, k_{27}, k_{67}, k_{7f}} | k_{12}, k_{13}, k_{14}, k_{1f}} F_{k_{38}, k_{39}, k_{98}, k_{8f}} | k_{12}, k_{13}, k_{14}, k_{1f}} F_{k_{12}, k_{13}, k_{14}, k_{1f}, k_{4f}}$$

Note that the original problem of determining a distribution on 13 target variables has been reduced to finding two 8 dimensional problems and one 5 dimensional problem. The solution scheme based on the influence diagram, would be similar to the solution scheme as presented for the example.

Figure 4.5: *The influence diagram of Pu-model.*

4.4 Cyclic Compartmental Models

Compartmental models are widely used in different fields. Most of these compartmental models are not acyclic, they include recirculation processes. In performing an uncertainty analysis, the aim is not to predict the amount of a quantity in different compartments, but to predict the uncertainty in these amounts. By definition, models are simplified representations of reality, hence it may be that an acyclic compartmental model may predict the uncertainty as well as a cyclic compartmental model. Hence, in performing uncertainty analysis on compartmental models it may be worthwhile to explore the possibility to perform the uncertainty analysis on an acyclic compartmental model closely related to the original cyclic compartmental model.

4.5 Implications to Methodology

Set in an expert judgment context, the relation between acyclic compartmental models and influence diagrams will determine the structuring of questioning to a large extent. It could very well determine the choice of elicitation variables as well. However, it should not be the driving force. The driving force in determining elicitation variables should be the extent to which the elicitation variable satisfies the criteria of being potentially observable. Recall that under the Structured Expert Judgment Elicitation Methodology, the experts are selected, among other issues, to allow for a diversity of viewpoints. A consequence of this selection procedure could be that some experts use different models. Designing the elicitation solely on the influence diagram of the ACM could result in elicitation questions which are specific for the ACM. This approach may lead to a situation in which the expert doesn't feel comfortable; he is being queried on specific model quantities which are not familiar to him. In performing an expert judgment study, situations in which the expert feels uncomfortable have to be avoided at any time.

For these reasons, the uncertainty analyst has to design the elicitation in such a way that he obtains information on quantities which can be con-

sidered as potentially measurable and model independent, and such that he can use the information in all steps of the solution scheme.

4.6 Conclusions

The use of models to capture uncertainty involves mathematical and conceptual problems. Experts and decision makers need to consider what purpose a model serves. If enough experimental data are available, any model can be qualified, with any particular choice of model input parameters, as incorrect. A more relevant question is to investigate the extent to which a model is able to capture the uncertainty associated.

In performing an uncertainty analysis over a model, the first question an uncertainty analyst should answer is to what extent the model can be simplified and still capture the uncertainty. Unless the uncertainty analyst is a competent modeler, this task should be performed in close cooperation with users/developers of the model on which the uncertainty analysis will be conducted. The uncertainty analysis will then be performed on the simplified model. In this chapter the attention is focused on ACMs; two ACMs were used to illustrate the construction of a simplified model. In case the simplified model requires probabilistic inversion, the current probabilistic inversion techniques cannot handle models much larger than the compartmental model describing the systemic retention of Pu in the human body, see Figure 4.4. Future research will focus on developing new solution algorithms (like the relationship between acyclic compartmental models and influence diagrams) and new heuristics for dealing with more complex models. To date the choice of elicitation variables has been driven by the Structured Expert Judgment Elicitation Methodology and model, but in the future may be driven also by specific solution schemes. However, elicitation variables must be always familiar quantities to the experts, model independent and their number must remain relatively small.

Chapter 5

Conclusions

The aim of this thesis has been the development of new mathematical techniques which support the Structured Expert Judgment Elicitation Methodology. The mathematical techniques Probabilistic Inversion and the elicitation of dependencies have been successfully applied in the Joint CEC/USNRC Uncertainty Analysis, [19] and [21], whereas more research is needed on the scoring of calibration when confronted with measurement variability and the effect of the relationship between acyclic compartmental models and influence diagrams. Results for the latter 2 research areas are preliminary at this stage, however encouraging and promising.

The reader is referred to the conclusions section at the end of each chapter for conclusions for the respective chapters. Here possible improvements and connections between the different subjects will be discussed.

Probabilistic inversion

It is common that given a model and some experimental results, values for model input parameters are sought which would minimize the Sum of Squares distance between model outcome and experimental results. Probabilistic inversion can be seen as an extension of this approach; the experimental results are considered uncertain, which results in a distribution on the model input parameters.

A quote concerning the use of probabilistic inversion is given in Chapter 7 of [19]:

”..probabilistic inversion is described as a powerful tool to quantitatively verify whether the selected model refinement is adequate in view of uncertainty in the process, which the model aims to describe. However it is costly in terms of computation time and in its current form it requires a skilled operator. Hence the technique is not suitable in the context of design practice.”

The quote is valid, however in [19] the version of PREJUDICE was used in which every sample is an optimization variable was used. At the time neither the iterative and/or efficient version of PREJUDICE were available. The efficient version of PREJUDICE requires less computation time, however probabilistic inversion remains a time consuming task. If models of similar complexity as the dispersion coefficient example (with a relative small number of target variables) are used in the design practice, then probabilistic inversion is suitable. If the computer implementation of PREJUDICE is done professionally, then solving the dispersion coefficient example using the iterative, efficient version of PREJUDICE can be done within 30 minutes.

Another point of criticism in Section 5.9 of [19] concerns the generation of physically acceptable scenarios. For example, in case of the dispersion coefficient example for stability class C the scenario

$$\mathbf{y} = (33, 64.8, 175, 448, 8250) \quad (5.1)$$

is physically acceptable. At first glance this scenario may seem reasonable; the lateral plume spreads are increasing with down wind distance. However it consists of the 5%-iles for 500 m. through 10 km and the 95%-ile of 30 km. This means that within stability class C, plumes exist which are narrow up to 10 km and then expand dramatically up to 30 km. It may be questioned whether scenarios with similar extreme behavior should be included in the set of potentially observable scenarios, hence the potentially observable scenarios should be checked to what extent they are realistic. A way to reduce the number of extreme scenarios is to account for dependence among the different elicitation variables. So far this has not been done, but it is not difficult to imagine, that if a strong positive correlation would have been specified among the lateral plume spreads at the down wind distances, scenario 5.1 could not occur.

Currently the implementation of Probabilistic inversion requires an interior point method in order to determine the minimum relative information distribution on the target variables, however in [8] a procedure based on the Iterative Proportional Fitting (IPF) procedure has been developed, which converges to the maximum entropy distribution if the PI problem formulation is feasible. See [15],[46] for details on iterative proportional fitting procedures. In case of infeasibility a ‘PARFUMized’ version of IPF can be shown to converge to a distribution minimizing (see Expression 1.17 for notation)

$$nH(\tilde{\mu}_k) - \sum_{j=1}^n H(\mu_{k|j}) \quad (5.2)$$

where $H(\mu)$ represents the entropy of μ .

Dependencies

The chosen dependence elicitation technique (conditional probability), proved to be easily understandable by the experts and they acknowledged its importance. Understanding the dependence elicitation technique in terms of frequency allows the experts to think of it in terms of scenarios. During an elicitation session it is easy to assist the expert without explaining mathematical concepts.

Currently, the treatment of dependencies inherent in the distribution on the target variables obtained from probabilistic inversion in an uncertainty analysis is considered to be crude. Usually the distribution over the model input parameters used in an uncertainty analysis is represented as marginal distributions with a rank correlation matrix. However, as seen in Section 2.3 it may be difficult to capture the dependence structure resulting from the probabilistic inversion in terms of rank correlations. It seems counter-intuitive to spend a large portion of time in the determination of a high dimensional distribution and then summarize it crudely by marginal distributions and a rank correlation matrix. A consequence is that the push-forward results based on the marginal distributions and rank correlation are not as good as the push-forward results of the high dimensional distribution, see Table 2.7. It is recommended to use/develop more advanced sampling techniques which are capable of sampling from a high dimensional distributions. For example the Acceptance-Rejection sampling scheme. A disadvantage of the Acceptance-Rejection sampling scheme is that it is not optimal in case the difference between the largest and smallest probability assigned to the sample vectors is large.

Calibration with uncertain observations

In [4] it is stated in the Sponsor Perspective that:

However, if an outlier interpretation persists, it is our firm belief - in agreement with the SSHAC¹ - that the approaches outlined will allow for essential down-weighting of that interpretation. This is preferable to the stiff adherence to an equal weighting scheme, which can result in final seismic hazard being driven by a single outlier input

Under the Structured Expert Judgment Elicitation Methodology all experts are considered as sources of data. Based on seed variable questions the expert's capacity to quantify uncertainty is measured, which will be used to assign weights to the experts. In this way experimental results determine if an expert's opinion should be down-weighted instead of the uncertainty analyst, which may not be as objective.

¹Senior Seismic Hazard Analysis Committee

Project staff of the joint CEC/USNRC questioned to what extent measurement variability might affect the capacity of experts to quantify the uncertainty. Eventhough more research is needed it was concluded that measurement variability makes a difference compared to not taking it into account. Still the number, quality and variety of seed variable questions and quantification of measurement variability are issues which have to be addressed for each study involving expert judgment.

Modeling Uncertainty

The intention of this chapter was to show the effect and considerations on how to reduce the complexity of a model before performing uncertainty analysis. A special class of compartmental models were used to illustrate this process. Based on compartmental models it may be possible to obtain a acyclic version which describes the uncertainty associated with the physical process as well. Since acyclic compartmental models and influence diagrams show great resemblance, it may be investigated whether conditional independence statements may be inferred to reduce the complexity even more. It has been shown that by formulating elicitation questions which describe the retention relative to the retention at a certain time point t , conditional independence among elicitation variables leads to conditional independence among target variables. The time point t to which the retention is considered may be t_{\max} ; the time point at which the maximum level of retention is attained, in formula $m(t_{\max}) = \max_t m(t)$.

Application of results

Finally, in performing an uncertainty analysis the mathematical techniques developed in this thesis have to be used in the reverse order as presented. Firstly, the model on which the uncertainty analysis will be performed has to be investigated if it can be simplified using the considerations of Chapter 4. Based on the simplified model the target variables are identified, upon which the elicitation variables will be based. If data are scarce, experts are queried on the marginal distribution of the elicitation variables and the dependence among them. Based on these results, the expert's calibration is scored using the techniques developed in Chapter 3, thus, if available, taking account of measurement variability. Based on the expert's calibration scores the marginal distributions and dependence information are aggregated using Strategy 2 of Chapter 2. If necessary, probabilistic inversion is performed on certain elicitation variables to obtain the distribution on target variables. The iterative, efficient version of PREJUDICE, introduced in Chapter 1 taking account of dependence is recommended and in generating the set of potentially observable scenarios the dependencies among the elicitation variables are used to avoid physically unrealistic scenarios. The final result will be a joint distribution on all relevant model input parameters

of the model, which will be propagated through the model using a certain representation of the joint distribution or a more sophisticated sampling scheme. The final step of the uncertainty analysis will be the analysis of the model output and to investigate which uncertain model input parameters influence the model output uncertainty the most.

Bibliography

- [1] G. Apostolakis and S. Kaplan. Pitfalls in risk calculations. *Reliability Engineering & System Safety*, 2:135–145, 1981.
- [2] L.H.J. Goossens B.C.P. Kraan, R.M. Cooke. Joint cec/usnrc accident consequence code uncertainty analysis using expert judgment; probabilistic inversion results. Technical report, Delft University of Technology, Delft, The Netherlands, 2002.
- [3] P. Billingsley. *Probability and Measure (second edition)*. John Wiley & Sons Inc., New York, 1986.
- [4] R.J. Budnitz, G. Apostolakis, D.M. Boore, L.S. Cluff, K.J. Copper-smith, C.A. Cornell, and P.A. Morris. Recommendations for probabilistic seismic hazard analysis: Guidance of uncertainty and use of experts: main report. Technical Report NUREG/CR-6372, United States Nuclear Regulatory Commission, 1997.
- [5] R.T. Clemen, G.W. Fischer, and R.L. Winkler. Assessing dependencies: Some experimental results. Available at www.duke.edu/~clemen.
- [6] R.T. Clemen and T. Reilly. Correlations and copulas for decision and risk analysis. *Management Science*, 45:208–224, 1999.
- [7] R. Cooke. Expert resolution. In *Proceedings of the 2nd IFAC conference on Analysis Design and Evaluation of Man-Machine Systems*. Pergamon Press, 1985.
- [8] R. Cooke and D. Kurowicka. Conditional rank correlations for the elliptic copulae: Dependence modelling in uncertainty analysis. In *Proceedings of ESREL 2001*, 2001. To appear.
- [9] R.M. Cooke. *Experts in Uncertainty: Opinion and Subjective Probability in Science*. Oxford University Press, New York, 1991.
- [10] R.M. Cooke. Parameter fitting for uncertain models; modeling uncertainty in small models. *Reliability Engineering & System Safety*, 44:89–102, 1994.

- [11] R.M. Cooke. *UNICORN; Methods and Code for Uncertainty Analysis*. AEA Technology, Warrington, Cheshire (UK), 1995.
- [12] R.M. Cooke and L.H.J. Goossens. Procedures guide for structured expert judgment. Technical Report EUR 18820 EN, European Commission, Directorate-General for Research, Brussels, Belgium, 2000.
- [13] R.M. Cooke, M. Mendel, and W. Thys. Calibration and information in expert resolution; a classical approach. *Automatica*, 24:87–94, 1988.
- [14] R.M. Cooke and D. Solomatine. *EXCALIBR: The User's Manual*. Delft University of Technology, 1993.
- [15] I. Csiszar. i -divergence geometry of probability distributions and minimization problems. *Annals of Probability*, 3(1):146–158, 1975.
- [16] A. Dawid. The well-calibrated bayesian. *Journal of the American Statistical Association*, 77(379):605–613, 1982.
- [17] A. Dawid. Calibration-based empirical probability. *Annals of Statistics*, 13(4):1251–1274, 1985.
- [18] M. de Groot and S. Feinberg. The comparison and evaluation of forecasters. *The Statistician*, 32:12–22, 1983.
- [19] M.S. de Wit. *Uncertainty analysis on Thermal comfort using expert judgement*. PhD thesis, Delft University of Technology, 2001.
- [20] R.M. Dudley. *Real Analysis and Probability*. Wadsworth Inc., Pacific Groove, California, 1989.
- [21] M. Frijters, R.M. Cooke, K. Slijkhuis, and J. van Noortwijk. Expertmeningen onzekerheidsanalyse; kwalitatief en kwantitatief verslag (in Dutch). Technical Report WB1083, Ministry of Transport, Public Works and Water Management, Utrecht, The Netherlands, 1999.
- [22] D. Gokhale and S. Press. Assessment of a prior distribution for the correlation coefficient in a bivariate normal distribution. *Journal of the Royal Statistical Society*, 145(2):237–249, 1982.
- [23] L.H.J. Goossens, R.M. Cooke, and B.C.P. Kraan. Methods for cec/usnrc accident consequence uncertainty analysis of dispersion and deposition; performance-based aggregating of expert judgments and parfum method for capturing modelling uncertainty. Technical Report EUR 15856 EN, European Commission, Directorate-General Science, Research and Development, Brussels, Belgium, 1994.
- [24] L.H.J. Goossens, B.C.P. Kraan, and R.M. Cooke. Evaluation of weighting schemes for expert judgement studies. Technical report, Delft University of Technology, Delft, The Netherlands, 1996.

- [25] F.T. Harper, L.H.J. Goossens, J. Boardman, M.L. Young, B.C.P. Kraan, R.M. Cooke, S.C. Hora, and J.A. Jones. Joint USNRC/CEC consequence uncertainty study: Summary of objectives, approach, application, and results for the deposited material and external doses assessments. Technical Report NUREG/CR-6526 and EUR 16772 and SAND97-2323, Sandia National Laboratories and Delft University of Technology, Washington DC, United States of America, 1997.
- [26] F.T. Harper, L.H.J. Goossens, J. Brown, F.E. Haskin, B.C.P. Kraan, M.L. Abbott, R.M. Cooke, M.L. Young, S.C. Hora, J.A. Jones, and A. Rood. Joint USNRC/CEC consequence uncertainty study: Summary of objectives, approach, application, and results for the food chain uncertainty assessment. Technical Report NUREG/CR-6523 and EUR 16771 and SAND97-0335, Sandia National Laboratories and Delft University of Technology, Washington DC, United States of America, 1997.
- [27] F.T. Harper, L.H.J. Goossens, R.M. Cooke, J.C. Helton, S.C. Hora, J.A. Jones, B.C.P. Kraan, C. Lui, M.D. McKay, L.A. Miller, J. Päsler-Sauer, and M.L. Young. Joint USNRC/CEC consequence uncertainty study: Summary of objectives, approach, application, and results for the dispersion and deposition uncertainty assessments. Technical Report NUREG/CR-6244 and EUR 15855 and SAND94-1453, Sandia National Laboratories and Delft University of Technology, Washington DC, United States of America, 1994.
- [28] F.T. Harper, L.H.J. Goossens, J.D. Harrison, B.C.P. Kraan, R.M. Cooke, and S.C. Hora. Joint USNRC/CEC consequence uncertainty study: Summary of objectives, approach, application, and results for the internal dosimetry uncertainty assessments. Technical Report NUREG/CR-6571 and EUR 16773 and SAND98-0119, Sandia National Laboratories and Delft University of Technology, Washington DC, United States of America, 1998.
- [29] F.T. Harper, L.H.J. Goossens, F.E. Haskin, B.C.P. Kraan, and J.B. Grupa. Joint USNRC/CEC consequence uncertainty study: Summary of objectives, approach, application, and results for the early health effects uncertainty assessments. Technical Report NUREG/CR-6545 and EUR 16775 and SAND97-2689, Sandia National Laboratories and Delft University of Technology, Washington DC, United States of America, 1997.
- [30] F.T. Harper, L.H.J. Goossens, M.P. Little, C.R. Muirhead, B.C.P. Kraan, R.M. Cooke, and S.C. Hora. Joint USNRC/CEC consequence uncertainty study: Summary of objectives, approach, application, and results for the late health effects uncertainty assessments. Technical Report NUREG/CR-6555 and EUR 16774 and SAND97-2322, Sandia

- National Laboratories and Delft University of Technology, Washington DC, United States of America, 1997.
- [31] ICRP. Age-dependent doses to members of the public from intakes of radionuclides. Technical Report ICRP Publication 67, International Committee of Radiation Protection, Elsevier Science Ltd., Oxford, 1994.
- [32] B. Jansen. *Interior point techniques in optimization*. PhD thesis, Delft University of Technology, 1995.
- [33] B.C.P. Kraan and T.J. Bedford. Foundations and application of probabilistic inversion in uncertainty analysis. To appear.
- [34] B.C.P. Kraan and R.M. Cooke. Dealing with dependencies in uncertainty analysis. In C. Cacciabue and I.A. Papazoglou, editors, *Proceedings of Probabilistic Safety Assessment and Management '96*, pages 625–630. Springer, 1996.
- [35] B.C.P. Kraan and R.M. Cooke. The effect of correlations in uncertainty analysis: two cases. In R.M. Cooke, editor, *Technical Committee Uncertainty Modelling: Benchmark Workshop*, 1997.
- [36] B.C.P. Kraan and R.M. Cooke. Post-processing techniques for the joint cec/usnrc uncertainty analysis of accident consequence codes. *Journal of Statistical Computation and Simulation*, 71:253–268, 1997.
- [37] B.C.P. Kraan and R.M. Cooke. Processing expert judgement in accident consequence modelling. *Radiation Protection Dosimetry; Expert Judgement and Accident Consequence Uncertainty Analysis*, 90(3):311–316, 2000.
- [38] B.C.P. Kraan and R.M. Cooke. Uncertainty in compartmental models for hazardous material- a case study. *Journal of Hazardous Material on Risk Analysis*, 57:243–260, 2000.
- [39] W. Kruskal. Ordinal measures of association. *Journal of the American Statistical Association*, 53:814–861, 1958.
- [40] J.P. Kullback. *Information Theory and Statistics*. Wiley, New York, 1959.
- [41] S. Lichtenstein, B. Fischhoff, and L. Phillips. Calibration of probabilities: the state of the art to 1980. *Judgment under Uncertainty, Heuristics and Biases*, pages 306–328, 1982.
- [42] A. McClelland and F. Bolger. The calibration of subjective probabilities: theories and models 1980-1994. *Subjective Probability*, 1994.

- [43] A.M.H. Meeuwissen and R.M. Cooke. Tree dependent random variables (revised version). Technical Report TUD-REPORT 94-28 and ISSN 0922-5641, Delft University of Technology, Delft, The Netherlands, 1994.
- [44] C. Roos and T. Terlaky. Non-linear optimization. Lecture notes (wi3031), Delft University of Technology, 1999.
- [45] C. Roos, T. Terlaky, and J.-Ph Vial. *Theory and Algorithms for Linear Optimization. An Interior Approach*. John Wiley & Sons, Chichester, 1997.
- [46] L. Rüschemdorf. Convergence of the iterative proportional fitting procedure. *Annals of Statistics*, 23(4):1160–1174, 1995.
- [47] R. Winkler. Scoring rules and the evaluation of probability assessors. *Journal of the American Statistical Association*, 64:1073–1078, 1969.
- [48] H. Yamashita. A globally convergent primal-dual interior point method for constraint optimization. Technical report, Mathematical Systems Institute Inc., Tokyo, Japan, 1992.

Appendix A

ϵ -neighborhood vs. Bin combinations

In this appendix the sensitivity of ϵ for the ϵ -neighborhood sampling scheme and the number of bins for the Bin-combination sampling scheme of PREJUDICE is investigated.

For all examples discussed in this appendix the set of 50 model inversions for A_y and B_y are equal to the model inversions of the dispersion coefficient example of Section 1.2.3. The number of samples equals 900.

A.1 ϵ -neighborhood sampling scheme

In **Step 1: ϵ -neighborhood** of Section 1.2.2, an ϵ_i may be chosen for each target variable X_i , $i \in \{1, \dots, m\}$. In this section, the choice of ϵ_i is the same for all target variables X_i . The ϵ -neighborhood sampling scheme has been applied to the dispersion coefficient example for $\epsilon = 0.05, 0.1, 0.25, 0.5$.

Table A.1 gives information on the distribution on the target variables for the different ϵ 's.

The column headed by $\epsilon = 0.05$ contains N/A, because no distribution on the propagated samples, which complies with the DM distribution, could be determined, i.e. this PI problem was infeasible. The relative information value increases when ϵ increases, which is in accordance with the intuitive understanding of relative information; with respect to the same background density, a density which is concentrated on an area A receives a smaller relative information value compared to a density concentrated on an area B , when $A \subset B$. No results on the push-forward distributions are given since they agreed with the quantile information of DM distributions, as presented in Table 1, perfectly. Figure A.1 compares the marginal distributions of A_y and B_y graphically.

For particular choices of ϵ , certain regions of the target variables space

		ϵ -neighborhood			
Quantile		$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.25$	$\epsilon = 0.50$
A_y	5%	N/A	1.66e-2	1.54e-2	2.24e-2
	50%	N/A	5.35e-1	3.83e-1	4.19e-1
	95%	N/A	5.63e+1	5.61	3.48
B_y	5%	N/A	5.84e-1	5.56e-1	5.87e-1
	50%	N/A	8.33e-1	8.75e-1	8.63e-1
	95%	N/A	1.27	1.27	1.19
ρ_{A_y, B_y}		N/A	-9.57e-1	-9.74e-1	-8.99e-1
Relative Information		N/A	3.76e-1	7.82e-1	1.44

Table A.1: Dispersion coefficient example (stability class C): quantile information, Spearman's rank correlations for different ϵ 's.

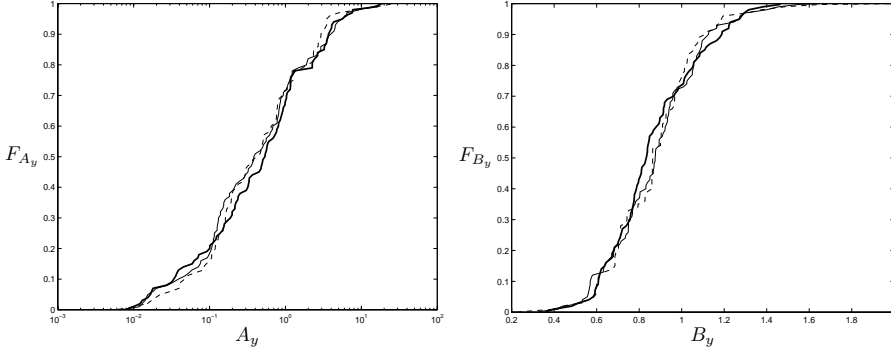


Figure A.1: Dispersion coefficient example (stability class C): graphical comparison of marginal distributions of A_y (left) and B_y (right) for $\epsilon = 0.1$ (—), $\epsilon = 0.25$ (--) and $\epsilon = 0.5$ (- -).

will not be considered. Looking at Figure A.1, flat or nearly flat sections of the distribution function may/can reflect no or very few samples are contained in the corresponding interval of the target variable. This effect is clear for target variable A_y for $\epsilon = 0.1$ and becomes less as ϵ becomes larger.

Loosely speaking, from Figure A.2 it can be observed that $M_{\epsilon=0.1} \subset M_{\epsilon=0.25} \subset M_{\epsilon=0.5}$. This observation is reflected by an increasing relative information value for increasing ϵ .

In conclusion, small values of ϵ could result in unsuccessful probabilistic inversion (see $\epsilon = 0.05$). In case of successful probabilistic inversion larger values of ϵ tend to result in more 'smooth' distribution function for the target variables.

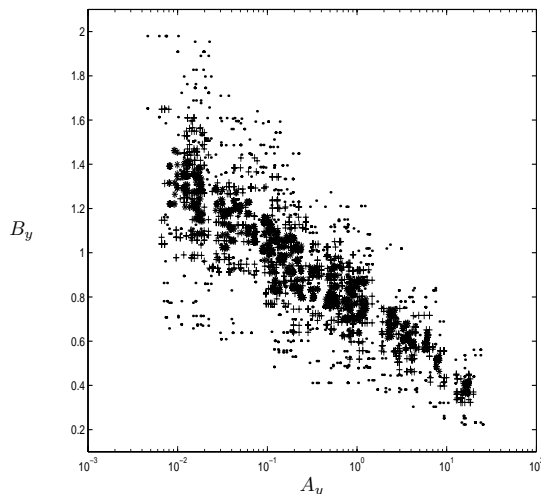


Figure A.2: *Dispersion coefficient example (stability class C): graphical comparison of domain M for $\epsilon = 0.1$ (*), $\epsilon = 0.25$ (+) and $\epsilon = 0.5$ (·).*

A.2 Bin-combination sampling scheme

In **Step 1: Bin combinations** of Section 1.2.2, the number of bins C_i considered for each target variable X_i may be different. Similar to the ϵ -neighborhood sampling scheme, the number of bins considered for each target variable were the same. The bin combination sampling scheme has been applied to the dispersion coefficient example for $C=10, 25, 50$.

In general, the sizes of the bins will be different. At this moment, the determination of the sizes of the bins is ad-hoc; the size of a bin is based on grouping model inversions. Figure A.3 presents samples and bins for $C = 10$ and $C = 25$.

Table A.2 contains information on the distribution on the target variables for the different C 's.

The relative information of the distribution on the target variables shows a tendency to increase when the number of bins decrease. Figure A.4 compares the marginal distributions for A_y and B_y graphically. Note the difference between the marginal distributions for A_y and B_y for the different C 's. This difference may be due to $M_{C=50} \not\subset M_{C=25} \not\subset M_{C=10}$ for the bin combination sampling scheme, see Figure A.5.

In conclusion; at this moment the determination of the size of a bin is done in an ad-hoc manner. It is shown that the bin combination sampling scheme is sensitive to the size and number of bins for each target variable, see Figure A.4.

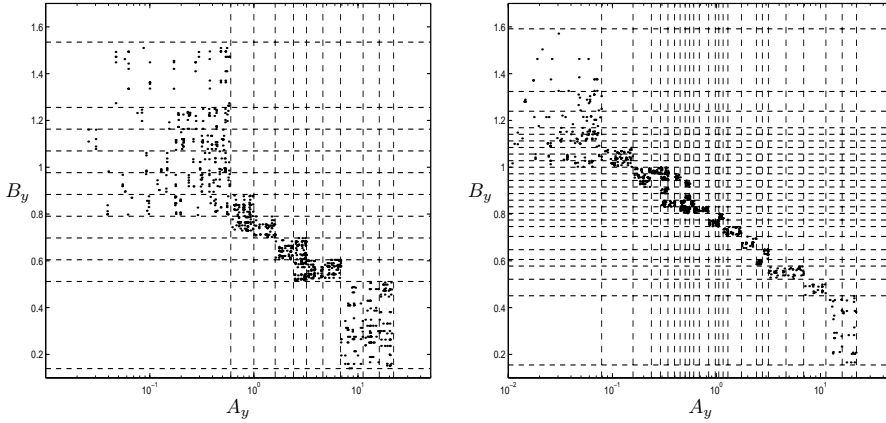


Figure A.3: *Dispersion coefficient example (stability class C): graphical representation of samples and bins for $C = 10$ (left) and $C = 25$ (right).*

		Bin combinations		
A_y	Quantile	$C = 10$	$C = 25$	$C = 50$
	5%	5.72e-2	2.93e-2	1.32e-2
	50%	7.12e-1	4.09e-1	3.71e-1
	95%	5.32	7.77	1.17e+1
B_y	5%	5.46e-1	4.99e-1	4.43e-1
	50%	8.17e-1	8.53e-1	8.66e-1
	95%	1.11	1.24	1.46
ρ_{A_y, B_y}		-9.22e-1	-9.63e-1	-9.76e-1
Relative Information		7.23e-1	1.55e-1	2.06e-1

Table A.2: *Dispersion coefficient example (stability class C): quantile information, Spearman's rank correlations for different bin combinations.*

A.3 Conclusion

The differences between the two sampling scheme can be summarized as: (i) using the ϵ -neighborhood sampling scheme, the sample regions¹ depend on the individual model inversions and ϵ , whereas in case of the bin combinations sampling scheme the sample regions depend on the spread of all model inversions; (ii) for particular choices of ϵ 's, certain regions of the target variable space will not be considered, whereas they are considered using the bin combinations sampling scheme.

It is recommended to use the ϵ -neighborhood sampling scheme instead of the bin combination sampling scheme for two reasons:

1. The results on the distributions based on ϵ -neighborhood sampling

¹By sample region, the Cartesian product of intervals, obtained from the sampling schemes, in the target variable space is meant.

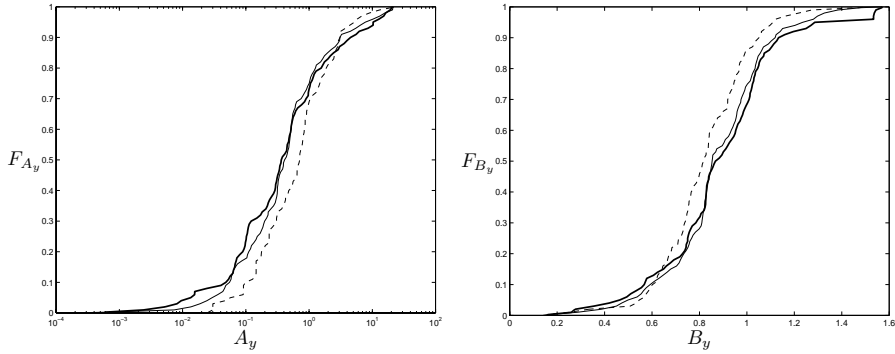


Figure A.4: *Dispersion coefficient example (stability class C): graphical comparison of marginal distributions of A_y (left) and B_y (right) for $C = 10$ (---), $C = 25$ (—) and $C = 50$ (—).*

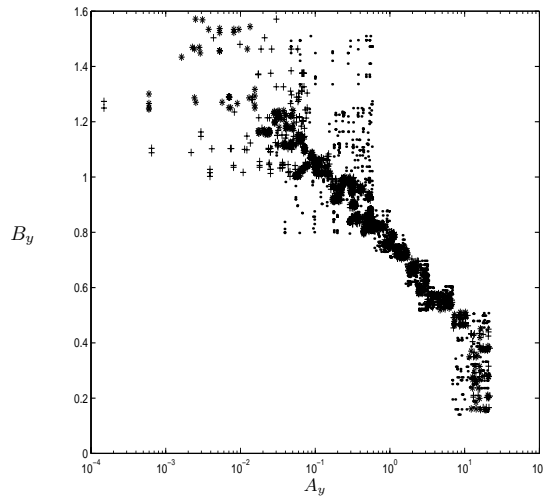


Figure A.5: *Dispersion coefficient example (stability class C): graphical comparison of domain M for $C = 10$ (*), $C = 25$ (+) and $C = 50$ (·).*

scheme (provided that the corresponding PI problem is feasible) suggest they are less sensitive to the choice of ϵ compared to the distributions based on the bin combination sampling scheme, which are regarded to be sensitive regarding the choice on the size and number of bins.

2. The intuitive understanding of the ϵ -neighborhood sampling scheme is easier compared to the bin combination solution scheme; the size of the sample region for a model inversion is given by ϵ , whereas the size

of sample regions in the bin-combination sampling scheme depends on the spread of all model inversions. At this moment there is no rule on how to select the ‘optimal’ size of bins for each target variable.

Appendix B

Determination of M .

The crucial step in probabilistic inversion is the determination of domain M . Since mapping T is not to be assumed invertible, the determination of M is based on determining model inversions \mathbf{x}_y , for which the Sum of Squares distance between the corresponding model reproducible observations and potential observations is minimal, see Expression 1.19. The mapping T under consideration may be very complex (non-linear, non-convex etc.), hence the optimization routines used to minimize the Sum of Squares distance may influence the determination of M .

The results in [2] have been derived using the `minimax` optimization routine of the optimization toolbox of Matlab4.2c; at the time, the most convenient optimization routine available. Since then new versions of Matlab have been released with new optimization toolboxes, whose performance is better than the optimization toolbox of Matlab4.2c. Besides using the optimization routine `lsqcurvefit` of the Matlab5.3 version 11 optimization toolbox, the Minos5 solver¹ of GAMS (General Algebraic Modeling System) was used.

Three problems from the Joint CEC/USNRC Uncertainty Analysis are considered; the dispersion coefficient example taken from the Dispersion & Deposition panel, lung morbidity from the Early Health Effects panel and systemic retention of Sr in the human body from the Internal Dosimetry panel. For each problem, for the same set of potentially observable scenarios, the model inversions \mathbf{x}_y are determined twice; once using `lsqcurvefit` and once using Minos5. Next, for both optimization routines the marginal distributions and rank correlation matrices are determined and compared to each other and to the results derived using Matlab 4.2c using the `minimax` optimization routine.

¹The Minos5 solver was chosen because it is considered to perform the best in case of non-constraint nonlinear optimization problems.

B.1 Dispersion coefficient example

For a description of the model, identification of target variables and formulation of elicitation variables the reader is referred to the Introduction or [2].

B.2 Results: Step 1

Based on the quantile information of the DM and physical considerations, the set of potentially observable scenarios N was determined. Next 50 scenarios were selected randomly from N . Using `lsqcurvefit` and `Minos5`, these 50 scenarios were used to determine model inversions A_y and B_y ; the starting values for A_y and B_y were the same for both optimization routines for all scenarios. Figure B.1 presents the model inversions A_y and B_y as determined using `lsqcurvefit` (+) and `Minos5` (o), for all scenarios; the model inversions for both optimization routines are identical².

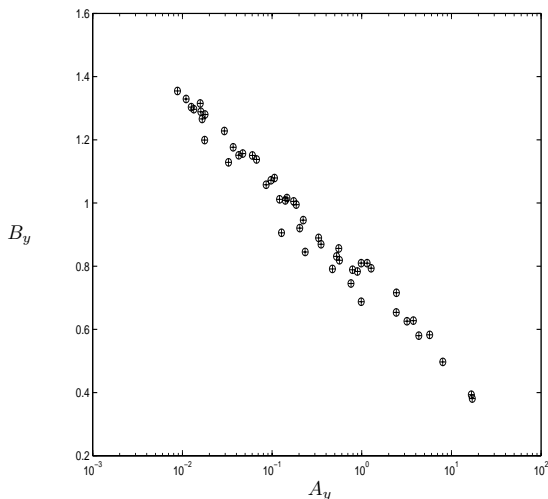


Figure B.1: *Dispersion coefficient example (stability class C): model inversions for 50 scenarios, `lsqcurvefit` (+) and `Minos5` (o).*

Figure B.2 compares the Sum of Squares distances resulting from the two optimization routines; every dot can be related to a scenario for which the Sum of Squares distance has been calculated based on model inversions determined using `lsqcurvefit` and `Minos5`.

²The symbol \oplus corresponds to same model inversions for `lsqcurvefit` and `Minos5`.

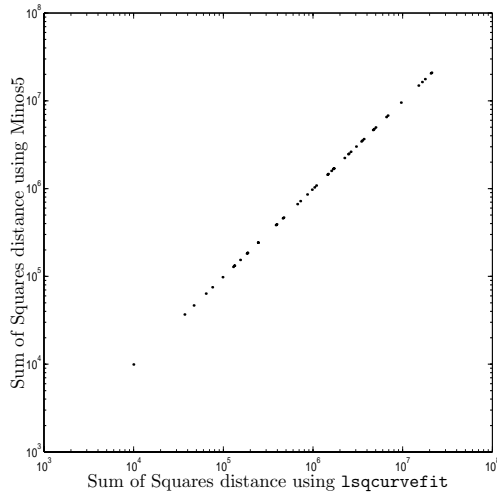


Figure B.2: *Dispersion coefficient example (stability class C): Sum of Squares distance comparison between `lsqcurvefit` and `Minos5` for 50 scenarios.*

B.2.1 Results: Step 2

From Figure B.1 it is observed that the model inversions for the 50 scenarios are the same. Hence, the distribution on (A_y, B_y) based on `lsqcurvefit` should be similar to the distribution based on `Minos5`. Table B.2.1 contains quantile information and Spearman's rank correlations of the distributions on A_y and B_y based on 900 samples. Due to the small number of samples there is some variation in the marginal distributions. The dispersion coefficient example was no probabilistic inversion problem itself for the Dispersion & Deposition panel, it was part of a larger and more complex probabilistic inversion problem see [2]; the information in the column `minimax` of Table B.2.1 is based on $1.19e+5$ samples. Comparing the marginal distributions of A_y and B_y based on `minimax` to the two optimization routines, it is observed that the marginal distributions for A_y and B_y based on `minimax` are a little wider and Spearman's rank correlation somewhat more negative.

A graphical comparison between distributions of A_y and B_y for the three optimizing routines is given in Figure B.3: `lsqcurvefit` (---), `Minos5` (-) and `minimax` (—).

Since the dispersion coefficient example involves two target variables only, the domain M based on `lsqcurvefit` and `Minos5` can be presented graphically, see Figure B.4. It is observed that both domains are specified roughly on the same region in the target variable space.

Table B.2.1 contains quantile information of the DM for the elicitation variables, the push-forward distributions based on `lsqcurvefit`, `Minos5`

	Quantile	lsqcurvefit	Minos5	minimax
A_y	5%	2.24e-2	1.86e-2	1.68e-2
	50%	4.19e-1	5.12e-1	2.39e-1
	95%	3.48	4.57	1.08e+1
B_y	5%	5.87e-1	5.43e-1	4.67e-1
	50%	8.63e-1	8.58e-1	9.08e-1
	95%	1.19	1.22	1.24
ρ_{A_y, B_y}		-8.99e-1	-8.69e-1	-9.67e-1

Table B.1: *Dispersion coefficient example (stability class C): quantile information and Spearman's rank correlations on target variables for three optimization routines.*

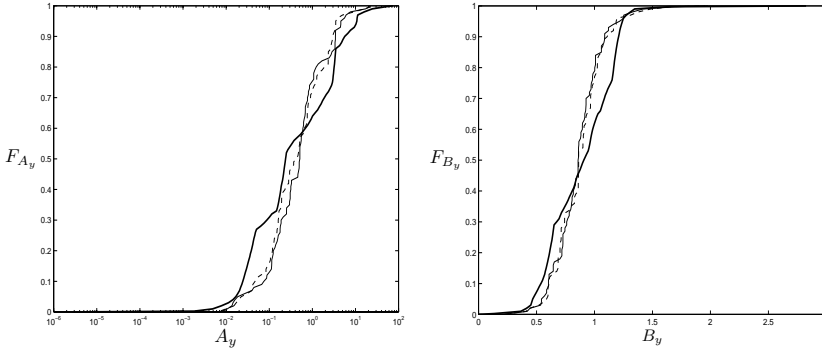


Figure B.3: *Dispersion coefficient example: graphical display of distribution of A_y (left), B_y (right). lsqcurvefit (---), Minos5 (—) and minimax (—).*

and `minimax`. The push-forward distributions based on `lsqcurvefit` and `Minos5` are based on 900 samples, whereas the push-forward distribution based on `minimax` is determined using 1.19×10^5 samples. Even with a small sample size of 900, the push-forward distributions based on `lsqcurvefit` and `Minos5` are considered to agree with the DM distributions rather well. The push-forward distributions based on `minimax` agree with the DM distributions very well also.

B.2.2 Discussion of results

The influence of optimization routines in determining the domain M for the dispersion coefficient example is considered to be very small. This probabilistic inversion problem is qualified as simple, since the powerlaw function can be transformed to a log-linear model. The model inversions for 50 scenarios were the same for both optimizing routines, which resulted in distributions on (A_y, B_y) being quite similar. Compared to the results used in the uncertainty analysis of COSYMA, it was concluded that these re-

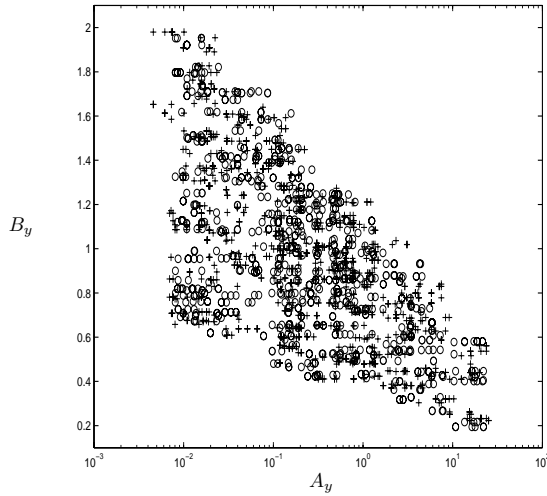


Figure B.4: Dispersion coefficient example (stability class C): graphical comparison between domain M as determined using `lsqcurvefit` (+) and `Minos5` (o).

Distance		$\sigma_y(z_i)$			
		DM	<code>lsqcurvefit</code>	<code>Minos5</code>	<code>minimax</code>
500 m.	5%	3.30e+1	3.38e+1	3.34e+1	3.30e+1
	50%	9.49e+1	9.55e+1	9.18e+1	9.49e+1
	95%	1.95e+2	1.94e+2	1.94e+2	1.95e+2
1 km.	5%	6.48e+1	6.57e+1	6.58e+1	6.48e+1
	50%	1.72e+2	1.74e+2	1.71e+2	1.72e+2
	95%	3.46e+2	3.41e+2	3.44e+2	3.46e+2
3 km.	5%	1.75e+2	1.77e+2	1.78e+2	1.75e+2
	50%	4.46e+2	4.53e+2	4.47e+2	4.46e+2
	95%	1.04e+3	1.03e+3	1.03e+3	1.04e+3
10 km.	5%	4.48e+2	4.49e+2	4.53e+2	4.48e+2
	50%	1.22e+3	1.26e+3	1.24e+3	1.22e+3
	95%	3.37e+3	3.36e+3	3.33e+3	2.89e+3
30 km.	5%	1.10e+3	1.11e+3	1.11e+3	1.10e+3
	50%	2.82e+3	2.83e+3	2.88e+3	2.82e+3
	95%	8.25e+3	8.17e+3	8.16e+3	8.25e+3

Table B.2: Dispersion coefficient example (stability class C): quantile information comparison between DM distributions vs. push-forward distributions based on `lsqcurvefit`, `Minos5` and `minimax` [m.].

sults are in-line with the results as presented in this section. Finally, the push-forward distributions of all optimization routines agreed with the DM distributions rather to very well.

B.3 Lung morbidity

For a more detailed description on the probabilistic inversion problem for lung morbidity, the reader is referred to [2].

B.3.1 Model

In case a population is exposed to an external gamma dose ED [Gy], received with dose rate DR [Gy/hr], the percentage of the exposed population which will suffer respiratory-functional morbidity, represented by $R_{Lung,MB}$, is modeled in COSYMA³ as

$$R_{Lung,MB} = 1 - e^{-\left(\ln(2)\left(\frac{ED}{D_{\infty,Lung,MB} + \frac{D_{0,Lung,MB}}{DR}}\right)^{\nu_{Lung,MB}}\right)} \quad (\text{B.1})$$

for $ED \geq ED_{thr,Lung}$ and where $D_{0,Lung,MB}$ [Gy²/hr], $D_{\infty,Lung,MB}$ [Gy] and $\nu_{Lung,MB}$ are the model parameters, and $ED_{thr,Lung}$ the threshold dose. Note, model B.1 is non-linear and strictly increasing in ED .

B.3.2 Target variables & Elicitation variables

The target variables are $(D_{0,Lung,MB}, D_{\infty,Lung,MB}, \nu_{Lung,MB})$. The target variables could not serve as elicitation variables, as they are model dependent and some are not physically observable. Instead, external gamma ray doses $ED_{i,j}$ were elicited which would lead to i -% of the exposed population to suffer respiratory-functional morbidity after receiving it with dose rate j . The percentages of exposed population considered are $i=\{10, 50, 90\}$, the dose rates [Gy/hr] considered are $j=\{0.2, 1, 10, 100\}$ and assumed to be constant during the exposure period. The exposed population to suffer from lung morbidity is assumed to consists of persons over 40 years old and receiving supportive medical treatment.

B.3.3 Results: Step 1

Based on the quantile information of the DM and physical considerations the set of potentially observable scenarios N was determined. The set N for this problem consisted of many scenarios, a random selection of 2000 scenarios from N was taken. These 2000 scenarios were used to determine model inversions for the target variables using `lsqcurvefit` and `Minos5`; the starting values were the same for all scenarios for both optimization routines. Figure B.5 compares the model inversions for pairs of target variables graphically, `lsqcurvefit` (+) and `Minos5` (o)⁴. Like the dispersion coefficient example, the majority of the model inversions are the same, although for certain scenarios `lsqcurvefit` tends to assign small values to $D_{0,Lung,MB}$

³COSYMA is the acronym of the accident consequence code of the CEC.

⁴The symbol \oplus corresponds to same model inversions for `lsqcurvefit` and `Minos5`.

unlike Minos5. Figure B.9 shows graphically the Sum of Squares distances for the 2000 scenarios resulting from `lsqcurvefit` and Minos5; every dot can be related to a scenario for which the Sum of Squares distance has been calculated based on model inversions determined using `lsqcurvefit` and Minos5. The straight line in Figure B.6 shows where the Sum of Squares distances of the optimization routines would be equal; for 74.1% of the scenarios the Sum of Squares distance using Minos5 was less than the Sum of Squares distance using `lsqcurvefit`.

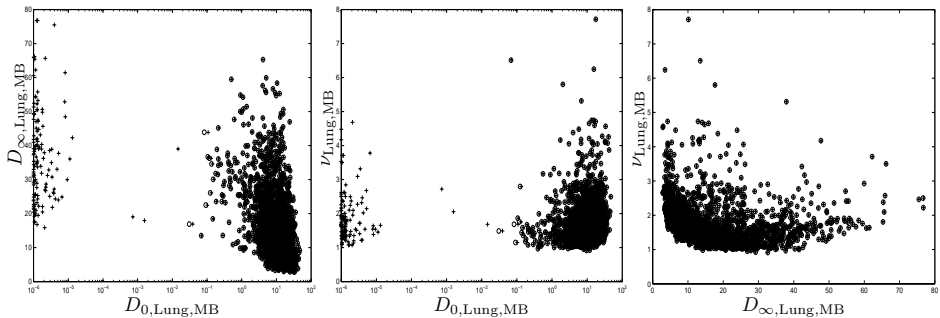


Figure B.5: Lung morbidity (supportive medical treatment, persons over 40 years old): model inversions for pairs of target variables for 2000 scenarios, `lsqcurvefit` (+) and Minos5 (o).

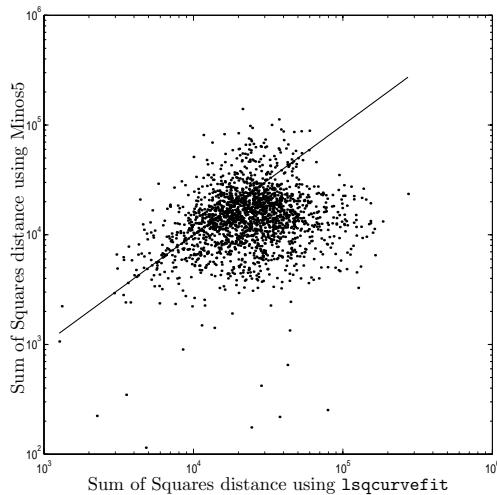


Figure B.6: Lung morbidity (supportive medical treatment, persons over 40 years old): Sum of Squares distance comparison between `lsqcurvefit` and Minos5 for 2000 scenarios.

B.3.4 Results: Step 2

The domains M , as determined using `lsqcurvefit` and `Minos5`, turned out to be such that a distribution could be determined, based on all information available on the elicitation variables. No distribution could be specified on the domain, as determined using `minimax`, using all elicitation variables, see [2]. Hence, a solution scheme based on reducing the dimension of the observable space was constructed, see Appendix C. The solution scheme is based on decomposing the probabilistic inversion problem using all elicitation variables into probabilistic inversion problems of lower complexity, i.e. the observable space associated with the probabilistic inversion problems of lower complexity consists of a selection of elicitation variables. For each problem of lower complexity a distribution on the target variables was determined, these distributions were combined to get an overall distribution on the target variables. Table B.3 gives quantile information of the marginal distributions on the target variables based on `lsqcurvefit`, `Minos5` and `minimax`.

	Quantile	<code>lsqcurvefit</code>	<code>Minos5</code>	<code>minimax</code>
$D_{0,\text{Lung,MB}}$	5%	9.93e-7	3.45e-4	5.51e-1
	50%	7.54	7.52	7
	95%	2.50e+1	2.45e+1	2.70e+1
$D_{\infty,\text{Lung,MB}}$	5%	2.74	2.91	2.96
	50%	5.40	5.43	6.40
	95%	7.53e+1	7.52e+1	5.82e+1
$\nu_{\text{Lung,MB}}$	5%	1.85	1.88	3.04
	50%	4.08	4.18	6.40
	95%	8.07	7.05	9.57

Table B.3: *Lung morbidity (supportive medical treatment, persons over 40 years old): quantile information of marginal distributions for three optimization routines.*

$$\begin{array}{l}
 D_{0,\text{Lung,MB}} \\
 D_{\infty,\text{Lung,MB}} \\
 \nu_{\text{Lung,MB}}
 \end{array}
 \begin{pmatrix}
 1 & -5.98e-1 & 1.86e-1 \\
 -5.98e-1 & 1 & -5.30e-1 \\
 1.86e-1 & -5.30e-1 & 1
 \end{pmatrix}$$

Table B.4: *Lung morbidity (supportive medical treatment, persons over 40 years old): Spearman's rank correlation matrix among target variables using `lsqcurvefit`.*

Since the majority of the model inversions for `lsqcurvefit` and `Minos5` were the same, the distribution on the domain should be similar too.

$$\begin{array}{l} D_{0,\text{Lung,MB}} \\ D_{\infty,\text{Lung,MB}} \\ \nu_{\text{Lung,MB}} \end{array} \begin{pmatrix} 1 & -5.55\text{e-1} & 1.53\text{e-1} \\ -5.55\text{e-1} & 1 & -5.73\text{e-1} \\ 1.53\text{e-1} & -5.73\text{e-1} & 1 \end{pmatrix}$$

Table B.5: *Lung morbidity (supportive medical treatment, persons over 40 years old): Spearman's rank correlation matrix among target variables using Minos5.*

$$\begin{array}{l} D_{0,\text{Lung,MB}} \\ D_{\infty,\text{Lung,MB}} \\ \nu_{\text{Lung,MB}} \end{array} \begin{pmatrix} 1 & 2.83\text{e-1} & -3.73\text{e-1} \\ 2.83\text{e-1} & 1 & -1.95\text{e-1} \\ -3.73\text{e-1} & 1.95\text{e-1} & 1 \end{pmatrix}$$

Table B.6: *Lung morbidity (supportive medical treatment, persons over 40 years old): Spearman's rank correlation matrix among target variables using minimax.*

This observation is supported by the results as presented in Tables B.3, B.4 and B.5; the marginal distributions and rank correlation matrices based on `lsqcurvefit` and `Minos5` are quite similar. The effect of random sampling is less than in the dispersion coefficient example as the number of samples equaled $5.00\text{e}+5$. The results of `minimax` in Tables B.3 and B.6 are based on $5.1\text{e}+4$ samples. Compared to the `lsqcurvefit` and `Minos5` results on the marginal distributions, the results of `minimax` are somewhat different. A large difference is identified in comparing the Spearman's rank correlation matrices. The reason could be due to the solution scheme which had to be constructed for `minimax`. Figure B.7 compares the marginal distributions on the target variables as determined by the three optimization routines graphically: `lsqcurvefit` (---), `Minos5` (—) and `minimax` (—). Loosely speaking, for $D_{\infty,\text{Lung,MB}}$ the marginal distributions are not too different, for $D_{0,\text{Lung,MB}}$ the marginal distribution based on `minimax` is considered to be different below the 25%-level and above the 95%-level compared to the marginal distributions based on `lsqcurvefit` and `Minos5` and the marginal distribution $\nu_{\text{Lung,MB}}$ based on `minimax` is considered to be different from the marginal distributions based on `lsqcurvefit` and `Minos5`.

The push-forward distributions of the three optimization routines are compared to the DM distributions in Table B.7. It is concluded that the push-forward distributions of `lsqcurvefit` and `Minos5` resemble the DM distributions very well. The push-forward distributions of `minimax` are con-

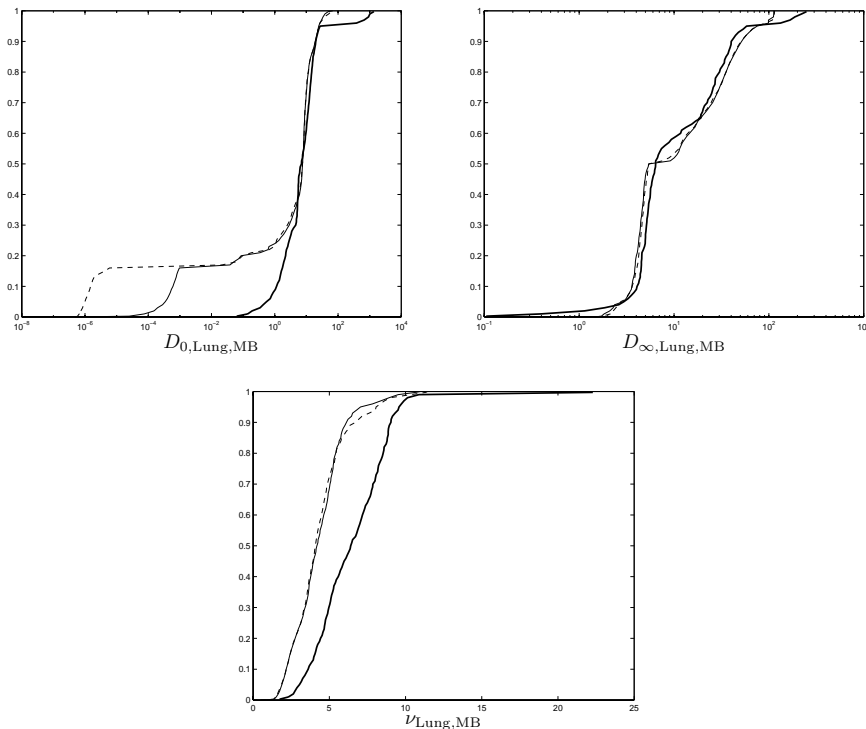


Figure B.7: Lung morbidity (supportive medical treatment, persons over 40 years old): graphical comparison of distribution of $D_{0,Lung,MB}$ (left), $D_{\infty,Lung,MB}$ (middle), $\nu_{Lung,MB}$ (right); `lsqcurvefit` (---), `Minos5` (—) and `minimax` (—).

sidered to resemble the DM distributions reasonably well; the differences occur at the 5%-iles for low dose rates (0.2 and 1 Gy/hr) and the 95%-iles for all dose-rates. Again, it is likely that these differences are due to the solution scheme which had to be constructed for `minimax`.

B.3.5 Discussion of results

A comparison between `lsqcurvefit` and `Minos5` was conducted for the probabilistic inversion problem for lung morbidity. The lung morbidity problem is regarded more ‘complicated’ than the dispersion coefficient example; (i) the non-linear model could not be transformed into a linear model, however it is strictly increasing in ED , (ii) the number of target variables was considered to be small, but the amount of information available on the elicitation variables was considered to be very large. The influence of the optimization routines in determining the domain M for this problem turned out to be insignificant. The domain M for both optimization routines were

	Quantile	DM.	lsqcurvefit	Minos5	minimax
$ED_{10,0.2}$	5%	1.12e+1	1.05e+1	1.02e+1	8.75
	50%	3.10e+1	3.10e+1	3.10e+1	3.00e+1
	95%	9.33e+1	9.33e+1	9.33e+1	1.22e+2
$ED_{50,0.2}$	5%	1.34e+1	1.30e+1	1.43e+1	1.14e+1
	50%	4.93e+1	4.93e+1	4.93e+1	4.50e+1
	95%	1.49e+2	1.35e+2	1.35e+2	1.49e+2
$ED_{90,0.2}$	5%	1.54e+1	1.65e+1	1.65e+1	1.36e+1
	50%	6.64e+1	6.64e+1	6.64e+1	5.79e+1
	95%	1.74e+2	1.74e+2	1.74e+2	2.40e+2
$ED_{10,1}$	5%	6.85	6.71	5.69	4.43
	50%	1.39e+1	1.39e+1	1.39e+1	1.21e+1
	95%	6.05e+1	5.39e+1	5.04e+1	6.18e+1
$ED_{50,1}$	5%	8.24	8.29	7.71	5.94
	50%	1.79e+1	1.79e+1	1.79e+1	1.80e+1
	95%	7.55e+1	7.55e+1	7.55e+1	7.58e+1
$ED_{90,1}$	5%	9.56	1.01e+1	9.56	6.98
	50%	2.13e+1	2.13e+1	2.13e+1	2.28e+1
	95%	1.29e+2	1.29e+2	1.29e+2	9.04e+1
$ED_{10,10}$	5%	3.00	2.92	3.00	2.43
	50%	7.05	7.05	6.75	5.25
	95%	6.04e+1	5.14e+1	5.33e+1	4.97e+1
$ED_{50,10}$	5%	3.80	3.67	3.77	3.32
	50%	8.80	8.80	8.80	7.42
	95%	7.53e+1	7.53e+1	7.51e+1	6.03e+1
$ED_{90,10}$	5%	4.36	4.35	4.36	4.02
	50%	1.03e+1	1.03e+1	1.03e+1	9.50
	95%	1.29e+2	1.29e+2	1.29e+2	7.30e+1
$ED_{10,100}$	5%	2.13	2.13	2.13	2.15
	50%	4.14	4.14	4.10	4.66
	95%	6.04e+1	5.14e+1	5.33e+1	4.85e+1
$ED_{50,100}$	5%	3.11	2.79	3.11	3.00
	50%	5.49	5.49	5.49	6.48
	95%	7.53e+1	7.53e+1	7.53e+1	5.94e+1
$ED_{90,100}$	5%	3.73	3.70	3.73	3.73
	50%	7.02	7.00	7.00	8.36
	95%	1.29e+2	1.29e+2	1.29e+2	7.19e+1

Table B.7: Lung morbidity (supportive medical treatment, persons over 40 years old): quantile information of marginal distributions of elicitation variables [Gy].

almost identical, which resulted in similar distributions on the target variables with relative information values of 3.97 and 3.98 for `lsqcurvefit` and `Minos5` respectively. The information used in the uncertainty analysis of COSYMA was determined using `minimax`. The probabilistic inversion problem based on `minimax` turned out to be such that no distribution could be determined using all information available on the elicitation variables. A solution scheme was constructed which decomposed the problem into problems of lower complexity. A consequence of applying such a solution scheme is that the push-forward distributions of the ‘overall’ distribution on the tar-

get variables will resemble the DM distributions not to well. However, the differences between the results for the marginal distributions on the target variables as determined using `minimax` and the marginal distributions on the target variables as determined using `lsqcurvefit` and `Minos5` were regarded to be insignificant, furthermore the push-forward distributions of `minimax` did resemble the DM distributions reasonably well. However, the difference between the marginal distributions for $\nu_{\text{Lung,MB}}$ and the lower and upper percentiles of the marginal distributions for $D_{0,\text{Lung,MB}}$ was acknowledged.

B.4 Systemic retention of Sr in the Human body

For a more detailed description on the probabilistic inversion problem for the systemic retention of Sr in the human body, the reader is referred to [2].

B.4.1 Compartmental model

Roughly, systemic retention of Sr in the human body in COSYMA is described by the compartmental model shown in Figure B.8. The parameters k_{ij} [1/d] are termed transfer coefficients and express the rate at which Sr is transferred from compartment i to compartment j . This rate is assumed to be constant within a short time interval. Based on this assumption and Figure B.8, a set of first order linear differential equations, which, together with appropriate initial conditions⁵, fully specifies the systemic retention of Sr in the human body.

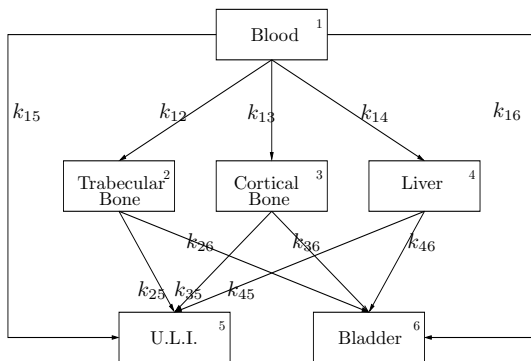


Figure B.8: *Systemic retention of Sr: Compartmental model.*

⁵ $m_1(t=0) = 1$ and $m_j(t=0) = 0$ for $j = 2, \dots, 6$, where $m_i(t)$ represents the retention of Sr in compartment i at time t .

Although the transfer coefficients cannot be measured directly, the following relationships are assumed to hold on the basis of physical considerations:

1. The transfer coefficients from compartment i to Bladder are modeled as

$$k_{i6} = U:F * k_{i5} \quad (\text{B.2})$$

with $i \in \{1, \dots, 4\}$ and $U:F$ represents the Urine-to-Faeces ratio for Sr. The project staff set the $U:F$ -ratio for Sr to 3.3.

2. It was assumed that the transfer coefficient from Blood to Trabecular Bone, k_{12} is correlated to the transfer coefficient from Blood to Cortical Bone, k_{13} in the following manner:

$$k_{12} = tc * k_{13} \quad (\text{B.3})$$

where tc is the Trabecular-to-Cortical factor. The project staff assigned a lognormal distribution to the tc -factor, with 5% and 95% quantile of 0.6 and 2.

B.4.2 Target variables & Elicitation variables

The transfer coefficients ($k_{13}, k_{14}, k_{15}, k_{25}, k_{35}, k_{45}$) are the model input parameters which are regarded uncertain, hence they are the target variables. These transfer coefficients cannot be measured directly and are regarded as model dependent, hence the target variables could not serve as elicitation variables. Elicitation variables were formulated on the amount of Sr retained in certain regions of the human body at certain times⁶ after being administered intravenously as a single injection. The regions of the human body for which the experts were queried were Skeleton+Liver (Skel+Liver) and Skeleton as a percentage of Skeleton+Liver ($\frac{\text{Skel}}{\text{Skel+Liver}}$), where Skeleton is the sum of Trabecular Bone and Cortical Bone. The functions describing the elicitation variables turned out to be non-linear and non-monotone.

B.4.3 Results: Step 1

The quantile information of the DM together with the information on tc and physical considerations were used to determine the set of potentially observable scenarios N . Like the lung morbidity example, the set N for this problem consisted of many scenarios. A random selection of 2000 scenarios from N was taken. These 2000 scenarios were used to determine model inversions for the target variables ($k_{13}, k_{14}, k_{15}, k_{25}, k_{35}, k_{45}$) using `lsqcurvefit` and `Minos5`. Note, each scenario in N contains a realization

⁶The timepoints t are $t \in \{1 \text{ day}, 1 \text{ week}, 1 \text{ month}, 1 \text{ year}, 10 \text{ years}, 50 \text{ years}\}$.

of tc as well, hence the model inversions for the target variables are determined conditional on this realization of tc . The starting values for the target variables are the same for `lsqcurvefit` and `Minos5` for all 2000 scenarios.

Figure B.9 shows graphically the Sum of Squares distance for the 2000 scenarios resulting from `lsqcurvefit` and `Minos5`; every dot can be related to a scenario for which the Sum of Squares distance has been calculated based on model inversions determined using `lsqcurvefit` and `Minos5`. The

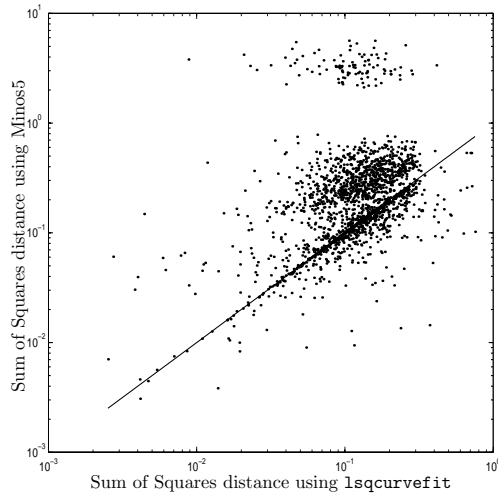


Figure B.9: *Systemic retention of S_r : Sum of Squares distance comparison between `lsqcurvefit` and `Minos5` for 2000 scenarios.*

straight line in Figure B.9 shows where the Sum of Squares distances of the optimization routines would be equal. Based on the 2000 scenarios selected, the Sum of Squares distance resulting from `Minos5` was less than the Sum of Squares distance resulting from `lsqcurvefit` in 62.1% of the scenarios. Looking at Figure B.9 somewhat closer, a cluster of points for which the Sum of Squares distance resulting from `Minos5` is roughly more than one order of magnitude larger than the Sum of Squares distance resulting from `lsqcurvefit`. It is difficult to give an explanation for this, but it may be due to the functions describing the elicitation variables being non-linear and non-monotone, which may cause `Minos5` to end its search in a point which is a local minimum. Dealing with functions which is non-linear and non-monotone, it is recommended to perform the minimization of the Sum of Squares distance using different starting values. For each target variable 10 different starting values were selected randomly from the interval $[\frac{1}{10} \cdot k_{ij}^*, 10 \cdot k_{ij}^*]$, where k_{ij}^* represents the initial starting value.

Here, the selection of the 10 different starting values is ad-hoc, however the choice of the starting values may be less ad-hoc if they could be based on the certain characteristics of the model, expert assessments and/or physical

	k_{13}	k_{14}	k_{15}	k_{25}	k_{35}	k_{45}
sv.1	4.42e-1	1.94e-3	5.43e-1	5.56e-2	1.11e-4	8.29e-5
sv.2	4.33e-1	1.90e-3	5.33e-1	5.45e-2	1.09e-4	8.13e-5
sv.3	1.32	5.78e-3	1.62	1.65e-1	3.31e-4	2.47e-4
sv.4	5.93e-1	2.60e-3	7.29e-1	7.46e-2	1.49e-4	1.11e-4
sv.5	3.92e-1	1.30e-3	6.23e-1	3.45e-2	1.69e-4	2.13e-5
sv.6	3.33e-2	1.46e-4	4.09e-2	4.19e-3	8.37e-6	6.25e-6
sv.7	1.63	7.15e-3	2.00	2.05e-1	4.09e-4	3.05e-4
sv.8	9.70e-1	4.26e-3	1.19	1.22e-1	2.44e-4	1.82e-4
sv.9	2.03	8.92e-3	2.50	2.55e-1	5.11e-4	3.81e-4
sv.10	1.02	4.46e-3	1.25	1.28e-1	2.55e-4	1.91e-4

Table B.8: *Systemic retention of Sr: 10 different starting values (sv.) for transfer coefficients k_{ij} (1/d).*

insights. Nonetheless, the cluster of points identified in Figure B.9 has disappeared in Figure B.10. Furthermore the Sum of Squares distances in Figure B.10 are concentrated more around the straight line. This shows the value of performing the minimization of the Sum of Squares distance using multiple starting values, in case of a non-linear and non-monotone model predictor.

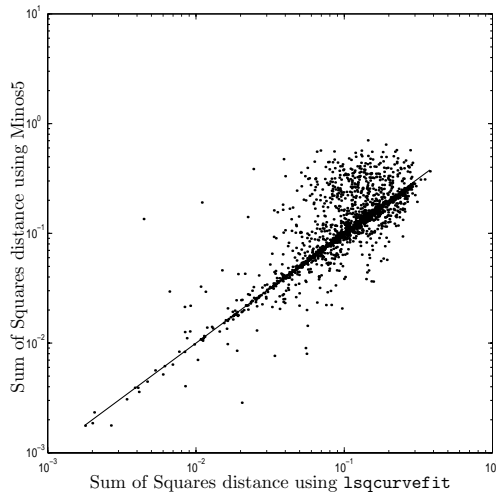


Figure B.10: *Systemic retention of Sr: Sum of Squares distance comparison between lsqcurvefit and Minos5 for Sr-problem using 10 different starting values for 2000 scenarios.*

Figures B.11 and B.12 compare graphically the model inversions for the 15 pairs of target variables. Unlike the dispersion coefficient and Lung morbidity example, the majority of the model inversions are not the same;

it is observed that for certain combinations of target variables the model inversions from Minos5 (\circ) are very different from the model inversions from lsqcurvefit (+).

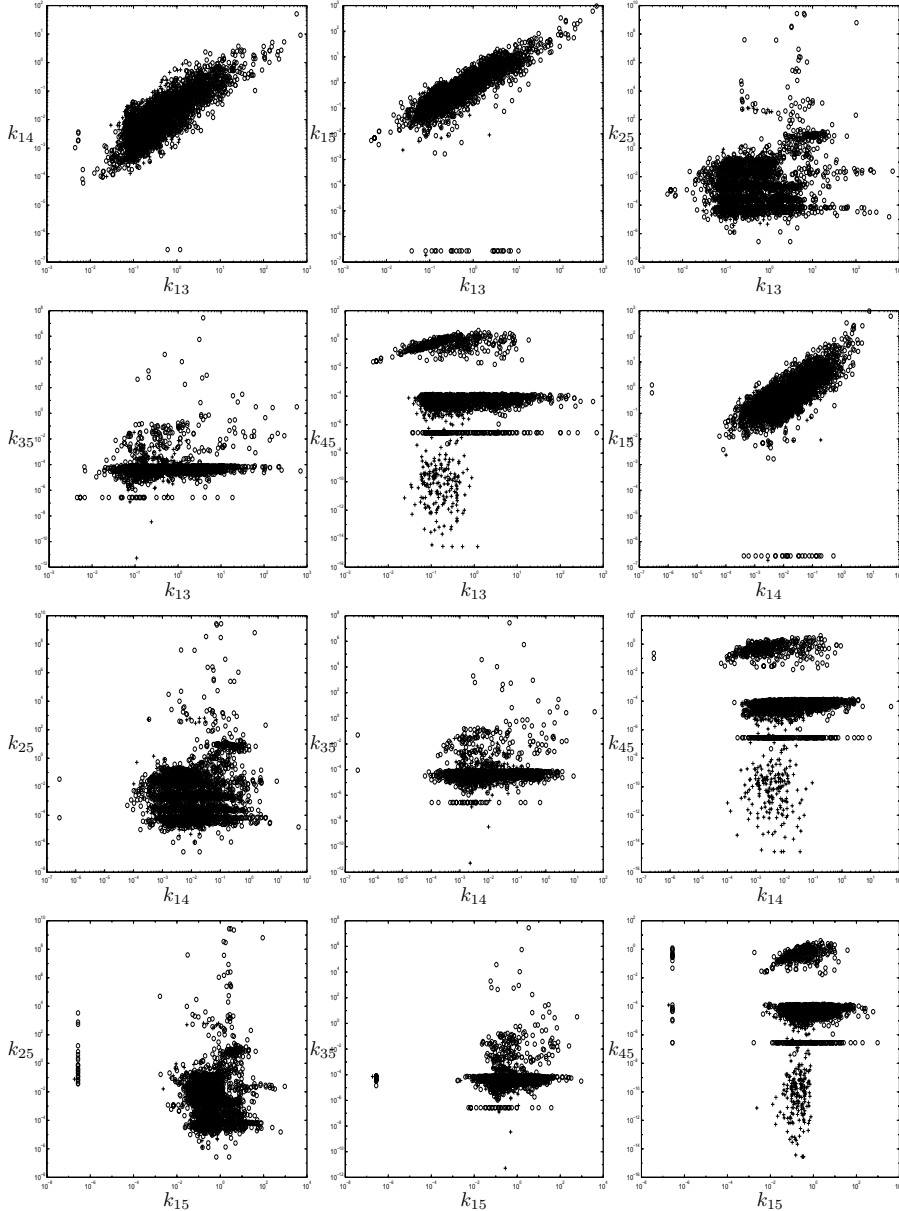


Figure B.11: Systemic retention of Sr: model inversions for 12 pairs of target variables for 2000 scenarios, lsqcurvefit (+) and Minos5 (\circ).

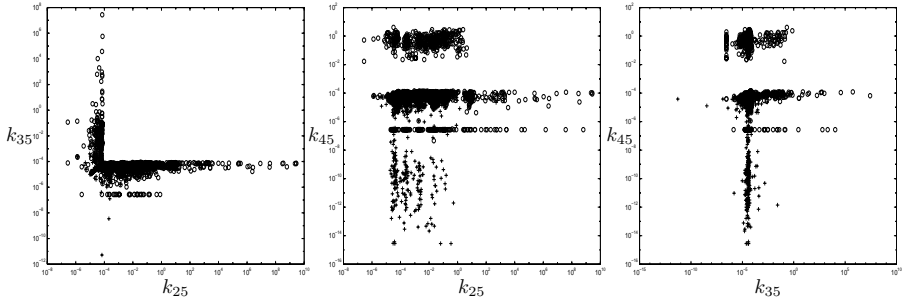


Figure B.12: *Systemic retention of Sr: model inversions for 3 pairs of target variables for 2000 scenarios, lsqcurvefit (+) and Minos5 (o).*

B.4.4 Results: Step 2

For each domain M , samples were propagated through the functions describing the elicitation variables to obtain a set of points in the observable space. The observable space consisted of 13 dimensions: the elicitation variables and the Trabecular-to-Cortical factor tc . Based on quantile information presented in Table B.9, it is concluded that the marginal distributions as determined using `lsqcurvefit`, `Minos5` and `minimax` are quite different. Figure B.13 presents the similarities/differences of the various marginal distributions graphically; `lsqcurvefit` (---), `Minos5` (—) and `minimax` (—).

The Spearman’s rank correlation matrices determined from the distributions using `lsqcurvefit`, `Minos5` and `minimax` are given in Tables B.10, B.11 and B.12, respectively. Looking at these correlation matrices, it is observed that the correlation matrices are very different.

The push-forward distributions of the distributions on the target variables are compared to the DM distributions in Table B.4.4. The push-forward distributions based on `lsqcurvefit` and `Minos5` are based on $2.00e+5$ samples, whereas the push-forward distributions based on `minimax` is based on $1.27e+6$ samples; it is concluded that the push-forward distributions of `lsqcurvefit`, `Minos5` and `minimax` resemble the DM distributions very well.

B.4.5 Discussion of results

A comparison between `lsqcurvefit` and `Minos5` was conducted for the probabilistic inversion problem for systemic retention of Sr in the human body. The systemic retention of Sr is regarded as a ‘complex’ probabilistic inversion problem; (i) non-linear, non-monotone functions describing the elicitation variables, (ii) the number of target variables is large, and the amount of information available on the elicitation variables is very large also. The influence of the optimization routines in determining the domain M for

	Quantile	lsqcurvefit	Minos5	minimax
k_{13}	5%	1.57e-1	1.02e-1	6.89e-2
	50%	4.23e-1	1.02	2.18e-1
	95%	1.21	3.24e+1	6.97e-1
k_{14}	5%	1.21e-3	2.44e-3	7.70e-4
	50%	1.19e-2	5.17e-2	9.57e-3
	95%	3.32e-1	6.55e-1	7.25e-2
k_{15}	5%	2.44e-1	1.39e-1	5.54e-2
	50%	5.59e-1	1.14	2.68e-1
	95%	9.07e-1	3.84e+1	3.53e-1
k_{25}	5%	5.83e-5	3.19e-5	1.45e-5
	50%	2.49e-3	2.09e-3	2.74e-2
	95%	1.17e-1	1.15e-1	1.22e-1
k_{35}	5%	4.28e-6	1.24e-5	2.63e-5
	50%	3.08e-5	4.61e-5	5.48e-5
	95%	9.13e-5	5.66e-3	7.41e-2
k_{45}	5%	1.02e-5	1.07e-5	7.33e-6
	50%	4.47e-5	5.33e-5	4.09e-5
	95%	9.68e-5	1.45e-4	9.58e-5

Table B.9: *Systemic retention of Sr: quantile information on target variables based on lsqcurvefit, Minos5 and minimax (1/d).*

k_{13}	1	4.47e-1	4.31e-1	3.01e-1	-5.03e-2	8.02e-2	-5.53e-2
k_{14}	4.47e-1	1	2.58e-1	-2.13e-1	-8.80e-3	4.49e-1	3.14e-1
k_{15}	4.31e-1	2.58e-1	1	-2.60e-3	-3.74e-1	-1.45e-2	-1.13e-1
k_{25}	3.01e-1	-2.13e-1	-2.60e-3	1	-1.96e-1	-2.91e-1	1.78e-2
k_{35}	-5.03e-2	-8.80e-3	-3.74e-1	-1.96e-1	1	4.02e-1	-2.04e-1
k_{45}	8.02e-2	4.49e-1	-1.45e-1	-2.91e-1	4.02e-1	1	2.80e-3
t_c	-5.53e-2	3.14e-1	-1.13e-1	1.78e-2	-2.04e-1	2.80e-3	1

Table B.10: *Systemic retention of Sr: Spearman's rank correlation matrix resulting from lsqcurvefit.*

this problem turned out to be significant. Some observations; firstly, dealing with functions which are non-linear and non-monotone it is recommended to perform the minimization of the Sum of Squares distance using a multiple starting values. Secondly, unlike for the dispersion coefficient example and lung morbidity, the majority of the model inversions obtained from the optimization routines were not the same. This resulted in distributions on

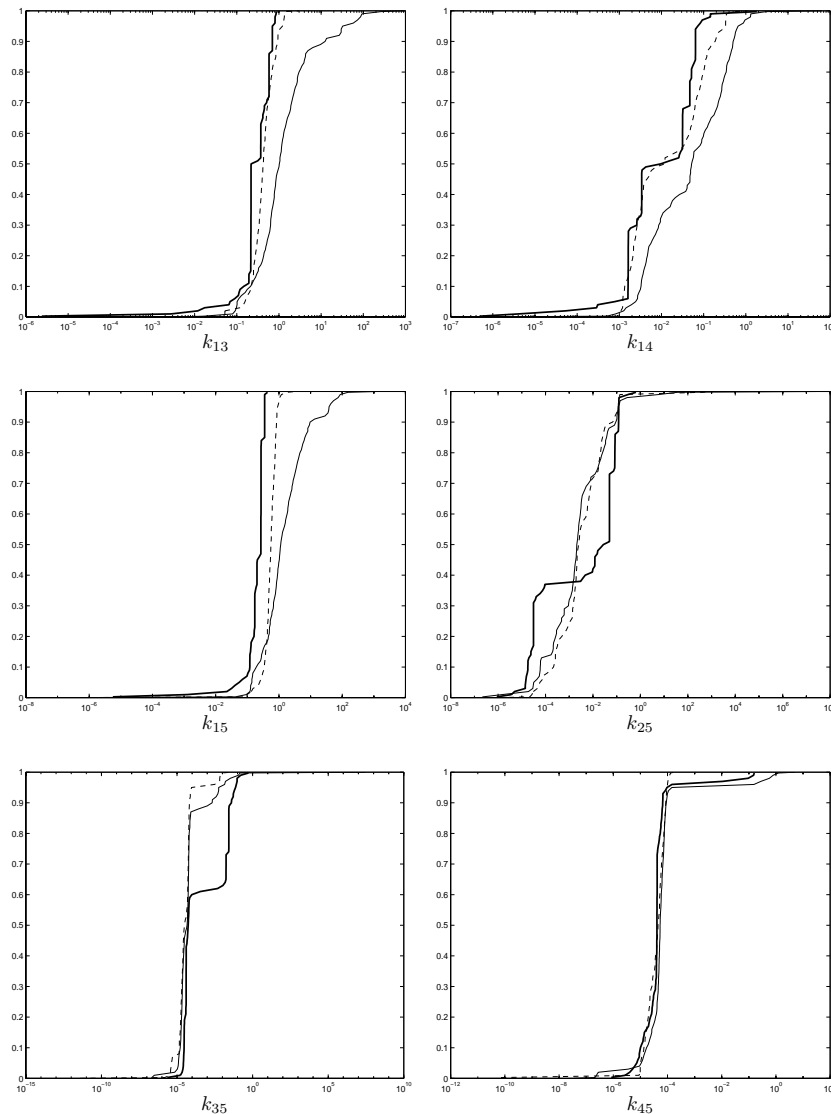


Figure B.13: *Systemic retention of Sr: graphical comparison of distribution of k_{13} (top-left), k_{14} (top-middle), k_{15} (top-right), k_{25} (below-left), k_{35} (below-middle) and k_{45} (below-right). `lsqcurvefit` (- -), `Minos5` (-) and `minimax` (-).*

the target variables which were different, which is confirmed by the relative information values: 5.66 and 5.06 for `lsqcurvefit` and `Minos5`, respectively. Furthermore the distribution based on `minimax` was regarded to be different

$$\begin{array}{l}
k_{13} \\
k_{14} \\
k_{15} \\
k_{25} \\
k_{35} \\
k_{45} \\
tc
\end{array}
\left(\begin{array}{cccccccc}
1 & 6.75e-1 & 8.70e-1 & -1.42e-1 & -4.33e-2 & -2.96e-2 & -1.17e-1 \\
6.75e-1 & 1 & 6.52e-1 & -2.19e-1 & -1.38e-1 & 2.44e-1 & 1.94e-2 \\
8.70e-1 & 6.52e-1 & 1 & -1.82e-1 & -1.68e-1 & 1.32e-2 & -2.63e-2 \\
-1.42e-1 & -2.19e-1 & -1.82e-1 & 1 & -2.15e-1 & -7.05e-2 & 2.40e-1 \\
-4.33e-2 & -1.38e-1 & -1.68e-1 & -2.15e-1 & 1 & 1.94e-1 & -3.29e-1 \\
-2.96e-2 & 2.44e-1 & 1.32e-2 & -7.05e-2 & 1.94e-1 & 1 & 9.19e-2 \\
-1.17e-1 & 1.94e-2 & -2.63e-2 & 2.40e-1 & -3.29e-1 & 9.19e-2 & 1
\end{array} \right)$$

Table B.11: *Systemic retention of Sr: Spearman's rank correlation matrix resulting from Minos5.*

$$\begin{array}{l}
k_{13} \\
k_{14} \\
k_{15} \\
k_{25} \\
k_{35} \\
k_{45} \\
tc
\end{array}
\left(\begin{array}{cccccccc}
1 & -3.99e-2 & 4.15e-1 & -6.15e-1 & 4.60e-1 & -4.46e-1 & -6.19e-1 \\
-3.99e-2 & 1 & -4.51e-1 & 1.19e-1 & -2.73e-1 & -2.61e-2 & 4.28e-2 \\
4.15e-1 & -4.51e-1 & 1 & -2.02e-1 & 2.70e-1 & 1.67e-1 & -2.51e-1 \\
-6.15e-1 & 1.19e-1 & -2.02e-1 & 1 & -8.45e-1 & 1.15e-1 & 6.03e-1 \\
4.60e-1 & -2.73e-1 & 2.70e-1 & -8.45e-1 & 1 & 1.35e-2 & -8.75e-1 \\
-4.46e-1 & -2.61e-2 & 1.67e-1 & 1.15e-1 & 1.35e-2 & 1 & -2.03e-2 \\
-6.19e-1 & 4.28e-2 & -2.51e-1 & 6.03e-1 & -8.75e-1 & -2.03e-2 & 1
\end{array} \right)$$

Table B.12: *Systemic retention of Sr: Spearman's rank correlation matrix resulting from minimax.*

from the distributions obtained using `lsqcurvefit` and Minos5 as well. However, the push-forward distributions of the three distributions on the target variables did resemble the DM distributions very well.

B.5 Conclusion

The most important step in a probabilistic inversion solution scheme is the determination of the domain M in the target variable space. The heuristics used in determining M consist of three elements: (i) the assessments of the experts, (ii) the mapping T , and (iii) the physics underlying the problem. These three elements meet when minimizing the Sum of Squares distance between potentially observable scenarios based on expert assessments, and T , which is a function of the target variables. Minimization of the Sum

Time		Skel+Liver			
		DM.	lsqcurvefit	Minos5	minimax
1 day	5%	1.70e-1	1.30e-1	1.69e-1	1.70e-1
	50%	3.24e-1	3.24e-1	3.24e-1	3.24e-1
	95%	5.76e-1	5.76e-1	5.76e-1	5.76e-1
1 week	5%	1.17e-1	1.17e-1	1.17e-1	1.17e-1
	50%	2.29e-1	2.29e-1	2.29e-1	2.29e-1
	95%	4.76e-1	4.76e-1	4.76e-1	4.76e-1
1 month	5%	1.04e-1	1.04e-1	1.04e-1	1.04e-1
	50%	2.11e-1	2.11e-1	2.11e-1	2.11e-1
	95%	3.51e-1	3.51e-1	3.51e-1	3.51e-1
1 year	5%	6.74e-2	6.65e-2	6.74e-2	6.74e-2
	50%	1.38e-1	1.38e-1	1.38e-1	1.38e-1
	95%	2.43e-1	2.43e-1	2.43e-1	2.43e-1
10 years	5%	1.81e-2	1.80e-2	1.80e-2	1.80e-2
	50%	6.45e-2	6.45e-2	6.45e-2	6.45e-2
	95%	1.37e-1	1.37e-1	1.37e-1	1.37e-1
50 years	5%	1.11e-3	1.10e-3	1.10e-3	1.10e-3
	50%	1.85e-2	1.85e-2	1.85e-2	1.85e-2
	95%	8.88e-2	8.87e-2	8.88e-2	8.86e-2
Time		Skel Skel+Liver			
		DM.	lsqcurvefit	Minos5	minimax
1 day	5%	8.46e-1	8.45e-1	8.49e-1	8.46e-1
	50%	9.56e-1	9.56e-1	9.56e-1	9.56e-1
	95%	9.98e-1	9.98e-1	9.98e-1	9.98e-1
1 week	5%	8.22e-1	8.20e-1	8.29e-1	8.20e-1
	50%	9.57e-1	9.56e-1	9.58e-1	9.58e-1
	95%	9.98e-1	9.98e-1	9.98e-1	9.98e-1
1 month	5%	8.51e-1	8.16e-1	8.50e-1	8.50e-1
	50%	9.84e-1	9.56e-1	9.58e-1	9.80e-1
	95%	9.99e-1	9.98e-1	9.99e-1	9.99e-1
1 year	5%	7.70e-1	7.70e-1	7.70e-1	7.77e-1
	50%	9.94e-1	9.60e-1	9.60e-1	9.90e-1
	95%	9.99e-1	9.98e-1	9.99e-1	9.99e-1
10 years	5%	6.79e-1	6.77e-1	6.89e-1	6.80e-1
	50%	9.95e-1	9.87e-1	9.79e-1	9.90e-1
	95%	9.99e-1	9.99e-1	9.99e-1	9.99e-1
50 years	5%	6.39e-1	6.40e-1	6.40e-1	6.40e-1
	50%	9.96e-1	9.90e-1	9.90e-1	9.90e-1
	95%	9.99e-1	9.99e-1	9.99e-1	9.99e-1

Table B.13: Systemic retention of Sr: Quantile information comparison of distributions of DM vs. push-forward distributions based on lsqcurvefit, Minos5 and minimax.

of Squares distance is a global minimization problem. In this appendix the effect of different optimization programs capable of dealing with global minimization problems has been investigated.

The optimization routines considered were the `lsqcurvefit` function of the Matlab5.3 version 11 optimization toolbox, the Minos5 solver of GAMS. It is concluded that the influence of optimization programs is negligible in case of non-linear but monotone model predictors (dispersion coefficient example and lung morbidity). However, a difference between the optimization programs was identified in dealing with non-linear and non-monotone functions describing the elicitation variables (systemic retention of Sr in the human body). Functions which are non-monotone may result in optimization routines to end its search in the target variable space in a local minimum instead of a global minimum. The following is recommended; try to avoid non-monotone functions in doing probabilistic inversion, if this is not possible be sure to minimize the Sum of Squares distance a number of times, each time using different starting values.

The distributions used in the uncertainty analysis of COSYMA were based on the minimization of the Sum of Squares distance using the `minimax` optimization routine of the Matlab4.2 optimization toolbox. The distributions on the target variables based on `minimax` were compared to the distributions based on `lsqcurvefit` and Minos5, i.e. the marginal distributions and Spearman's rank correlation matrices were compared. For the dispersion coefficient example and lung morbidity, the comparison was considered to be reasonably good. The comparison for the systemic retention of Sr was considered to be poor, although the push-forward distributions of all three distributions available on the target variables resembled the DM distributions rather well.

Appendix C

Reduction of Dimension

The reduction of dimension technique reduces the dimension of the observable space. In this way (i) probabilistic inversion problems which were infeasible considering all dimensions may become feasible and (ii) probabilistic inversion problems with an observable space consisting of a large number of dimensions may be broken down in probabilistic inversion problems which are computationally tractable.

Consider n elicitation variables and suppose the corresponding PI problem is infeasible. The reduction of dimension strategy looks at all problems involving $n - 1$ elicitation variables, i.e. $\binom{n}{n-1}$ problems with observable space dimension $n - 1$. For each of the $\binom{n}{n-1}$ problems, the respective domains are determined in **Step 1**. Next, let $M_{\text{RD},n-1}$ be the union of the $\binom{n}{n-1}$ search grids. Samples from $M_{\text{RD},n-1}$ are propagated through mapping T . For each $\binom{n}{n-1}$ observable space, select the push-forward samples of mapping T corresponding to the elicitation variables which make up the respective observable space. In this way $\binom{n}{n-1}$ PI problems are constructed. Let n^* ($n^* \leq \binom{n}{n-1}$) denote the number of PI problems for which a distribution on $M_{\text{RD},n-1}$ in **Step 2** can be determined; assume n^* ($n^* > 0$) distributions on $M_{\text{RD},n-1}$ are obtained. The problem of finding a distribution over the target variables which ‘best fits’ these n^* distributions is described in **Step 2** of the PARFUM solution scheme (see Section 1.2.1).

If $n^* = 0$, the dimension of the observable space is reduced once more, and **Step 1** and **Step 2** of the PREJUDICE solution scheme are performed for $\binom{n}{n-2}$ problems with observable space dimension $n - 2$ in order to determine a distribution on $M_{\text{RD},n-2}$.

Note, if the dimension of the observable space is 2 and its corresponding PI-problem is infeasible, then reduction of dimension results in a solution scheme which resembles the PARFUM solution scheme closely.

Appendix D

Minimal solution

Consider the following example; 5%, 50% and 95% quantile points are available for elicitation variables Y_j ($j \in \{1, 2\}$), and domain M_{\min} is such that $T(M_{\min}) \subset \cup_{k=1}^4 I_{1,k} \times I_{2,k}$ (Figure D.1), where $I_{j,k}$ represents the k -th interquartile interval of the j -th elicitation variable.

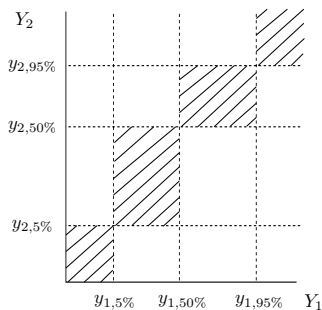


Figure D.1: *The shaded area visualizes $T(M_{\min})$.*

Clearly, the quantile constraints are satisfied in this case; 5% probability mass is assigned to samples in hypercubes (1,1) and (4,4) and 45% probability mass is assigned to the samples in hypercubes (2,2) and (3,3). In principle, this problem could be solved if M_{\min} consisted of 4 samples; for example the k -th sample maps into each $I_{1,k} \times I_{2,k}$ ($k = 1, \dots, 4$) and assigning these samples 0.05, 0.45, 0.45, 0.05 respectively. Propagation of domains resulting in a coverage of hypercubes as depicted in Figure D.2 could also be solved using 4 samples only. Generally speaking, any probabilistic inversion problem with information on the elicitation variables available as k quantile points can be solved using $k + 1$ points only. This type of solution is referred to as a minimal solution of a probabilistic inversion problem. Clearly, any convex combination of minimal solutions is a solution to the probabilistic inversion problem as well.

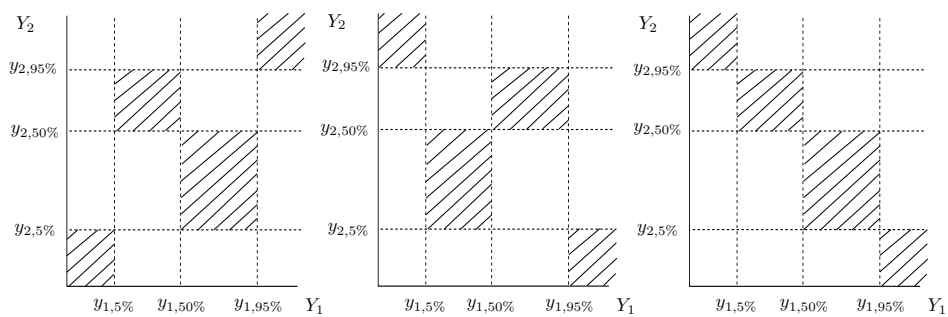


Figure D.2: Visualization of coverage of hypercubes which could be solved using a minimal number of samples.

Appendix E

Dependencies: assessments and results.

In this appendix, the $\rho\pi$ -table is presented graphically together with the expert assessments and results of Example 1 and 2 of Chapter 2.

E.1 The $\rho\pi$ -table

Figures E.1 through E.6 illustrate graphically the $\rho\pi$ -tables for values $r_2 = 0.05$ through $r_2 = 0.95$, respectively, where

$$\pi_{r_1, r_2}(Y_1, Y_2) := P(F_{Y_1}(Y_1) > r_1 \mid F_{Y_2}(Y_2) > r_2). \quad (\text{E.1})$$

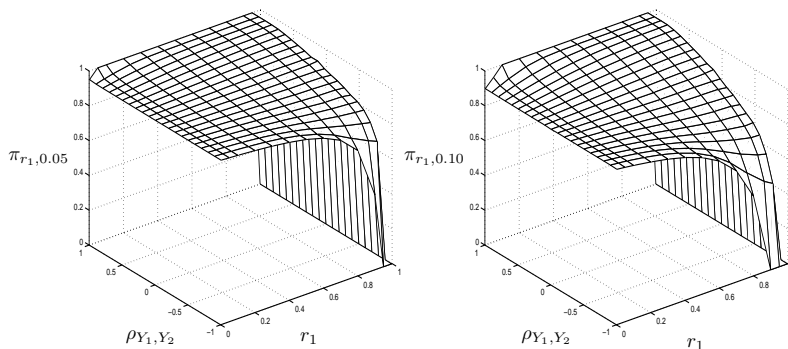


Figure E.1: Graphical illustration of $\rho\pi$ -table for $r_2 = 0.05$ (left) and $r_2 = 0.10$ (right).

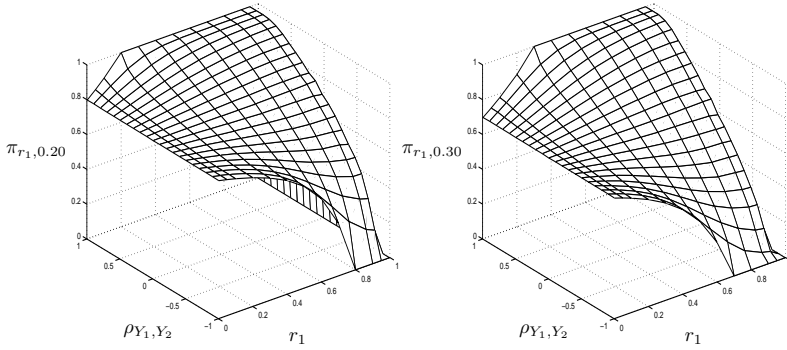


Figure E.2: Graphical illustration of $\rho\pi$ -table for $r_2 = 0.20$ (left) and $r_2 = 0.30$ (right).

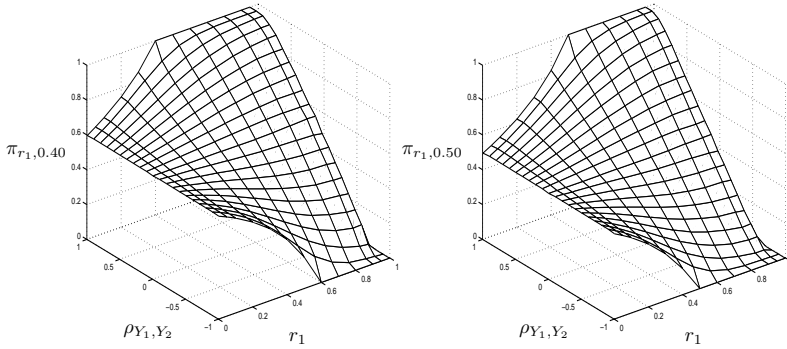


Figure E.3: Graphical illustration of $\rho\pi$ -table for $r_2 = 0.40$ (left) and $r_2 = 0.50$ (right).

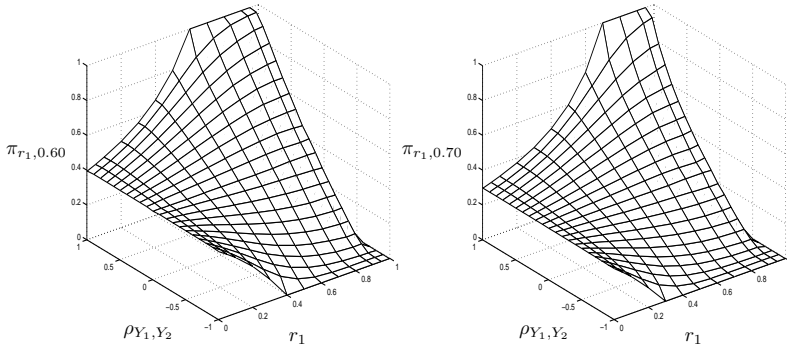


Figure E.4: Graphical illustration of $\rho\pi$ -table for $r_2 = 0.60$ (left) and $r_2 = 0.70$ (right).

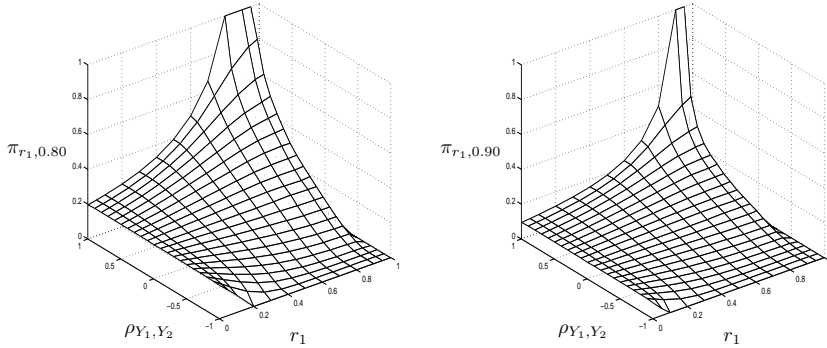


Figure E.5: Graphical illustration of $\rho\pi$ -table for $r_2 = 0.80$ (left) and $r_2 = 0.90$ (right).

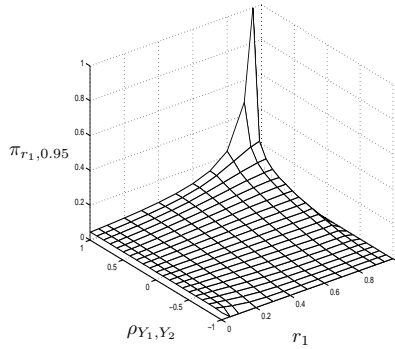


Figure E.6: Graphical illustration of $\rho\pi$ -table for $r_2 = 0.95$.

E.2 Example 1

Table E.1 presents the conditional probabilities among elicitation variables obtained from the experts. The abbreviation N/A refers to no assessments provided by the expert for the pair of elicitation variables. Table E.2 presents the Spearman's rank correlations based on the conditional probabilities using $\rho\pi$ -table. From Tables E.1 and E.2 it is concluded that the expert assessments are not the same. Obviously experts did not agree on the sign and strength of the dependence.

Tables E.3 and E.4 present the Spearman's rank correlation resulting from Strategy 1 and Strategy 2, respectively.

Elic. questions		$\pi_{\frac{1}{2}, \frac{1}{2}}(Y_1, Y_2)$							
Y_1	Y_2	Exp.1	Exp.2	Exp.3	Exp.4	Exp.5	Exp.6	Exp.7	Exp.8
DCO	DPO	0.85	0.25	1	0.5	0.52	0.7	N/A	N/A
DPO	DSO	N/A	N/A	1	0.1	N/A	0.7	N/A	N/A
DSI	DCI	0.85	0.7	1	0.1	0.4	0.7	N/A	N/A
DCO	DCI	0.90	0.5	1	0.6	0.6	0.54	N/A	N/A
BPO	BSO	0.5	N/A	1	0.5	0.4	0.7	N/A	N/A
BPO	BCO	0.5	0.84	1	0.1	0.52	0.7	N/A	N/A
BSI	BCI	0.5	0.84	1	0.1	0.4	0.7	N/A	N/A
BCI	BCO	0.99	0.84	1	0.6	0.6	0.54	N/A	N/A
BPO	DPO	0.95	0.75	0.75	0.6	0.5	0.90	0.90	N/A

Table E.1: *Example 1: conditional probability assessments $\pi_{\frac{1}{2}, \frac{1}{2}}(Y_1, Y_2)$ given by 8 experts from the Animal expert panel.*

Elic. questions		Spearman's ρ_{Y_1, Y_2}							
Y_1	Y_2	Exp.1	Exp.2	Exp.3	Exp.4	Exp.5	Exp.6	Exp.7	Exp.8
DCO	DPO	0.79	-0.62	1	0	0.06	0.45	N/A	N/A
DPO	DSO	N/A	N/A	1	-0.91	N/A	0.45	N/A	N/A
DSI	DCI	0.79	0.45	1	-0.91	-0.27	0.45	N/A	N/A
DCO	DCI	0.88	0	1	0.24	0.24	0.12	N/A	N/A
BPO	BSO	0	N/A	1	0	-0.27	0.45	N/A	N/A
BPO	BCO	0	0.78	1	-0.91	0.06	0.45	N/A	N/A
BSI	BCI	0	0.78	1	-0.91	-0.27	0.45	N/A	N/A
BCI	BCO	0.99	0.78	1	0.24	0.24	0.12	N/A	N/A
BPO	DPO	0.94	0.61	0.61	0.24	0	0.88	0.88	N/A

Table E.2: *Example 1: Spearman's rank correlations $\rho(Y_1, Y_2)$ for 8 experts from the Animal expert panel.*

	DPO	DSO	DCO	DPI	DSI	DCI	BPO	BSO	BCO	BPI	BSI	BCI
1	1.1e-1	6.0e-2	-1.0e-2	1.0e-2	0	4.3e-1	1.3e-1	3.0e-2	0	0	1.0e-2	0
1.1e-1	1	3.9e-1	-1.0e-2	0	4.0e-2	5.0e-2	3.0e-2	1.0e-2	-1.0e-2	0	0	0
6.0e-2	3.9e-1	1	-1.0e-2	2.0e-2	1.3e-1	4.0e-2	3.0e-2	2.0e-2	0	1.0e-2	2.0e-2	0
-1.0e-2	-1.0e-2	-1.0e-2	1	0	-1.0e-2	0	0	-2.0e-2	0	0	-1.0e-2	0
1.0e-2	0	2.0e-2	0	1	1.7e-1	0	1.0e-2	0	1.0e-2	0	0	0
0	4.0e-2	1.3e-1	-1.0e-2	1.7e-1	1	0	0	1.0e-2	-2.0e-2	1.0e-2	-1.0e-2	0
4.3e-1	5.0e-2	4.0e-2	0	0	0	1	3.0e-1	1.0e-1	0	1.0e-2	5.0e-2	0
1.3e-1	3.0e-2	3.0e-2	0	1.0e-2	0	3.0e-1	1	3.1e-1	0	5.0e-2	1.6e-1	0
3.0e-2	1.0e-2	2.0e-2	-2.0e-2	0	1.0e-2	1.0e-1	3.1e-1	1	-1.0e-2	1.5e-1	4.9e-1	0
0	-1.0e-2	0	0	1.0e-2	-2.0e-2	0	0	-1.0e-2	1	0	-2.0e-2	0
0	0	1.0e-2	0	0	1.0e-2	1.0e-2	5.0e-2	1.5e-1	0	1	3.2e-1	0
1.0e-2	0	2.0e-2	-1.0e-2	0	-1.0e-2	5.0e-2	1.6e-1	4.9e-1	-2.0e-2	3.2e-1	1	0

Table E.3: Example 1: Spearman's rank correlation matrix for input parameters using Strategy 1.

	DPO	DSO	DCO	DPI	DSI	DCI	BPO	BSO	BCO	BPI	BSI	BCI
1	2.3e-1	1.4e-1	0	3.8e-1	2.1e-1	4.6e-1	1.6e-1	7.0e-2	7.0e-2	-5.0e-2	1.1e-1	7.0e-2
2.3e-1	1	3.1e-1	-5.0e-2	2.4e-1	3.3e-1	-2.0e-2	4.0e-1	3.1e-1	3.1e-1	-8.0e-2	-1.2e-1	2.4e-1
1.4e-1	3.1e-1	1	2.6e-1	1.2e-1	5.9e-1	6.0e-2	2.3e-1	3.2e-1	3.2e-1	-1.9e-1	-1.0e-2	2.5e-1
0	-5.0e-2	2.6e-1	1	-3.0e-2	1.9e-1	4.0e-2	-1.0e-2	4.0e-2	4.0e-2	-1.3e-1	9.0e-2	3.0e-2
3.8e-1	2.4e-1	1.2e-1	-3.0e-2	1	1.3e-1	2.2e-1	2.8e-1	-2.0e-2	-2.0e-2	-3.0e-2	1.4e-1	-3.0e-2
2.1e-1	3.3e-1	5.9e-1	1.9e-1	1.3e-1	1	1.1e-1	1.5e-1	4.6e-1	4.6e-1	-1.5e-1	-6.0e-2	4.3e-1
4.6e-1	-2.0e-2	6.0e-2	4.0e-2	2.2e-1	1.1e-1	1	1.7e-1	7.0e-2	7.0e-2	-2.0e-2	4.1e-1	1.2e-1
1.6e-1	4.0e-1	2.3e-1	-1.0e-2	2.8e-1	1.5e-1	1.7e-1	1	1.6e-1	1.6e-1	-7.0e-2	1.5e-1	1.0e-1
7.0e-2	3.1e-1	3.2e-1	4.0e-2	-2.0e-2	4.6e-1	7.0e-2	1.6e-1	1	1	-6.0e-2	-5.0e-2	8.2e-1
-5.0e-2	-8.0e-2	-1.9e-1	-1.3e-1	-3.0e-2	-1.5e-1	-2.0e-2	-7.0e-2	-6.0e-2	-6.0e-2	1	-3.0e-2	-4.0e-2
1.1e-1	-1.2e-1	-1.0e-2	9.0e-2	1.4e-1	-6.0e-2	4.1e-1	1.5e-1	-5.0e-2	-5.0e-2	-3.0e-2	1	-1.0e-2
7.0e-2	2.4e-1	2.5e-1	3.0e-2	-3.0e-2	4.3e-1	1.2e-1	1.0e-1	8.2e-1	8.2e-1	-4.0e-2	-1.0e-2	1

Table E.4: *Example 1: Spearman's rank correlation matrix for input parameters using Strategy 2.*

E.3 Example 2

Table E.5 presents the conditional probabilities among elicitation variables obtained from the experts.

From Table E.5 it is observed that for some questions some experts express a positive correlation whereas other experts are indifferent. Overall, most of the assessments of the experts are in line with each other. Table E.6 presents the Spearman's rank correlations based on the conditional probabilities using the $\rho\pi$ -table.

Comparing the Spearman's rank correlation matrices resulting from Strategy 1 (Table E.7) and Strategy 2 (Table E.3), a large difference is observed; the Spearman's rank correlations for Strategy 1 are very close to zero, whereas the majority of the Spearman's rank correlations resulting from Strategy 2 are not.

Elicitation questions		$\pi_{\frac{1}{2}, \frac{1}{2}}(Y_1, Y_2)$								
Y_1	Y_2	Exp.1	Exp.2	Exp.3	Exp.4	Exp.5	Exp.6	Exp.7	Exp.8	Exp.9
$Y_{\text{Bone},1}$	$Y_{\text{Bone},2}$	0.55	0.95	0.80	0.82	0.50	0.75	0.50	0.95	0.95
$Y_{\text{Colon},1}$	$Y_{\text{Colon},2}$	0.85	0.80	0.80	0.82	0.50	0.90	0.50	0.95	0.95
$Y_{\text{Breast},1}$	$Y_{\text{Breast},2}$	0.85	0.80	0.80	0.82	0.50	0.75	0.50	0.95	0.95
$Y_{\text{Leukemia},1}$	$Y_{\text{Leukemia},2}$	0.55	0.95	0.80	0.89	0.75	0.50	0.50	0.95	0.95
$Y_{\text{Liver},1}$	$Y_{\text{Liver},2}$	0.85	0.70	0.80	0.82	0.50	0.90	0.50	0.95	0.95
$Y_{\text{Lung},1}$	$Y_{\text{Lung},2}$	0.85	0.80	0.80	0.82	0.50	0.90	0.50	0.95	0.95
$Y_{\text{Pancreas},1}$	$Y_{\text{Pancreas},2}$	0.85	0.60	0.80	0.82	0.50	0.90	0.50	0.90	0.95
$Y_{\text{Skin},1}$	$Y_{\text{Skin},2}$	0.85	0.60	0.80	0.82	0.50	0.90	0.50	0.95	0.95
$Y_{\text{Stomach},1}$	$Y_{\text{Stomach},2}$	0.85	0.80	0.80	0.82	0.50	0.90	0.50	0.95	0.95
$Y_{\text{Thyroid},1}$	$Y_{\text{Thyroid},2}$	0.85	0.80	0.80	0.82	0.50	0.90	0.50	0.95	0.95
$Y_{\text{Remainder},1}$	$Y_{\text{Remainder},2}$	0.85	0.90	0.80	0.82	0.50	0.90	0.50	0.95	0.95
$Y_{\text{Bone},1}$	$Y_{\text{Bone},3}$	0.55	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Colon},1}$	$Y_{\text{Colon},3}$	0.55	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Breast},1}$	$Y_{\text{Breast},3}$	0.55	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Leukemia},1}$	$Y_{\text{Leukemia},3}$	0.65	0.60	0.80	0.77	0.50	0.90	0.50	0.50	0.75
$Y_{\text{Liver},1}$	$Y_{\text{Liver},3}$	0.55	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Lung},1}$	$Y_{\text{Lung},3}$	0.55	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Pancreas},1}$	$Y_{\text{Pancreas},3}$	0.55	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Skin},1}$	$Y_{\text{Skin},3}$	0.60	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Stomach},1}$	$Y_{\text{Stomach},3}$	0.55	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Thyroid},1}$	$Y_{\text{Thyroid},3}$	0.60	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Remainder},1}$	$Y_{\text{Remainder},3}$	0.55	0.60	0.80	0.80	0.50	0.80	0.50	0.50	0.75
$Y_{\text{Bone},2}$	$Y_{\text{Lung},2}$	0.55	0.50	0.50	0.67	0.50	0.60	0.50	0.50	0.55
$Y_{\text{Colon},2}$	$Y_{\text{Lung},2}$	0.70	0.50	0.50	0.67	0.50	0.70	0.50	0.50	0.65
$Y_{\text{Breast},2}$	$Y_{\text{Lung},2}$	0.70	0.50	0.50	0.67	0.50	0.70	0.50	0.50	0.65
$Y_{\text{Leukemia},2}$	$Y_{\text{Lung},2}$	0.55	0.50	0.50	0.60	0.50	0.60	0.50	0.50	0.65
$Y_{\text{Skin},2}$	$Y_{\text{Lung},2}$	0.65	0.50	0.50	0.67	0.50	0.50	0.50	0.50	0.55
$Y_{\text{Thyroid},2}$	$Y_{\text{Lung},2}$	0.65	0.50	0.50	0.67	0.50	0.50	0.50	0.50	0.65
$Y_{\text{Remainder},2}$	$Y_{\text{Lung},2}$	0.70	0.50	0.50	0.67	0.50	0.60	0.50	0.50	0.65
$Y_{\text{Stomach},2}$	$Y_{\text{Colon},2}$	0.70	0.50	0.50	0.67	0.50	0.75	0.50	0.50	0.65
$Y_{\text{Liver},2}$	$Y_{\text{Colon},2}$	0.70	0.50	0.50	0.67	0.50	0.75	0.50	0.50	0.65
$Y_{\text{Pancreas},2}$	$Y_{\text{Colon},2}$	0.70	0.50	0.50	0.67	0.50	0.70	0.50	0.50	0.55

Table E.5: *Example 2: conditional probability assessments $\pi_{\frac{1}{2}, \frac{1}{2}}(Y_1, Y_2)$ given by 9 experts from the Late Health expert panel.*

Elicitation questions		$\rho(Y_1, Y_2)$								
Y_1	Y_2	Exp.1	Exp.2	Exp.3	Exp.4	Exp.5	Exp.6	Exp.7	Exp.8	Exp.9
$Y_{Bone,1}$	$Y_{Bone,2}$	0.12	0.94	0.71	0.74	0	0.61	0	0.94	0.94
$Y_{Colon,1}$	$Y_{Colon,2}$	0.79	0.71	0.71	0.74	0	0.88	0	0.94	0.94
$Y_{Breast,1}$	$Y_{Breast,2}$	0.79	0.71	0.71	0.74	0	0.61	0	0.94	0.94
$Y_{Leukemia,1}$	$Y_{Leukemia,2}$	0.12	0.94	0.71	0.86	0	0.61	0	0.94	0.94
$Y_{Liver,1}$	$Y_{Liver,2}$	0.79	0.45	0.71	0.74	0	0.88	0	0.94	0.94
$Y_{Lung,1}$	$Y_{Lung,2}$	0.79	0.71	0.71	0.74	0	0.88	0	0.94	0.94
$Y_{Pancreas,1}$	$Y_{Pancreas,2}$	0.79	0.24	0.71	0.74	0	0.88	0	0.88	0.94
$Y_{Skin,1}$	$Y_{Skin,2}$	0.79	0.24	0.71	0.74	0	0.88	0	0.94	0.94
$Y_{Stomach,1}$	$Y_{Stomach,2}$	0.79	0.71	0.71	0.74	0	0.88	0	0.94	0.94
$Y_{Thyroid,1}$	$Y_{Thyroid,2}$	0.79	0.71	0.71	0.74	0	0.88	0	0.94	0.94
$Y_{Remainder,1}$	$Y_{Remainder,2}$	0.79	0.88	0.71	0.74	0	0.88	0	0.94	0.94
$Y_{Bone,1}$	$Y_{Bone,3}$	0.12	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Colon,1}$	$Y_{Colon,3}$	0.12	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Breast,1}$	$Y_{Breast,3}$	0.12	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Leukemia,1}$	$Y_{Leukemia,3}$	0.35	0.24	0.71	0.65	0	0.88	0	0	0.61
$Y_{Liver,1}$	$Y_{Liver,3}$	0.12	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Lung,1}$	$Y_{Lung,3}$	0.12	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Pancreas,1}$	$Y_{Pancreas,3}$	0.12	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Skin,1}$	$Y_{Skin,3}$	0.24	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Stomach,1}$	$Y_{Stomach,3}$	0.12	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Thyroid,1}$	$Y_{Thyroid,3}$	0.24	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Remainder,1}$	$Y_{Remainder,3}$	0.12	0.24	0.71	0.71	0	0.71	0	0	0.61
$Y_{Bone,2}$	$Y_{Lung,2}$	0.12	0	0	0.39	0	0.24	0	0	0.12
$Y_{Colon,2}$	$Y_{Lung,2}$	0.45	0	0	0.39	0	0.45	0	0	0.35
$Y_{Breast,2}$	$Y_{Lung,2}$	0.45	0	0	0.39	0	0.45	0	0	0.35
$Y_{Leukemia,2}$	$Y_{Lung,2}$	0.12	0	0	0.24	0	0.24	0	0	0.35
$Y_{Skin,2}$	$Y_{Lung,2}$	0.35	0	0	0.39	0	0	0	0	0.12
$Y_{Thyroid,2}$	$Y_{Lung,2}$	0.35	0	0	0.39	0	0	0	0	0.35
$Y_{Remainder,2}$	$Y_{Lung,2}$	0.45	0	0	0.39	0	0.24	0	0	0.35
$Y_{Stomach,2}$	$Y_{Colon,2}$	0.45	0	0	0.39	0	0.61	0	0	0.35
$Y_{Liver,2}$	$Y_{Colon,2}$	0.45	0	0	0.39	0	0.61	0	0	0.35
$Y_{Pancreas,2}$	$Y_{Colon,2}$	0.45	0	0	0.39	0	0.45	0	0	0.12

Table E.6: Example 2: Spearman's $\rho(Y_1, Y_2)$ for 9 experts from the Late Health effects panel.

BM	Bo	Br	Lu	St	Co	Li	Pa	Th	Sk	OC
1	-2.0e-2	-1.0e-2	3.0e-2	0	0	-5.0e-2	1.0e-2	-6.0e-2	0	-3.0e-2
-2.0e-2	1	4.0e-2	-4.0e-2	-2.0e-2	1.0e-2	1.0e-2	1.0e-2	-1.0e-2	-4.0e-2	-2.0e-2
-1.0e-2	4.0e-2	1	-2.0e-2	1.0e-2	-6.0e-2	-4.0e-2	-1.0e-2	-3.0e-2	0	-1.0e-2
3.0e-2	-4.0e-2	-2.0e-2	1	5.0e-2	3.0e-2	-3.0e-2	1.0e-2	3.0e-2	-1.0e-2	-5.0e-2
0	-2.0e-2	1.0e-2	5.0e-2	1	2.0e-2	-1.0e-2	1.0e-2	-1.0e-2	-9.0e-2	-4.0e-2
0	1.0e-2	-6.0e-2	3.0e-2	2.0e-2	1	-2.0e-2	-6.0e-2	5.0e-2	3.0e-2	-4.0e-2
-5.0e-2	1.0e-2	-4.0e-2	-3.0e-2	-1.0e-2	-2.0e-2	1	-1.0e-2	-3.0e-2	1.0e-2	2.0e-2
1.0e-2	1.0e-2	-1.0e-2	1.0e-2	1.0e-2	-6.0e-2	-1.0e-2	1	-2.0e-2	0	-1.0e-2
-6.0e-2	-1.0e-2	-3.0e-2	3.0e-2	-1.0e-2	5.0e-2	-3.0e-2	-2.0e-2	1	-2.0e-2	-3.0e-2
0	-4.0e-2	0	-1.0e-2	-9.0e-2	3.0e-2	1.0e-2	0	-2.0e-2	1	1.0e-2
-3.0e-2	-2.0e-2	-1.0e-2	-5.0e-2	-4.0e-2	-4.0e-2	-2.0e-2	-1.0e-2	-3.0e-2	1.0e-2	1

Table E.7: Example 2: Spearman's rank correlation matrix among target variables using Strategy 1.

BM	Bo	Br	Lu	St	Co	Li	Pa	Th	Sk	OC
1	1.0e-1	3.0e-2	1.3e-1	1.6e-1	1.5e-1	1.4e-1	1.0e-1	1.2e-1	-1.0e-1	1.1e-1
1.0e-1	1	8.0e-2	3.2e-1	2.9e-1	3.5e-1	2.7e-1	2.1e-1	3.4e-1	-2.1e-1	3.2e-1
3.0e-2	8.0e-2	1	2.0e-1	8.0e-2	7.0e-2	1.3e-1	7.0e-2	1.0e-1	-1.0e-2	1.0e-1
1.3e-1	3.2e-1	2.0e-1	1	3.4e-1	3.8e-1	3.5e-1	2.2e-1	3.5e-1	-2.4e-1	3.4e-1
1.6e-1	2.9e-1	8.0e-2	3.4e-1	1	3.7e-1	4.1e-1	2.0e-1	3.9e-1	-2.5e-1	3.5e-1
1.5e-1	3.5e-1	7.0e-2	3.8e-1	3.7e-1	1	3.7e-1	2.1e-1	3.8e-1	-2.8e-1	3.6e-1
1.4e-1	2.7e-1	1.3e-1	3.5e-1	4.1e-1	3.7e-1	1	2.2e-1	4.3e-1	-1.9e-1	3.8e-1
1.0e-1	2.1e-1	7.0e-2	2.2e-1	2.0e-1	2.1e-1	2.2e-1	1	2.4e-1	-1.2e-1	2.1e-1
1.2e-1	3.4e-1	1.0e-1	3.5e-1	3.9e-1	3.8e-1	4.3e-1	2.4e-1	1	-2.7e-1	4.0e-1
-1.0e-1	-2.1e-1	-1.0e-2	-2.4e-1	-2.5e-1	-2.8e-1	-1.9e-1	-1.2e-1	-2.7e-1	1	-2.2e-1
1.1e-1	3.2e-1	1.0e-1	3.4e-1	3.5e-1	3.6e-1	3.8e-1	2.1e-1	4.0e-1	-2.2e-1	1

Table E.8: *Example 2: Spearman's rank correlation matrix among target variables using Strategy 2.*

Appendix F

Calibration with Uncertain observations: data

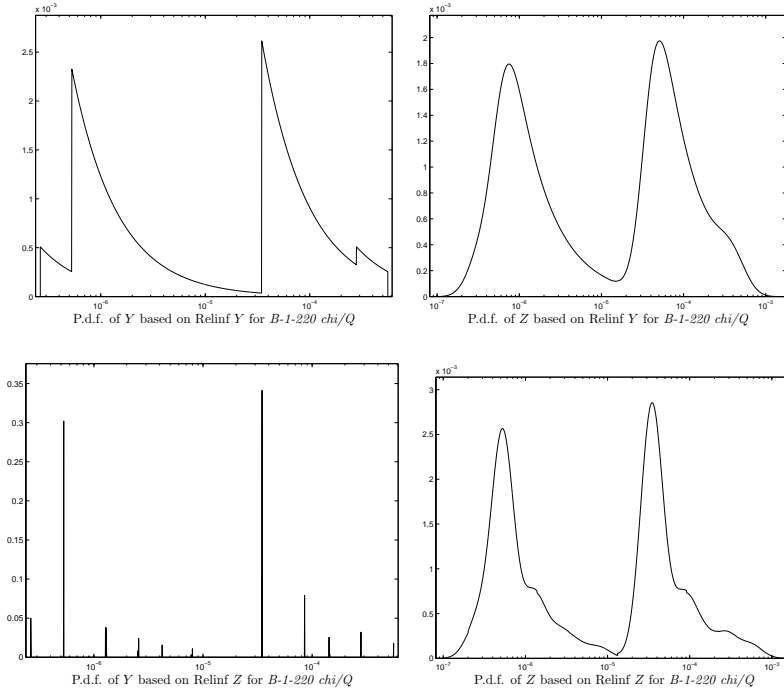
In this appendix the data used in Chapter 3 are presented together with figures displaying the various probability density functions.

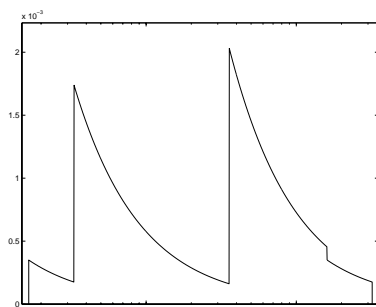
F.1 Dispersion example

The quantile assessments of the DM in Table F.1 are taken from [23],[27]. The reader is referred to these references for a detailed explanation of the abbreviations/seed variables and descriptions of the various experiments. It is assumed that the variability of the measurement is represented adequately via a log-normal distribution, specified by median and error factor EF_{95} . The medians listed in Table F.1 resulted from the various experiments and the information contained on the error factor EF_{95} are specified by the author; the abbreviations L and U in the column Background density denote the log-uniform and uniform background density, respectively.

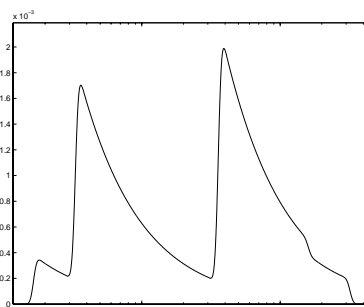
The figures show the various probability density functions (p.d.f.) for the different elicitation variables resulting from the Relinf Y and Relinf Z approach for $c = 1$, see Section 3.3.1.

Seed var. Questions	Quantile assessments for the DM			Meas. variability		Background density
	$y_{5\%}$	$y_{50\%}$	$y_{95\%}$	Median	EF_{95}	
B-1-220 chi/Q	5.28e-7	3.46e-5	2.79e-4	2.39e-5	1.67	L
B-1-315 chi/Q	3.31e-6	3.57e-5	1.60e-4	4.19e-5	1.07	L
B-2-220 chi/Q	3.56e-7	4.31e-5	2.39e-4	3.07e-5	1.14	L
B-2-315 chi/Q	2.10e-6	3.44e-5	1.55e-4	5.38e-5	1.07	L
B-3-300 chi/Q	2.15e-6	1.62e-5	9.44e-5	5.26e-5	1.05	L
B-3-600 chi/Q	1.05e-6	6.22e-6	4.95e-5	5.19e-5	1.05	L
B-4-300 chi/Q	3.05e-6	5.15e-5	3.71e-4	2.10e-5	1.19	L
B-4-600 chi/Q	5.30e-6	2.78e-5	1.85e-4	3.79e-5	1.06	L
B-5-600 sig-y	1.51e+1	3.85e+1	1.33e+2	2.80e+1	1.07	U
B-5-600 sig-z	4.04	1.22e+1	2.91e+1	8.36	1.08	U

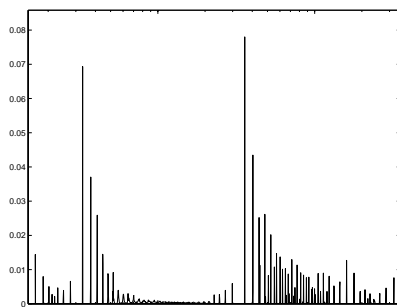
Table F.1: *Quantile assessments for the DM and measurement variability.*



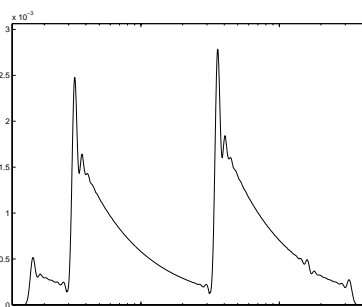
P.d.f. of Y based on Relinf Y for B-1-315 χ/Q



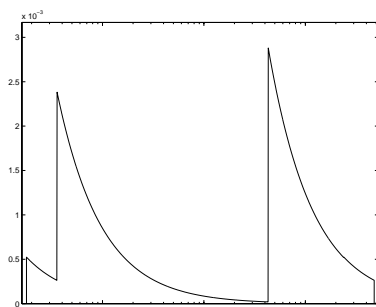
P.d.f. of Z based on Relinf Y for B-1-315 χ/Q



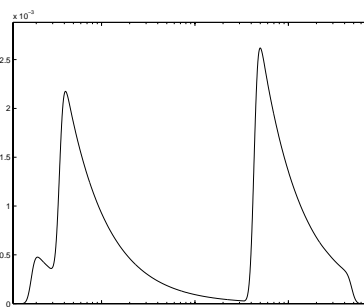
P.d.f. of Y based on Relinf Z for B-1-315 χ/Q



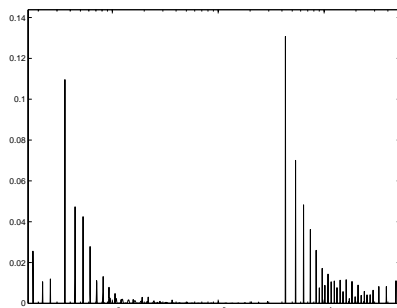
P.d.f. of Z based on Relinf Z for B-1-315 χ/Q



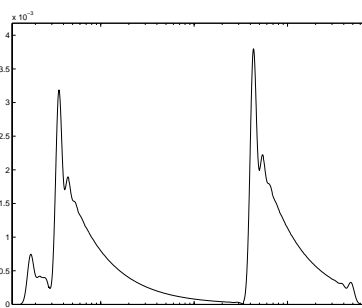
P.d.f. of Y based on Relinf Y for B-2-220 χ/Q



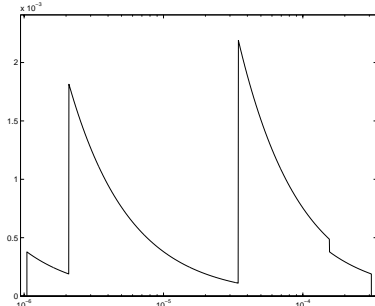
P.d.f. of Z based on Relinf Y for B-2-220 χ/Q



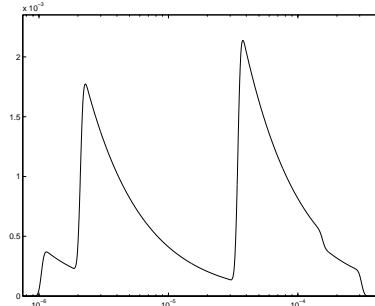
P.d.f. of Y based on Relinf Z for B-2-220 χ/Q



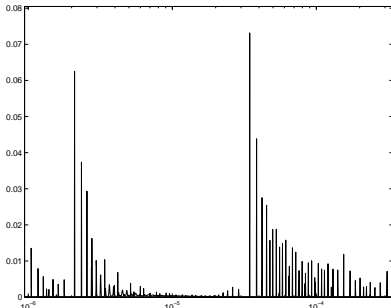
P.d.f. of Z based on Relinf Z for B-2-220 χ/Q



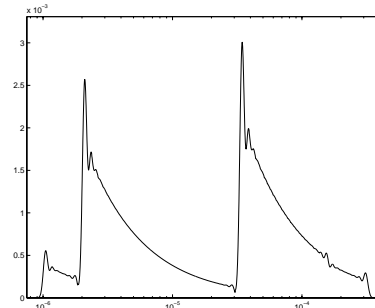
P.d.f. of Y based on Relin Y for $B-2-315 \chi/Q$



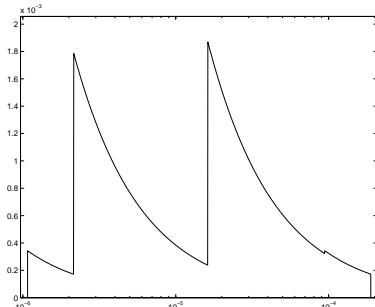
P.d.f. of Z based on Relin Y for $B-2-315 \chi/Q$



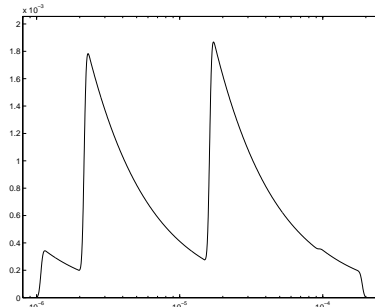
P.d.f. of Y based on Relin Z for $B-2-315 \chi/Q$



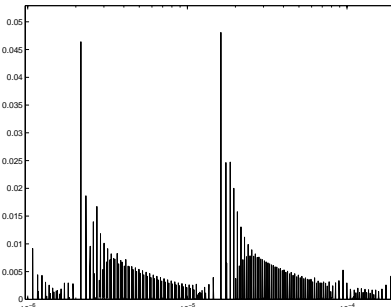
P.d.f. of Z based on Relin Z for $B-2-315 \chi/Q$



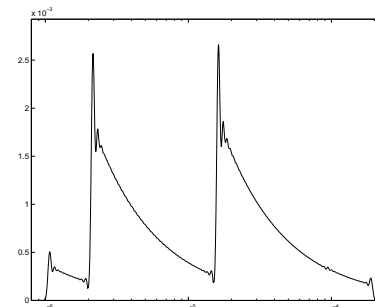
P.d.f. of Y based on Relin Y for $B-3-300 \chi/Q$



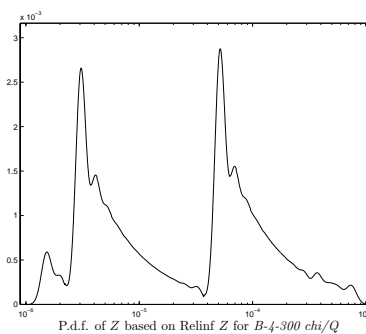
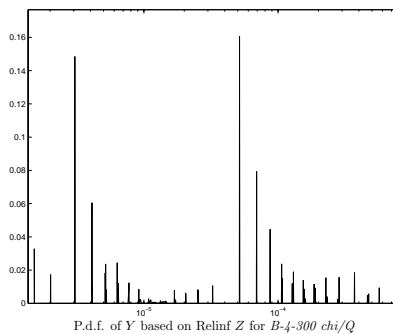
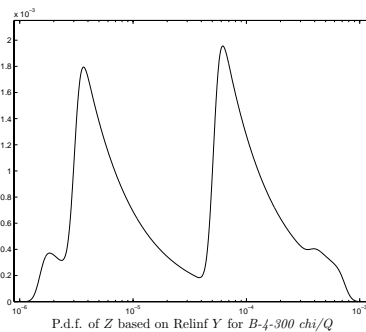
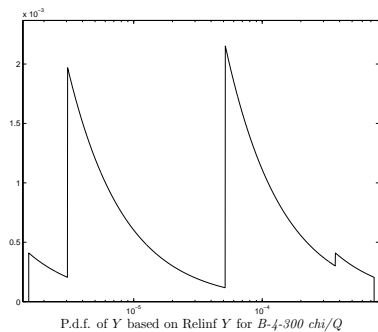
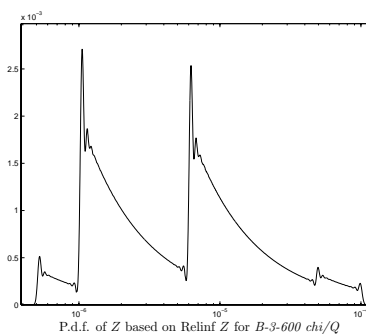
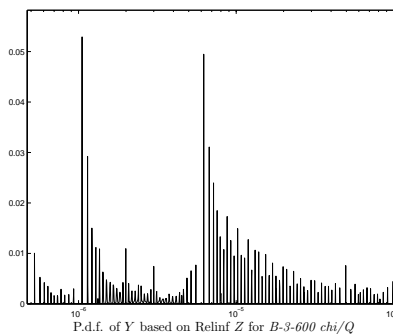
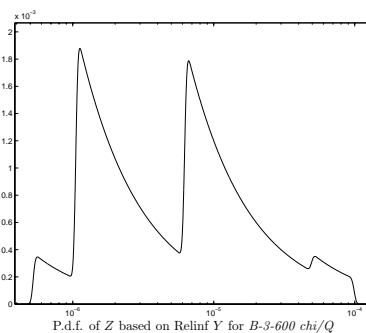
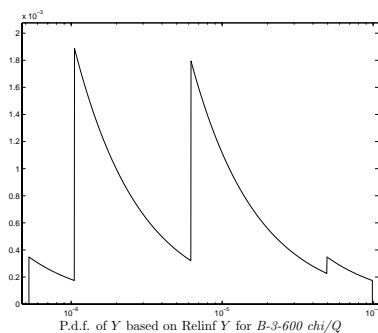
P.d.f. of Z based on Relin Y for $B-3-300 \chi/Q$

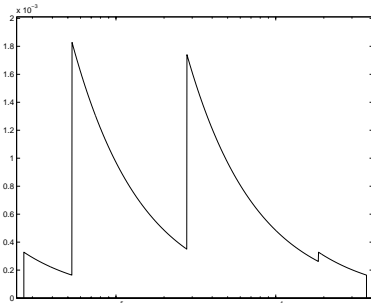


P.d.f. of Y based on Relin Z for $B-3-300 \chi/Q$

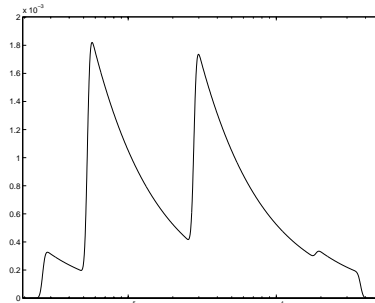


P.d.f. of Z based on Relin Z for $B-3-300 \chi/Q$

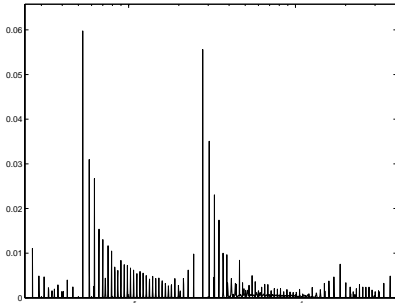




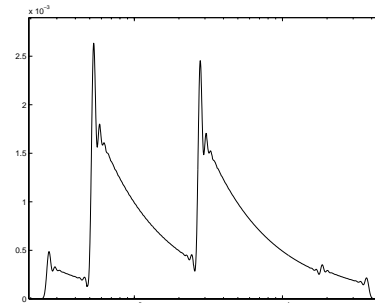
P.d.f. of Y based on Relinf Y for B-4-600 χ/Q



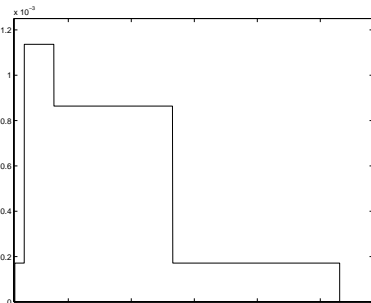
P.d.f. of Z based on Relinf Y for B-4-600 χ/Q



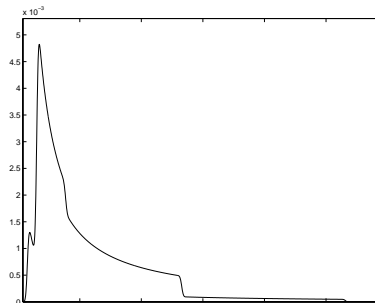
P.d.f. of Y based on Relinf Z for B-4-600 χ/Q



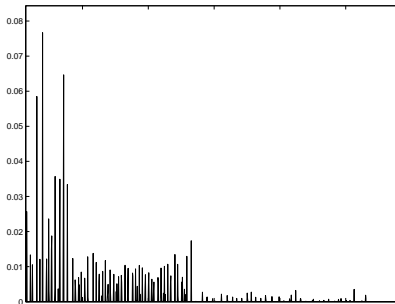
P.d.f. of Z based on Relinf Z for B-4-600 χ/Q



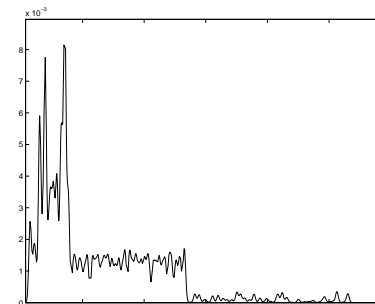
P.d.f. of Y based on Relinf Y for B-5-600 sig-y



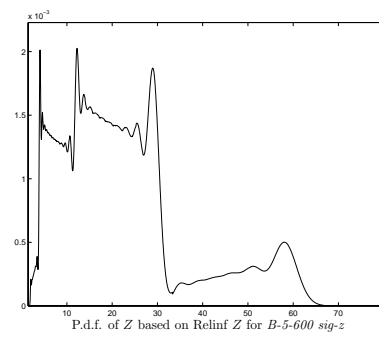
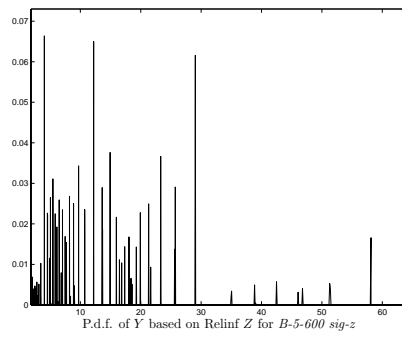
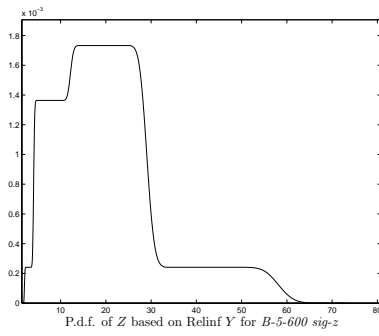
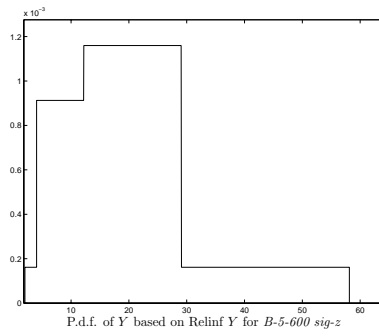
P.d.f. of Z based on Relinf Y for B-5-600 sig-y



P.d.f. of Y based on Relinf Z for B-5-600 sig-y



P.d.f. of Z based on Relinf Z for B-5-600 sig-y



F.2 Public Works and Water Management example

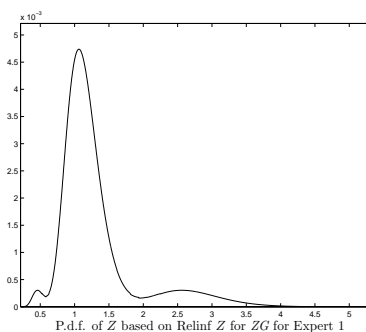
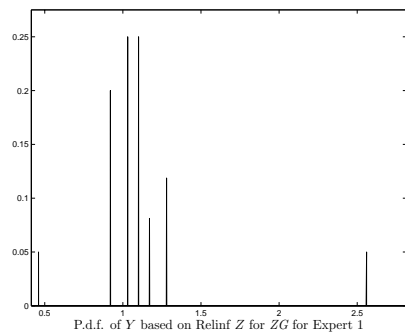
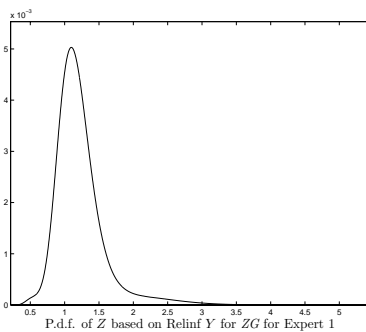
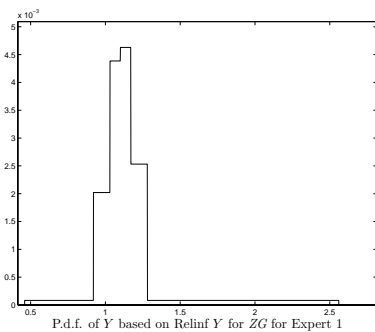
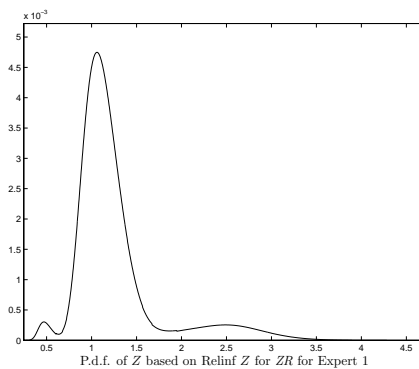
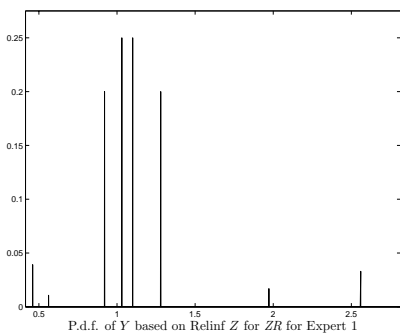
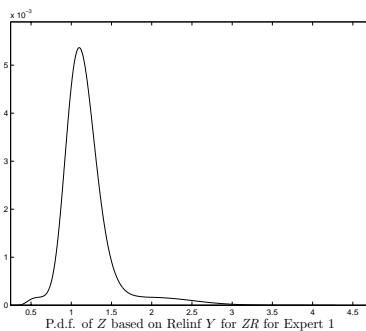
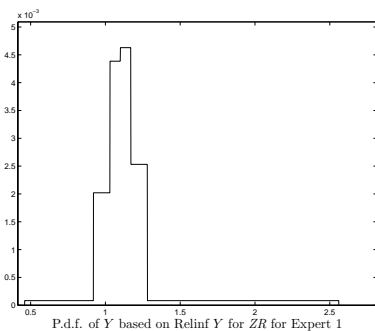
	Experimental results	Median	EF_{95}
ZR	[0.92,1.39,1.20,1.18,1.35,0.97,1]	1.18	1.25
ZG	[1.14,1.29,1.55,1.49,0.98,1.34,1.68,1.06]	1.31	1.35
ZD	[1.31,1.17,1.57,1.46,1.50,0.99]	1.38	1.27
HS	[1.1,0.94,1.09,0.99,1.19,1.09,1.12]	1.09	1.12
TS	[0.95,1.03,1.02,0.88,0.89,1.01,1.01]	1.01	1.07
mod1	[0.35,0.18,0.32,0.22,0.44]	3.20e-1	1.63
mod2	[6.4,3.6,7.3,7.5]	6.85	1.53
mod3	[9.30e+1,7.70e+1,6.40e+1]	7.70e+1	1.36

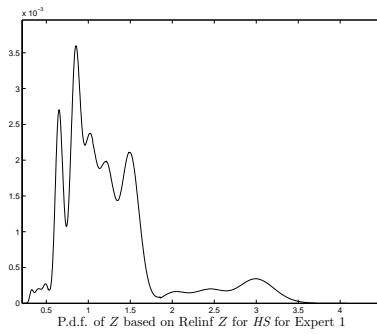
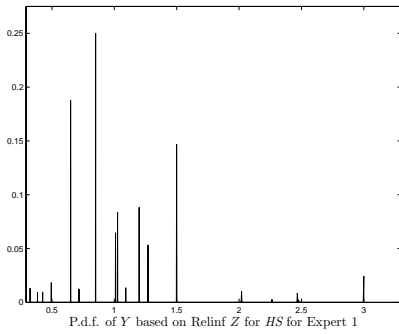
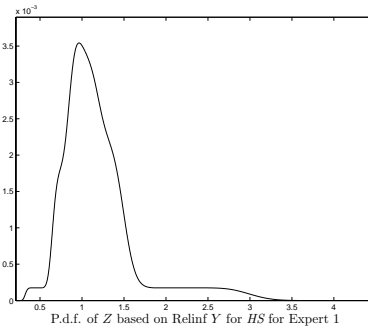
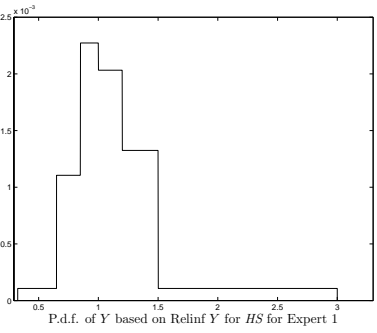
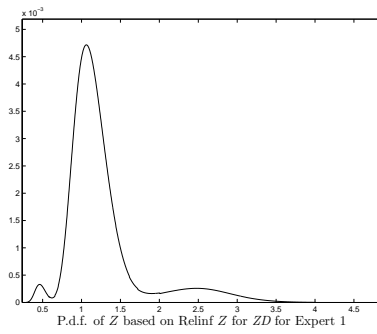
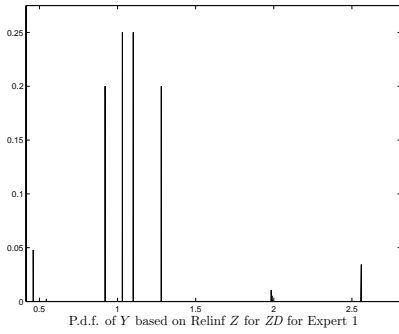
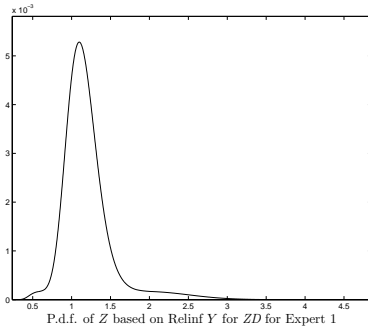
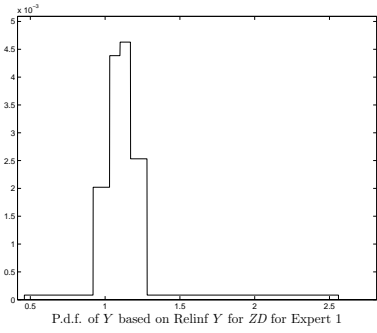
Table F.2: Overview of experimental results, median and EF_{95} for each seed variable question.

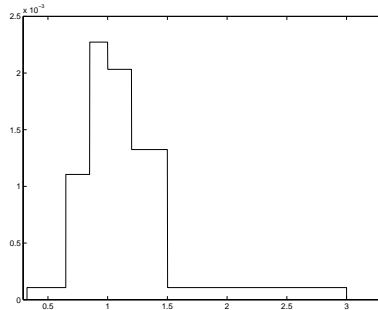
Seed var.	Quantile assessments for Expert 1					Measurement	
Questions	$y_{5\%}$	$y_{25\%}$	$y_{50\%}$	$y_{75\%}$	$y_{95\%}$	Median	EF_{95}
ZR	9.20e-1	1.03	1.10	1.17	1.28	1.18	1.25
ZG	9.20e-1	1.03	1.10	1.17	1.28	1.31	1.35
ZD	9.20e-1	1.03	1.10	1.17	1.28	1.38	1.27
HS	6.50e-1	8.50e-1	1	1.20	1.50	1.09	1.12
TS	6.50e-1	8.50e-1	1	1.20	1.50	1.01	1.07
mod1	1.00e-1	3.00e-1	5.00e-1	1	3	3.20e-1	1.63
mod2	1	3	5	1.00e+1	3.00e+1	6.85	1.53
mod3	1.00e+1	3.00e+1	5.00e+1	1.00e+2	3.00e+2	7.70e+1	1.36
Seed var.	Quantile assessments for Expert 6					Measurement	
Questions	$y_{5\%}$	$y_{25\%}$	$y_{50\%}$	$y_{75\%}$	$y_{95\%}$	Median	EF_{95}
ZR	2.00e-1	7.00e-1	1	1.30	1.80	1.18	1.25
ZG	2.00e-1	7.00e-1	1	1.30	1.80	1.31	1.35
ZD	2.00e-1	7.00e-1	1	1.30	1.80	1.38	1.27
HS	4.00e-1	7.50e-1	1	1.25	1.60	1.09	1.12
TS	5.00e-1	7.50e-1	1	1.50	2	1.01	1.07
mod1	5.00e-1	8.00e-1	1	1.50	2.50	3.20e-1	1.63
mod2	5	8	1.00e+1	1.50e+1	2.50e+1	6.85	1.53
mod3	5.00e+1	8.00e+1	1.00e+2	1.50e+2	2.50e+2	7.70e+1	1.36
Seed var.	Quantile assessments for Expert 10					Measurement	
Questions	$y_{5\%}$	$y_{25\%}$	$y_{50\%}$	$y_{75\%}$	$y_{95\%}$	Median	EF_{95}
ZR	5.00e-1	8.00e-1	1	1.20	1.50	1.18	1.25
ZG	5.00e-1	8.00e-1	1	1.20	1.50	1.31	1.35
ZD	5.00e-1	8.00e-1	1	1.20	1.50	1.38	1.27
HS	6.00e-1	8.50e-1	1	1.15	1.40	1.09	1.12
TS	7.00e-1	8.50e-1	1	1.15	1.30	1.01	1.07
mod1	1.00e-1	5.00e-1	1	4	6	3.20e-1	1.63
mod2	4	6	1.00e+1	2.00e+1	3.00e+1	6.85	1.53
mod3	3.00e+1	6.00e+1	1.00e+2	1.50e+2	2.00e+2	7.70e+1	1.36

Table F.3: Overview of expert quantile assessments and measurement variability.

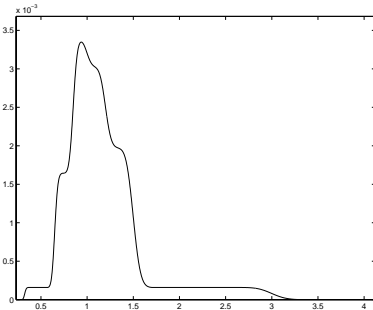
The figures below show the resulting probability density functions (p.d.f.) for the seed variable questions for Expert 1 only.



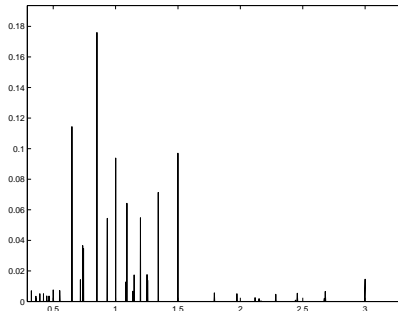




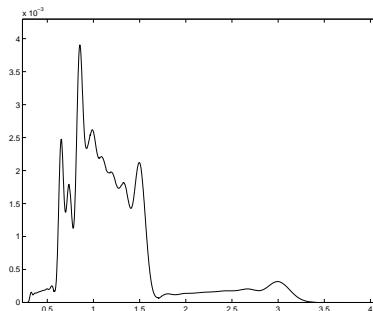
P.d.f. of Y based on Relinf Y for TS for Expert 1



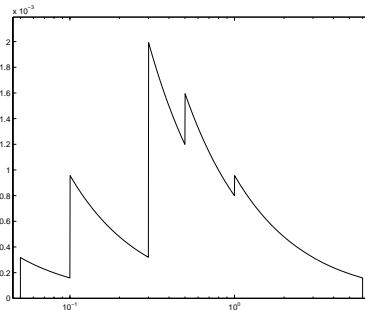
P.d.f. of Z based on Relinf Y for TS for Expert 1



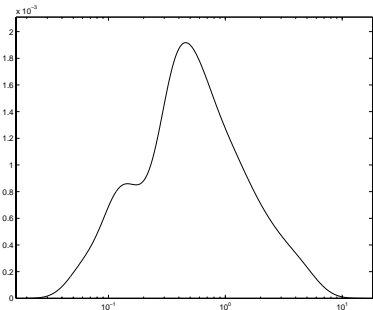
P.d.f. of Y based on Relinf Z for TS for Expert 1



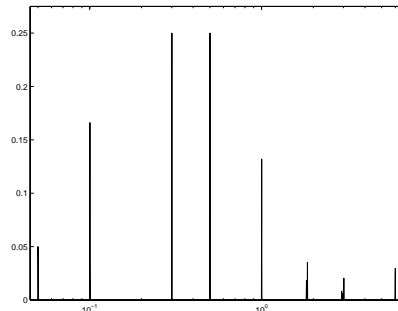
P.d.f. of Z based on Relinf Z for TS for Expert 1



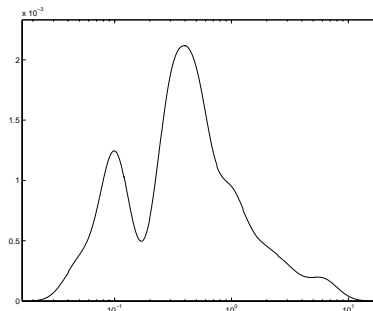
P.d.f. of Y based on Relinf Y for mod1 for Expert 1



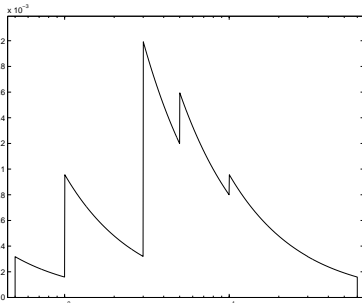
P.d.f. of Z based on Relinf Y for mod1 for Expert 1



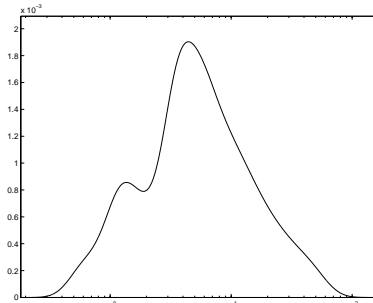
P.d.f. of Y based on Relinf Z for mod1 for Expert 1



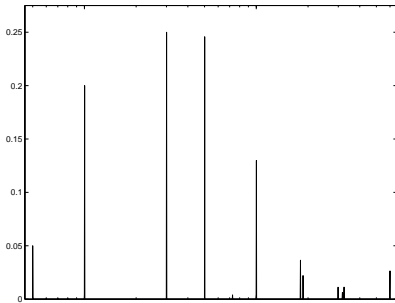
P.d.f. of Z based on Relinf Z for mod1 for Expert 1



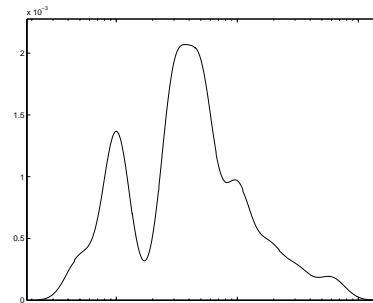
P.d.f. of Y based on Relinf Y for $mod2$ for Expert 1



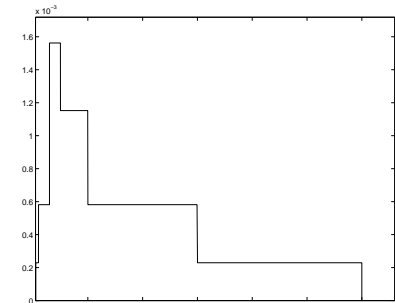
P.d.f. of Z based on Relinf Y for $mod2$ for Expert 1



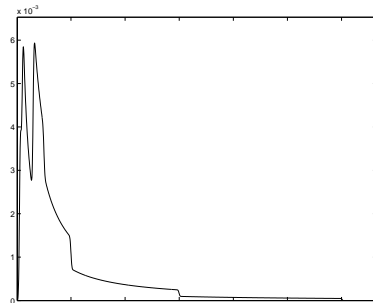
P.d.f. of Y based on Relinf Z for $mod2$ for Expert 1



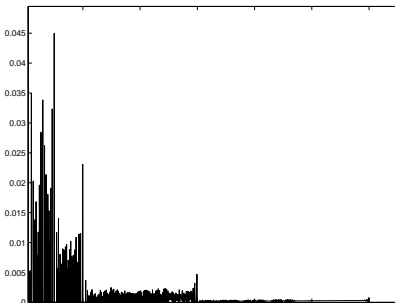
P.d.f. of Z based on Relinf Z for $mod2$ for Expert 1



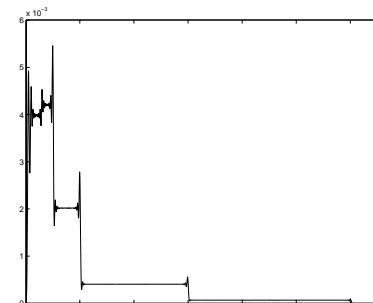
P.d.f. of Y based on Relinf Y for $mod3$ for Expert 1



P.d.f. of Z based on Relinf Y for $mod3$ for Expert 1



P.d.f. of Y based on Relinf Z for $mod3$ for Expert 1



P.d.f. of Z based on Relinf Z for $mod3$ for Expert 1

Expert 1	
Seed variable question	Interquantile intervals which may contain observation and q_{j,l_j} for Relinf Y and Relinf Z
ZR	Relinf Y
	$[I_{1,1}, I_{1,2}, I_{1,3}, I_{1,4}, I_{1,5}, I_{1,6}]$ $[1.13e-3, 1.08e-1, 2.68e-1, 3.41e-1, 2.72e-1, 9.61e-3]$
	Relinf Z
ZG	Relinf Y
	$[I_{2,1}, I_{2,2}, I_{2,3}, I_{2,4}, I_{2,5}, I_{2,6}]$ $[1.46e-3, 9.54e-2, 2.29e-1, 3.19e-1, 3.23e-1, 3.17e-2]$
	Relinf Z
ZD	Relinf Y
	$[I_{3,1}, I_{3,2}, I_{3,3}, I_{3,4}, I_{3,5}, I_{3,6}]$ $[1.31e-4, 3.86e-2, 1.59e-1, 3.11e-1, 4.35e-1, 5.63e-2]$
	Relinf Z
HS	Relinf Y
	$[I_{4,3}, I_{4,4}, I_{4,5}]$ $[1.19e-1, 8.28e-1, 5.36e-2]$
	Relinf Z
TS	Relinf Y
	$[I_{5,3}, I_{5,4}]$ $[4.31e-1, 5.69e-1]$
	Relinf Z
mod1	Relinf Y
	$[I_{6,2}, I_{6,3}, I_{6,4}]$ $[1.50e-1, 7.66e-1, 8.41e-2]$
	Relinf Z
mod2	Relinf Y
	$[I_{7,3}, I_{7,4}, I_{7,5}]$ $[1.30e-1, 8.18e-1, 5.26e-2]$
	Relinf Z
mod3	Relinf Y
	$I_{8,4}$ 1
	Relinf Z
	$I_{8,4}$ 1

Table F.4: *Public Works and Water Management example: overview of interquantile intervals and likelihoods for Expert 1.*

Expert 6	
Seed variable question	Interquantile intervals which may contain observation and q_{j,l_j} for Relinf Y and Relinf Z
ZR	Relinf Y
	$[I_{1,3}, I_{1,4}, I_{1,5}]$ [9.10e-2, 7.36e-1, 1.73e-1]
	Relinf Z
	$[I_{1,4}, I_{1,5}]$ [7.99e-1, 2.01e-1]
ZG	Relinf Y
	$[I_{2,3}, I_{2,4}, I_{2,5}, I_{2,6}]$ [6.47e-2, 5.34e-1, 3.98e-1, 3.92e-3]
	Relinf Z
	$[I_{2,4}, I_{2,5}]$ [6.22e-1, 3.78e-1]
ZD	Relinf Y
	$[I_{3,3}, I_{3,4}, I_{3,5}, I_{3,6}]$ [1.21e-2, 4.39e-1, 5.45e-1, 3.38e-3]
	Relinf Z
	$[I_{3,4}, I_{3,5}]$ [4.69e-1, 5.31e-1]
HS	Relinf Y
	$[I_{4,3}, I_{4,4}, I_{4,5}]$ [8.33e-2, 9.00e-1, 1.66e-2]
	Relinf Z
	$[I_{4,4}, I_{4,5}]$ [9.98e-1, 1.73e-3]
TS	Relinf Y
	$[I_{5,3}, I_{5,4}]$ [4.97e-1, 5.03e-1]
	Relinf Z
	$[I_{5,3}, I_{5,4}]$ [1.66e-1, 8.34e-1]
mod1	Relinf Y
	$[I_{6,1}, I_{6,2}, I_{6,3}]$ [6.38e-1, 3.62e-1, 3.32e-4]
	Relinf Z
	$[I_{6,1}, I_{6,2}, I_{6,3}]$ [3.48e-1, 6.31e-1, 2.09e-2]
mod2	Relinf Y
	$[I_{7,1}, I_{7,2}, I_{7,3}, I_{7,4}, I_{7,5}]$ [1.17e-2, 4.74e-1, 4.26e-1, 8.77e-2, 1.84e-4]
	Relinf Z
	$[I_{7,2}, I_{7,3}, I_{7,4}]$ [2.60e-1, 5.63e-1, 1.77e-1]
mod3	Relinf Y
	$[I_{8,2}, I_{8,3}]$ [9.70e-1, 3.0e-2]
	Relinf Z
	$[I_{8,2}, I_{8,3}]$ [8.48e-1, 1.52e-1]

Table F.5: *Public Works and Water Management example: overview of interquantile intervals and likelihoods for Expert 6.*

Expert 10	
Seed variable question	Interquantile intervals which may contain observation and q_{j,l_j} for Relinf Y and Relinf Z
ZR	Relinf Y
	$[I_{1,2}, I_{1,3}, I_{1,4}, I_{1,5}, I_{1,6}]$ [4.36e-4, 1.11e-1, 5.48e-1, 3.38e-1, 2.47e-3]
	Relinf Z
	$[I_{1,3}, I_{1,4}, I_{1,5}]$ [2.13e-2, 7.67e-1, 2.12e-1]
ZG	Relinf Y
	$[I_{2,2}, I_{2,3}, I_{2,4}, I_{2,5}, I_{2,6}]$ [1.26e-3, 8.83e-2, 4.02e-1, 4.89e-1, 1.97e-2]
	Relinf Z
	$[I_{2,3}, I_{2,4}, I_{2,5}]$ [2.49e-2, 4.01e-1, 5.74e-1]
ZD	Relinf Y
	$[I_{3,3}, I_{3,4}, I_{3,5}, I_{3,6}]$ [1.86e-2, 2.86e-1, 6.67e-1, 2.76e-2]
	Relinf Z
	$[I_{3,4}, I_{3,5}]$ [1.94e-1, 8.06e-1]
HS	Relinf Y
	$[I_{4,3}, I_{4,4}, I_{4,5}]$ [1.01e-1, 7.60e-1, 1.39e-1]
	Relinf Z
	$[I_{4,3}, I_{4,4}]$ [6.20e-1, 9.38e-1]
TS	Relinf Y
	$[I_{5,3}, I_{5,4}]$ [3.72e-1, 6.28e-1]
	Relinf Z
	$[I_{5,3}, I_{5,4}]$ [7.93e-2, 9.21e-1]
mod1	Relinf Y
	$[I_{6,2}, I_{6,3}]$ [7.30e-1, 2.70e-1]
	Relinf Z
	$[I_{6,2}, I_{6,3}]$ [7.93e-2, 9.21e-1]
mod2	Relinf Y
	$[I_{7,1}, I_{7,2}, I_{7,3}, I_{7,4}]$ [1.99e-3, 2.75e-1, 6.57e-1, 6.64e-2]
	Relinf Z
	$[I_{7,2}, I_{7,3}, I_{7,4}]$ [7.06e-2, 6.72e-1, 2.57e-1]
mod3	Relinf Y
	$I_{8,3}$ 1
	Relinf Z
	$I_{8,3}$ 1

Table F.6: Public Works and Water Management example: overview of interquantile intervals and likelihoods for Expert 10.

Samenvatting

In het laatste decennia heeft onzekerheidsanalyse een steeds prominentere rol binnen de beslissingsanalyse ingenomen. Het is algemeen erkend dat als beslissingen genomen worden op basis van wiskundige modellen dat het verstandig is om de onzekerheid in de model invoer parameters te onderzoeken. Immers, een kleine verandering van de model invoer parameter kan resulteren in een andere beslissing.

Om een onzekerheidsanalyse uit te voeren is het noodzakelijk om een gemeenschappelijke verdeling over de onzekere model invoer parameters te specificeren. Gebruikmakend van Monte-Carlo technieken wordt deze verdeling door het model gepropageerd en een verdeling over de waarden van de model uitvoer parameters verkregen, die nader wordt geanalyseerd. Derhalve is het bepalen van de gemeenschappelijke verdeling over de onzekere model invoer parameters cruciaal in het doen van een onzekerheidsanalyse. Als experimentele gegevens over de onzekere model invoer parameters beschikbaar zijn, dan kunnen die gebruikt worden om een gemeenschappelijke verdeling te bepalen. Als er weinig of geen experimentele gegevens beschikbaar zijn, is een andere aanpak nodig. In dit proefschrift wordt de situatie waarin weinig of geen experimentele gegevens beschikbaar zijn beschouwd.

In het geval dat er weinig of geen experimentele gegevens beschikbaar zijn, is het benaderen van experts op het betreffende gebied een natuurlijke stap. Experts zijn bij uitstek diegenen die het beste overzicht over de literatuur hebben, de beschikking over wiskundige modellen hebben en in staat zijn om de beschikbare data naar waarde te schatten en te gebruiken voor interpolatie en extrapolatie. Om op een gestructureerde, open en verdedigbare manier gebruik te maken van de kennis/vaardigheden van experts is de gestructureerde Expert Meninge Methodologie ontwikkeld, [12].

In dit proefschrift worden nieuwe wiskundige technieken geïntroduceerd ter ondersteuning van de gestructureerde Expert Meninge Methodologie. Ruwweg kan het proefschrift in 4 delen onderverdeeld worden: een deel wat probabilistische inversie behandelt, een deel waarin de rol van afhankelijkheden in onzekerheidsanalyse besproken wordt, een deel wat aandacht besteedt aan het ontwikkelen van grootheden die de capaciteit van de expert meten in hoe goed een expert zijn/haar kennis kan kwantificeren en een deel

die het modeleren van onzekerheid behandelt. Voorbeelden om de verschillende wiskundige technieken te illustreren zijn van de Joint CEC/USNRC Uncertainty Analysis (Contract F13P-CT092-0023 en 93-ET-001) genomen.

Probabilistische inversie

Een van de belangrijkste elementen in de gestructureerde Expert Mening Methodologie is dat experts hun kennis kwantificeren voor (potentieel) meetbare grootheden. Derhalve kan het zijn dat de onzekere model invoer parameters niet geschikt zijn om aan een expert voor te leggen. Probabilistische inversie biedt uitkomst. Als experts hun onzekerheid kwantificeren voor (potentieel) meetbare grootheden dan kan, door probabilistische inversie toe te passen, deze kennis ‘terugvertaald’ worden naar informatie over de onzekere model invoer parameters. In Hoofdstuk 1 worden de beginselen en implementaties van probabilistische inversie technieken besproken.

Afhankelijkheden

Om een onzekerheidsanalyse te kunnen uitvoeren is het nodig om een gemeenschappelijke verdeling over de onzekere model invoer parameters te specificeren. Vaak is deze gemeenschappelijke verdeling opgebouwd uit de afzonderlijke verdelingen van, en afhankelijkheden tussen de onzekere model invoer parameters. In Hoofdstuk 2 wordt er allereerst een strategie geselecteerd die het meest geschikt is om afhankelijkheidsinformatie van experts te eliciteren. Op dit moment worden binnen de gestructureerde Expert Mening Methodologie alleen de meningen van experts gecombineerd. Twee strategieën worden geïntroduceerd die zowel de meningen van experts als afhankelijkheidsinformatie combineren. Als afsluiting wordt het gebruik van afhankelijkheidsinformatie in probabilistische inversie besproken.

Calibratie met onzekere waarnemingen

Onder de gestructureerde Expert Mening Methodologie zijn er grootheden ontwikkeld die, in statistische zin, de capaciteit van een expert meten hoe ‘goed’ een expert is in het kwantificeren van zijn onzekerheid. Experts wordt gevraagd om hun onzekerheid te kwantificeren voor (potentieel) meetbare grootheden waarvoor een meetwaarde ter beschikking is; deze meetwaarde is alleen bekend bij het team wat de onzekerheidsanalyse uitvoert en niet bij de expert. Als er een redelijke hoeveelheid van dit soort vragen beschikbaar zijn, dan kan, in statistische zin, de capaciteit van een expert in hoe ‘goed’ hij/zij is in het kwantificeren van onzekerheid gemeten worden. Omdat het reproduceren van meetwaarden vaak niet mogelijk is en ook omdat andere meetopstellingen andere meetwaarden kunnen opleveren, is de meetwaarde van een experiment ook onzeker. In Hoofdstuk 3 worden de grootheden

die de capaciteit van een expert in het kwantificeren van zijn/haar onzekerheid uitgebreid door ook de meetwaarde van een experiment als onzeker te beschouwen. Het effect van meetvariabiliteit is aangetoond gebruikmakend van voorbeelden uit het Joint CEC/USNRC Uncertainty Analysis project en een studie voor Rijkswaterstaat.

Modeleren van onzekerheid

In dit hoofdstuk worden overwegingen in het kwantificeren van de onzekerheid voor onzekere model invoer parameters en effect/overwegingen in het reduceren van de complexiteit van wiskundige modellen voor het uitvoeren van de onzekerheidsanalyse besproken. Speciale aandacht gaat uit naar een speciale klasse van acyclische compartiment modellen (ACMs). ACMs worden veel gebruikt in milieu-modelering en worden, vanwege hun grafische voorstelling, als makkelijk in het gebruik gezien. Maar ACMs verdienen speciale aandacht als ze gebruikt worden in een onzekerheidsanalyse. Net als ACMs zijn Invloeds Diagrammen (IDs) ook acyclische grafen. De relatie tussen ACMs en IDs wordt onderzocht, wat leidt tot een decompositie strategie voor complexe ACM's. De decompositie zal invloed hebben op de vraagstelling richting de experts. Maar de vraagstelling dient ten allen tijden in overeenstemming te zijn met de gestructureerde Expert Meninge Methodologie.

Toepassing van resultaten

In ieder hoofdstuk is een nieuwe wiskundige techniek geïntroduceerd ter ondersteuning van de gestructureerde Expert Meninge Methodologie. De meerderheid van deze technieken zijn al gevalideerd en praktisch gebleken in de Joint CEC/USNRC Uncertainty Analysis [25], [26], [27], [28], [29], [30], een studie voor Rijkswaterstaat [21] en in het proefschrift 'Uncertainty in predictions of thermal comfort inbuildings' [19]. De wiskundige technieken van dit proefschrift dienen in omgekeerde volgorde van presentatie gebruikt te worden in het uitvoeren van een onzekerheidsanalyse. Allereerst dient er nagedacht te worden over de overwegingen zoals gepresenteerd in Hoofdstuk 4. Ten tweede, als experts gebruikt worden om de onzekerheid te kwantificeren en hun capaciteit daarin gemeten wordt, dan kunnen de technieken zoals beschreven in Hoofdstuk 3 gebruikt worden. Ten derde, de expert meninge kunnen worden gecombineerd gebruikmakend van de geëliciteerde afhankelijkheidsinformatie en strategieën zoals in Hoofdstuk 2 geïntroduceerd. En indien noodzakelijk, kan probabilistische inversie uit Hoofdstuk 1 gebruikt worden om een verdeling over de onzekere model invoer parameters te verkrijgen.

Dankwoord

Vaak heb ik het schrijven van een proefschrift vergeleken met een solo zeiltocht rond de wereld. Er zijn tijden dat het je voor de wind gaat en er veel progressie geboekt wordt, maar er zijn ook tijden dat je in een windstille gebied ligt en er geen tot weinig progressie wordt geboekt. Een kaart van de route is niet beschikbaar en soms kies je een koers waarvan je naderhand merkt dat het niet het gewenste resultaat oplevert en dan kan het lastig zijn om weer op koers te komen. Daarom is het goed om bakens/vuurtorens te hebben waarop je kunt varen, mensen met wie je radiocontact hebt en die je kunt vragen welke koers te varen. Dat maakt het schrijven van een proefschrift niet helemaal een solo-activiteit. Veel mensen staan je (gelukkig) met raad en daad bij, wat zeker op momenten van windstille of slecht weer een stimulans is. Aan het einde van de tocht komt dan eindelijk de haven in zicht, wat betekent dat de reis voltooid is en in het geval van een promotieonderzoek, dat er een proefschrift ligt. Persoonlijk had ik het gevoel dat ik al een heel tijdje voor de haven lag, maar dat ik geen wind in de zeilen had (of dat ik de zeilen niet gehesen had!).

Nu ik aangekomen ben in de haven wil ik graag de mensen bedanken met wie ik veel radiocontact had tijdens mijn reis.

Allereerst wil ik Roger Cooke bedanken. Zonder zijn steun, motivatie, inzet en enthousiasme had ik dit proefschrift nooit tot een einde kunnen brengen. Wij hebben veel meegemaakt en ik ben ervan overtuigd dat we nog veel mee zullen meemaken. Ik heb veel van je navigatie- en zeil kwaliteiten opgestoken en overgenomen. Ik ben je zeer dankbaar voor alles wat je voor me gedaan hebt.

Louis Goossens, Europees projectleider van het Joint CEC/USNRC Uncertainty Analysis project, heeft voor mij een rol binnen het project gecreëerd die zeer belangrijk is geweest voor mijn ontwikkeling. Zijn vertrouwen in mij gaf mij de vrijheid om me verder te ontwikkelen. Vanuit België heeft Neale Kelly, opdrachtgever van het Joint CEC/USNRC Uncertainty Analysis project, er grotendeels voor gezorgd dat ik überhaupt aan een dergelijke onderneming kon beginnen. Zijn analyserend vermogen en scherpe vragen hebben er altijd voor gezorgd dat je wel twee keer nadacht voordat je iets zegt of opschrijft. Verder zijn er nog een aantal Amerikaanse navigators/zeilers waarvan ik op verschillende vlakken veel geleerd heb, zo zijn

daar Steve Hora die mij geïntroduceerd heeft in de wereld van het eliciteren van experts en mij inspiratie gaf voor het ontwikkelen van PREJUDICE, Mike McKay wiens filosofie in hoe om te gaan met wiskundige problemen mij zeer heeft geïnspireerd en Fred Harper, die als amerikaans projectleider, samen met Louis op een zeer professionele, objectieve manier het project leidde.

Binnen de academische wereld is daar Tim Bedford. Tim Bedford heeft me op sleeptouw genomen in de woeste wateren van de maattheorie en persoonlijk vind ik het nog steeds spijtig dat hij naar Glasgow, Schotland vertrokken is.

Andere solo-zeezeilers: Sten de Wit, Jan van Noortwijk, Benjamin Jansen, Etienne de Klerk en Erling Andersen. Veel koersinformatie hebben wij in de loop der tijd uitgewisseld. Vaak heb ik gekeken watvoor koers zij vaarden. Vaak moest ik tegenover hun mijn koers verantwoorden. Sten en Jan hebben mij vooral bijgestaan in het ontwikkelen van PREJUDICE. Door hun inzichten is PREJUDICE geëvolueerd tot een wiskundig gereedschap dat van praktische waarde is. Benjamin, Etienne en Erling zijn vooral de mensen geweest die mij wegwijs maakte in verschillende optimaliserings-technieken/programmatuur.

Frank Härte en Nicole van Elst hebben als verkenners gefungeerd voor Hoofdstuk 3. De verkenning die zijn gedaan hebben en het rapport wat daarna geschreven is was zodanig dat het de moeite waard was om die richting/koers verder te verkennen.

En natuurlijk is daar het secretariaat, het logistiek centrum met Cindy Bosman, Diana Droog en Netty Zuidervaart. Geen verzoek is teveel, staan altijd klaar en voeren een goed georganiseerde administratie die het werken met ze zeer plezierig en makkelijk maakt.

Praten over welke koers te varen is belangrijk, maar net zo belangrijk is dat je je gedachten even op iets anders kan richten. Door afstand te nemen kom je vaak tot betere inzichten. Vrienden zijn hierin heel erg belangrijk. Daarom dank ik speciaal mijn vrienden Fred, Menno, Gert-Jan en Michiel. Vriendschap wat zolang terug gaat (in sommige gevallen meer dan 30 jaar) is schaars goed en iets waar ik zeer zuinig op ben. De weekenden, activiteiten etc. die altijd georganiseerd worden waren prima gelegenheden om de aandacht op andere dingen te richten.

Behalve nederlandse vrienden ben ik op mijn reis ook Amerikaanse vrienden tegengekomen. Allereerst zijn daar Andy en Christine Reiter wiens gastvrijheid, flexibiliteit en vriendschap mij zeer dierbaar is, Pete and Lisa Easton hebben een nieuwe dimensie gecreëerd in mijn fietsplezier door VeloClassic Tours op te richten (www.veloclassic.com).

Dan zijn er nog twee vrienden, c.q. zeer prominente bakens/vuurtorens, die ik zeker niet mag en wil vergeten. Het zijn Ronald 'CooZ' Cozijn en Bas Meyberg. Met CooZ en Bas heb ik veel gepraat over welke koers te nemen, en niet alleen qua promotieonderzoek. Bas pakte het op een gegeven moment zeer gestructureerd aan met het DIRECT project plan. De vele

ontmoetingen, lunches etc. hebben mij enorm gesteund.

Verder wil ik Annette Moelker en Ronald Cozijn hartelijk bedanken voor het prachtige ontwerp van de omslag.

Mijn broer Pieter, zus Marisa, zwager Ibbel en nichtjes Nadine en Merle wil ik bedanken voor hun steun en het steeds maar weer herinneren aan het feit dat ik wel mijn proefschrift moet afmaken.

En als laatste zijn daar de mensen die mij 'opgeleid' hebben: Pa en Ma. Mijn GPS systeem. Zij stonden in de haven bij mijn vertrek en zijn er ook weer op het moment dat ik de haven binnenvaar. Alhoewel zij niet de gehele reis mij fysiek hebben vergezeld, heb ik altijd hun steun en toewijding gevoeld. Dat is iets wat niet meetbaar is, maar wat ongelooflijk belangrijk voor mij is geweest en zeker een, zonet de, beslissende factor is geweest in het afronden van dit proefschrift.

Curriculum Vitae

Bernd Kraan werd geboren op 13 juli 1967 in Den Helder. In 1979 begon hij op de MAVO-Wieringen in Hippolytushoef, waar hij in 1983 zijn MAVO diploma behaalde. Vervolgens werd aan de scholengemeenschap Wiringherlant in Wieringerwerf het HAVO en Atheneum diploma in, respectievelijk, 1985 en 1987 behaald. Aansluitend studeerde hij Technische Wiskunde aan de Technische Universiteit Delft (TUD). Het ingenieursdiploma werd in 1994 behaald. In 1994 trad hij in dienst bij de TUD als Assistent In Opleiding bij de faculteit Technische Wiskunde en Informatica. Van 1994-1999 deed hij zijn promotieonderzoek en was lid van het projectteam van de Joint CEC/USNRC Uncertainty Analysis project. In 1999 trad hij in dienst bij de TUD als toegevoegd onderzoeker. In mei 2000 richtte hij het adviesbureau RandUM op, wat zich specialiseert in het optimaal beslissen onder onzekerheid.