Eduard Hofer

The Uncertainty Analysis of Model Results

A Practical Guide

Springer

The Uncertainty Analysis of Model Results

Eduard Hofer

The Uncertainty Analysis of Model Results

A Practical Guide

Springer

Eduard Hofer
Dorfen, Germany

ISBN 978-3-319-76296-8 ISBN 978-3-319-76297-5 (eBook)
https://doi.org/10.1007/978-3-319-76297-5

Library of Congress Control Number: 2018934848

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Computer models are everywhere in use, and the uncritical acceptance of results from computer models is also found everywhere. The developers of computer models are generally aware of the weaknesses and limitations of their models and frequently also of the influence of various individual uncertainties on model results. However, the model users and those who have to base their decisions on model results are unfortunately often less well informed. For those who are informed, it is often impossible to adequately express this information by grey areas around the model results obtained or presented to them. Consequently, model results are frequently taken at face value and are interpreted up to a detail that would only be justified in the case of precision measurements. This may provoke embarrassing situations such as the following:

Dr. A is in a conference where he is to compare two design options for a technical system with respect to their compliance with a safety limit. The audience is made up of fellow researchers, representatives of a funding organization, government departments and prospective users of the system. One of his PowerPoint slides shows the computed evolution of a critical temperature over the planned system operation time. Above this time history runs a straight line indicating the safety limit that must not be exceeded. The computed temperature increases from the ambient value at time zero to the maximum value of 532 °C at 3865 s, which is 78 °C below the limit, and then decreases to 320 °C where it settles for the rest of the operation time. Everybody present is satisfied with this result. However, a Mister B, seated in the third row, wishes to ask a question.

B: How did you arrive at this temperature curve?
A: We ran the computer model XYZ with our case data. It took several hours on our computer, but, as you can see, the result is worth the wait.
B: Are you sure that XYZ models all relevant phenomena?
A: We take it from the model description that all phenomena relevant to our case are in the model.

B: Is it imaginable that those that are not in the model could drive the critical temperature above the safety limit?
A: This would have to be some weird combination of very improbable phenomena.
B: What do you mean by "improbable" in this context?
A: That there is very little degree of belief among experts that a combination of phenomena, such as the one you are thinking of, could be of relevance in our case.
B: How much degree of belief?
A: Difficult to say.
B: How about the modelling assumptions and their parameter values used by XYZ? Could they have been chosen differently?
A: Wherever we saw a choice, we used the best estimate assumptions and values.
B: That means, other assumptions and values could have also been possibly true, but you trusted yours most. How much less do you trust the others? Might some combinations of those have driven the maximum temperature above the safety limit?
A: May be, but again, those would be very unlikely combinations of assumptions and parameter values.
B: What you mean by "unlikely" in this context?
A: The same as for phenomena.
B: Thank you, and I am sure you have an impression of the influence of measurement errors in your case-specific input data.
A: The situation is much clearer here, as these values are based on our wide experience in the field and on measurements with very good information as to their error distribution.
B: Bias of your field data can be excluded, I assume.

In the meantime, Doctor C, the head of Doctor A's department, has lost patience with Mister B and wishes to close the discussion.

C: Why are we having this discussion? We are below the safety limit by almost 100 °C. This suggests to me that, if taken into account, none of the possibilities discussed here at length would have the potential of driving the computed temperature above the limit.
A: It would be very unlikely.
C: I thought it could be excluded. What do you mean by unlikely? Either we are below the limit or we are not. Why are we using terms like "improbable" and "unlikely"?
B: This is also my opinion. Either the maximum temperature is above or below the limit. Unfortunately, Doctor A cannot answer this question in that clarity. He has shown us the result from a computer model that does comply with the limit but he cannot exclude that using other model assumptions and different combinations of parameter values could drive the computed maximum temperature above the limit, nor can he tell us how confident he is that such a combination is not true.

C: Doctor A, what would you need to bring this discussion to an end? Where would you need to know more and how confident could you then be that the limit will not be exceeded?
A: I would have to perform an uncertainty analysis.
C: Well, then why are we wasting our time with this unnecessary discussion, go ahead with it and let's talk about your analysis results once you have them.
B: Doctor A, I believe you also applied the computer model to calculate the maximum temperature for another design of the system. How do those results compare to the limit?
A: The computed maximum temperature is below the limit by 100 °C.
B: Well, then this is clearly the better design option.
A: Unfortunately, one cannot say "clearly". For this design we have had to choose values for parameters that are less well known than those for the first option.
C: But the maximum temperature is significantly further below the safety limit than in the case of the first design.
A: This is the case in the calculation that uses our best estimate parameter values.
C: What else could you have used?
A: There is quite a range of values for some of these parameters. Any value from each of these ranges is possibly true. Using these in combination would certainly give different results. I can't say by how much but I assume that we would still be well below the limit.
C: This is all very fine, but which design option is now the better one?
A: To answer this question I would need the results from an uncertainty analysis of both computations.
C: Just go ahead and do it. After all this discussion I do not understand why we have bothered to look at the results you showed us today. They are nice but as we have learned today there is a range of possibly true maximum temperatures. In other words, there is more to it than just looking at these point results before we can make a meaningful decision.

Surely, you don't want to find yourself in Doctor A's predicament. The subsequent chapters will guide you through the steps of an uncertainty analysis. If Doctor A had performed such an analysis before going public with his model results, he could have answered the questions of Mister B in a professional manner. More importantly, he would have gained insight into the combined effect of the various uncertainties on the model result and thus could have provided information essential for a robust decision.

It is common knowledge that models

- Try to adequately represent functional relationships
- Introduce simplifications for reasons of computational efficiency
- Employ parameter values that could have been chosen differently
- Use data that are at best subject to measurement error and at worst only vaguely known
- Apply numerical algorithms to obtain approximate solutions

Prediction models additionally need to process data that will only be known some time in the future.

All these ingredients of the model application have in common that they introduce uncertainties due to lack of knowledge, also known as "epistemic uncertainties". The uncertainty analysis is carried out to assess their combined influence on the model result. To this end, the state of knowledge is quantified at the level of scenarios, phenomena, model formulations, parameters, input data and controls of numerical solution algorithms.

It would be naïve to ignore the epistemic uncertainties and to expect precise results. It is equally naïve to be satisfied when presented with numbers that look like "precise" results. Too often are model results discussed as if they were the truth. Whoever bases his decision on model results that are not supplemented by quantitative statements about their uncertainty runs the risk of making the wrong decision.

The model result obtained with so-called best estimates of all uncertainties may comply with a safety standard, while quantifying the states of knowledge may show a non-negligible subjective probability for violation.

On the other hand, the model result obtained with so-called conservative assumptions may violate the standard, while quantifying the state of knowledge may show compliance with the safety standard at a high subjective probability. Not to mention the fact that what appears to be "conservative" for one result may not be so for another result of the same model application.

Frequently, the decision is about the preference for one design or action out of a number of alternative designs or actions. The choice between the alternatives may be rather straightforward if based on model results obtained with "best estimates" for the uncertainties. The model application for each of the alternatives may, however, involve different uncertainties of differing state of knowledge and/or may react differently to the same uncertainties. It can be of vital importance to know the effect of these differences.

Of course, quantifying the combined influence of all identified uncertainties is one aspect of the analysis. The other would be to rank the uncertainties with respect to their contribution to model result uncertainty. Many model developers and model users are particularly interested in this by-product of the uncertainty analysis. The uncertainty importance information is in the form of a ranking of the epistemic uncertainties. It provides guidance as to where to invest in better measurements, additional field data collection, supplementary experiments, further theoretical investigations and elicitation of expert judgement, in order to reduce the uncertainty of the model result most effectively and most cost-efficiently.

Of course, it is always possible to perform so-called one-at-a-time variations or scenario studies where

- A single parameter is varied by some amount or
- A single model is varied from its standard formulation or
- A single phenomenon of uncertain relevance is included or excluded or
- The scenario description for the model application is changed.

This approach is still popular among some model developers although it does not give an impression of the combined influence of the uncertainties. The effort involved is often comparable to that of an uncertainty analysis, while the information gained is far less conclusive. It is simply a matter of efficiency to use well-established methods and tools from statistics in order to determine a range that contains the requested percentage of the model result uncertainty at a sufficiently high level of confidence: in other words, to obtain as much information as is achievable with the invested budget.

Half a century ago it was tradition among scientists and engineers to perform "error analysis". The "errors" were mostly due to small fluctuations around nominal values and to rounding in moderately complex functional relationships. Today, there are large uncertainties in complex model formulations that require tens of thousands of lines of coding and in numerous parameter values and input data. To conduct an error analysis, or more appropriately an uncertainty analysis, is obviously mandatory.

To summarize:

Models are applied to obtain input for decisions. Uncertainty analysis shows:

- Whether the pending decision can be meaningfully made at the present state of knowledge.
- Where the state of knowledge should be primarily improved in order to allow meaningful decision-making.
- Whether there is a chance to achieve the necessary improvement.
- The subjective probability at which a specific action or design, out of a set of alternatives, is to be preferred to the rest of the set. This information is not available without analysis but decides about success or failure.
- The subjective probability for compliance with a safety standard, i.e. limit value. If this probability is judged to be too low, action is required. Either improving the state of knowledge may do or a change to the system under investigation is needed or both. Without the analysis there is no indication as to whether and where a system change and/or state of knowledge improvement might be most effective.

In conclusion, the uncertainty analysis is performed in order to arrive at decisions that are less often regretted than would be the case without analysis.

This book aims to serve as your practical guide through the six steps of an uncertainty analysis.

Two types of uncertainty, namely epistemic and aleatoric uncertainty, are to be considered, and Chapter 1 explains their differences and gives practical examples of the common situation where only epistemic uncertainty is present and of the special situation where both types of uncertainty are involved and separation is required. Chapters 2 to 7 guide through the six analysis steps in situations where only epistemic uncertainty is present. Step 1 consists of a systematic walk through the computer model in search of the potentially important uncertainties. The task of Step 2 is the quantification of the respective states of knowledge and their expression by subjective probabilities. Step 3 propagates the states of knowledge through the

computer model application. In Step 4 one arrives at the state of knowledge that follows in a logically consistent manner for the model results given the state of knowledge quantifications at the level of the identified uncertainties. The evaluation of a set of uncertainty importance measures follows in Step 5. The measures serve to rank the uncertainties with respect to their contribution to the uncertainty of the model result. Finally, in Step 6, the analysis and its findings are presented. The chapter explains how the findings are to be interpreted and communicated in a manner that is scientifically sound and yet comprehensible for those who need the computer model result for their decision-making.

Chapter 8 looks at each of the six steps from the viewpoint of their practical execution and points out how they can be supported by dedicated software. The specifics of an uncertainty analysis, when separation of uncertainties is required by the question formulation, are considered in Chapter 9 for each of the six analysis steps. Finally, two practical examples are presented in Chapter 10. Separation is not required in the first example where the uncertainty is to be analysed for results that are functions of the independent variable "time". Uncertainty and uncertainty importance are both presented over time.

The model result of the second example is a probability distribution summarizing variability within a population of values. Of particular interest is the complementary cumulative probability at a given limit value. The combined effect of all quantified epistemic uncertainties on the computed complementary cumulative probability and the corresponding uncertainty importance information are determined and presented. This example requires the separation of the variability within the population of values from the epistemic uncertainty involved in their quantification.

Dorfen, Germany Eduard Hofer

Contents

Chapter 1
Introduction and Necessary Distinctions

1.1 The Application of Computer Models

Computer models are everywhere in use. Whether it is the manufacturing industry, the electricity generating utility, the airways, the railways, financial institutions, the oil and chemical industry, hospitals or the government, they all use the output from computer models as input to their decision-making.

At the base of the decisions are problems such as the:

- Prediction of economic growth.
- Assessment of health, technical and monetary risks.
- Reliable supply of electricity and drinking water.
- Reliable operation of waste water treatment plants.
- Maintenance of oil reservoirs for safe and optimum production.
- Preparation of response plans and intervention strategies for infectious diseases.
- Formulation of safety requirements for potentially hazardous industries.
- Choice of sustainable harvesting quotas for natural resources like forests and fish populations.
- Adequate appreciation of the human influence on global climate.
- Profitable management of investment portfolios.
- Optimum scheduling of aircrafts and aircrews.
- And so forth.

To solve these problems, a number of questions are to be asked. Computer models are applied to find the answers that are then used in decision-making. Applying a computer model is the only way of finding the answers since observations, tests or experiments are often out of the question. This is particularly so if the decisions are concerned with potentially hazardous events or with predictions far into the future.

A computer model is an encoded simplified mathematical abstraction of reality. Simplified, because only those influential measurable quantities are included that are

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_1

deemed important for the answer and so are only those interactions among them that are judged to influence the answer noticeably. Abstraction, because the simplified picture of reality is reduced to numbers describing a situation on a grid of a mesh size judged to be sufficiently fine, and to parameter studded mathematical equations operating on this description. An application of the computer model provides an approximation of how the situation evolves due to the modelled interactions and due to any modelled disturbances. The answers obtained from the output may be single numbers, sets of numbers, discretized functions of time or space or of both, sets of such functions, or a combination of some or all of these formats. Some questions will require answers in the form of probability distributions.

A sequence of modelling steps lead from the question to the answer. First comes the compilation and representation by numbers of all quantities needed to describe the situation and all relevant disturbances, i.e. the preparation of the scenario description. It follows the formulation of the conceptual model consisting of a compilation of the mechanisms, phenomena and processes driving the interactions between the components of the scenario description. This may be done in the form of an influence diagram. The next step is the representation of the conceptual model by a mathematical model consisting of sets of conditions, governing equations, constitutive relations and logic expressions. The values of its parameters have then to be chosen in accordance with the intended model application. Numerical algorithms are needed to represent the mathematical model by a sequence of arithmetic and logic operations that solve the equations. This numerical model contains controls with their parameter values often chosen by the user. It needs to be encoded into a computer program. Application of this program requires the choice of the model option best suited for the purpose. The computer program is translated into machine language and executed using the input data chosen for the application of the model option. Finally, from all the information contained in the model output, those parts need to be selected and properly interpreted that are suitable to answer the question, and the answer needs to be unambiguously formulated and communicated.

1.2 Sources of Epistemic Uncertainty

A responsible decision-maker will scrutinize the answer by asking the following questions:

- Scenario description
  Is the choice of scenario suitable for answering the question and is its description sufficiently accurate and comprehensive?
- Conceptual model
  Are all relevant phenomena, mechanisms and processes accounted for and are they considered in the required detail?

- Mathematical model
  Are the governing equations, constitutive relations and logic expressions the only proper representation of the conceptual model?
- Parameter values
  Are the parameter values uniquely defined or could they have been chosen differently?
- Numerical model
  Was the overall effect of errors that are due to the representation of continuous quantities on a finite mesh of grid points as well as the effect of errors that are due to iterative solution processes and rounding controlled and kept within prescribed limits? Which are these limits and are they chosen such as to guarantee a sufficiently accurate solution?
- Computer program
  What kind of precautions were taken to prevent and detect any programming errors?
- Model options
  The model developers have most likely provided for various options to run the model on a computer. Is the model option chosen by the user the intended one and is it appropriate for the purpose of the model application? Are the required input data provided correctly?
- Computing system
  What is the chance that any hardware, operating system or compiler error may have affected the model results?
- Model output
  Was the suitable model output chosen and interpreted correctly in order to properly answer the question?

Most likely he will learn that:

- Scenario description
  The situation is not uniquely defined and even if it were there would still be incomplete knowledge of the numerical values of the quantities needed in its description. Furthermore, the relevant disturbances could be different from the chosen ones with respect to type, number, location, as well as time of occurrence. The modellers may be of the opinion that they have chosen a set of disturbances that capture the main effects with regard to the question. They will not be able to exclude that they might be wrong, nor will they be able to claim that the chosen set is unique.
- Conceptual model
  Inevitably, not all phenomena, mechanisms and processes acting on the quantities of interest are included in the conceptual model. Some of them might have been judged as having negligible effect and were therefore excluded. What is relevant was determined by expert judgment. Other experts may hold a different opinion.

- Mathematical model
  There are different ways of formulating the governing equations depending on what the modellers want to emphasize and how generously they think they can allocate computer time to their model runs. Since constitutive relations between model variables are often based on observations from experiments they too provide room for discrepancies among modellers.
- Parameter values
  For almost every parameter, a value different to the one chosen could be considered as applicable. For some parameters, the range of such values is large and for others small, yet it is not known for certain what the variation (and particularly the simultaneous variation) of their values would do to the model result.
- Numerical model
  The numerical solution process is controlled by parameters. Their values are often default values prescribed by the model developers or are free for the model user to choose. These parameters are responsible for speed and accuracy of the solution process. Often a trade-off between the two is made. Obviously, a different set of parameter values will produce a different solution. It is, however, not quite clear by how much different.
- Programming
  There is no guarantee that the program is free of errors although parts or all of it may have undergone a more or less elaborate verification process.
- Model option
  Again, there is no guarantee that the model option chosen via input is the actually intended one nor that the intended one is actually appropriate for answering the question.
- Computing system
  The chance of errors to exist and to have an effect on the model results is usually considered as remote.
- Model output
  Other experts may have looked at a different set of output values and may have drawn somewhat different conclusions leading to different answers.

Obviously, the decision-maker will want to know the overall effect of these uncertainties on the model result that is provided as an answer to his question and ultimately as input to his decision. The uncertainties are called epistemic since their root cause is a lack of knowledge. The term "epistemic" is derived from the Greek word "episteme" which means knowledge. Epistemic uncertainties could be eliminated given sufficient knowledge. The epistemic uncertainties are also known as Type B uncertainties,[1] and the analysis of their combined effect on the model results is the subject of Chaps. 2–7.

[1]Not to be confused with "Type B uncertainty evaluation" in (Joint Committee for Guides in Metrology 2008).

1.3 Verification and Validation

Computer models providing input to important decisions should have undergone a verification and validation process in order to diminish some of the uncertainties (Weathers et al. 2009). Verification aims at ensuring that the encoded mathematical/numerical model, i.e. the computer program, is free of programming errors. Validation, on the other hand, aims at ensuring that the mathematical/numerical model, including the scenario description and a typical selection from the model options, does what it should do in order to provide the required answer to the question as input to the decision-making. Neither verification nor validation of large complex computer models can ever be complete. Among other techniques and strategies, the former often tests the program with artificial data selected from the assumed range of applications while the latter often works with the comparison of results from critical sub-models and of the complete model to experimental data from a so-called validation matrix of experiments. Depending on how important the intended application of the computer model is, it may even be decided to go the expensive route of having several teams of programmers encode the same model and compare the results in order to reduce the chance of programming errors. An equivalent on the validation side are international benchmarks. In these benchmarks, several institutions or nations finance an experimental facility and the execution of experiments that are then part of the validation matrix for each institution's own computer model development (own choice of scenario description, own conceptual and mathematical/numerical model).

These efforts greatly diminish the uncertainties mentioned above. They can, however, not eliminate them completely. There will always remain a range of alternative scenario descriptions, conceptual models, mathematical model formulations, parameter and data values, numerical solution processes and model options that cannot be rationally excluded from consideration for the actual practical model application that has to answer the specific question at hand.

1.4 Why Perform an Analysis of Epistemic Uncertainty?

The model result presented to the decision-maker is often only a so-called point result obtained with "best estimates" for all uncertainties. In view of the uncertainties involved, the decision-maker will need to specifically request that an analysis of their overall effect be done, i.e. he will want to know where within this overall effect the point result is located. Here are some good reasons why he should want such an analysis to be done:

– Comparison of alternative designs or actions
  Frequently, the decision is about the preference for one design or action out of a number of alternative designs or actions (Sanchez et al. 2009; Koch et al. 2009). For each alternative, the model application may be subject to different uncertainties

and/or may react differently to the same uncertainties. The state of knowledge of the latter may also be different for each alternative. It can be of vital importance to know the effect of these differences on the uncertainty ranges. The ranges may, for instance, hardly overlap or the width of one may be a fraction of that of the other, while the point results, obtained with so-called best estimates of the uncertainties, may be very similar. Consequently, a comparison of the point results, in order to decide about the best alternative, only makes sense together with the output from the uncertainty analysis. In this way, a measure is obtained on which to base one's preference in the face of uncertainty.

– Compliance with safety limits
  Many decisions are concerned with limit values in safety standards, guidelines or business practices. Permission to operate a potentially hazardous industrial facility or to follow a specific strategy in exploiting a natural resource will depend on compliance with prescribed limit values. Compliance is often demonstrated by comparing the result from a computer model application to the limit. While the point result may comply with the limit, a considerable part of the quantified uncertainty may indicate violation of the limit. Uncertainty analysis can provide the necessary insight to decide whether and where the state of knowledge should be improved or whether and where the design of the technical system or the strategy should be changed in order to achieve compliance in the most cost-effective manner.

In some instances, uncertainty analysis may indicate that the decision cannot be made at the present state of knowledge since the uncertainty of the model result is simply too large for it to be a meaningful input. Or, on a more positive note, one may learn that the point result, including its large uncertainty range, is still on the safe side of the prescribed limit value. In this case, the decision can be safely made despite large uncertainties. Decisions made on the basis of an uncertainty analysis of model results are robust and less frequently regretted than those based only on point results.

1.5 Source of Aleatoric Uncertainty

Up to this point, it was assumed that the decision-maker's question has a single true answer. However, some questions are such that several different answers are simultaneously true (Hoffman and Hammonds 1994; Hofer 1996; Aven 2010). For instance, the generic question "What is the financial damage caused by an oil spill?" has a variety of values as true answers. As long as no additional detail is provided, the question cannot be answered by a single value. However, the question "What is the financial damage of yesterday's oil spill in XY bay?" clearly has only one true value as its answer. This value is most likely not known right now but will be known once the clean-up operation is completed and all damage claims have been dealt with.

Not every detail necessary to narrow down the population of true answers to a unique answer is given in the formulation of generic questions. In many instances, it is simply not possible nor intended to provide all the necessary details. The population of answers corresponds to the many ways in which the question formulation could theoretically be made sufficiently specific so that a single answer is true. It is uncertain which of these specifications the answer should refer to and therefore it is uncertain which answer to choose from the population. The associated uncertainty is called aleatoric or Type A uncertainty.[2] The term "aleatoric" is derived from the Latin word for die. It should imply that any sufficient specification of the question and therefore the true answer arises by chance similar to the outcome of the throw of a die. Some authors (Glorennec 2006; Ragas et al. 2008; Warren-Hicks and Hart 2010) prefer not to speak of uncertainty but simply of variability that is to be determined and expressed by a summarizing quantitative expression.

As against epistemic uncertainties, aleatoric uncertainties cannot be eliminated by acquiring more knowledge. The population of true answers can only be reduced by increasing the detail of the question, i.e. by making the question more specific as in the oil spill example given above. The variability of the true answers in the population is summarized by a frequency distribution or, after normalization, by a probability distribution. This distribution is the answer to be presented to the decision-maker. It is a so-called probabilistic answer. In short, the probability distribution quantifying the aleatoric uncertainty (or summarizing the variability) is the answer produced by a probabilistic computer model, i.e. one that works with random variables. The analysis of the epistemic uncertainty of the result from this computer model produces a quantitative expression of the epistemic uncertainty of the probabilistic answer that then permits uncertainty statements for distribution moments like mean value and variance as well as for cumulative or complementary cumulative probabilities at any values of interest.

In some instances, the question formulation could include enough details so as to have a single true answer but some of the detail is deliberately left unspecified in order to learn about the corresponding population of answers and in particular about the distribution that summarizes it.

The decision-maker may be less interested in the full information about the distribution quantifying aleatoric uncertainty but may only want to know mean value and standard deviation or some specific cumulative or complementary cumulative probabilities. The analysis of the epistemic uncertainties provides state of knowledge expressions for all of these values.

[2]Not to be confused with "Type A uncertainty evaluation" in (Joint Committee for Guides in Metrology 2008).

1.6 Two Different Interpretations of "Probability"

The mathematical concept of probability is used to quantify uncertainty. A different interpretation of the term "probability" is being used for each of the two types of uncertainty mentioned above. There is the classical frequentistic interpretation (probability as the limit of relative frequencies obtained with increasing sample sizes) and the subjectivistic interpretation (probability as a measure of degree of belief). With both interpretations, the wealth of well-established concepts and tools of probability calculus and statistics are at the disposal of the uncertainty analyst.

The aleatoric uncertainty is quantified using the frequentistic interpretation where one simply speaks of "probability" while the subjectivistic interpretation, where one speaks of "subjective probability", is used for the epistemic uncertainty. Aleatoric uncertainty leads to a population of true answers that is summarized by a probability distribution. Since anything subject to epistemic uncertainty has only one true answer, limits of relative frequencies do not make sense. Rather degrees of belief are held for the validity of an answer. In the case of epistemic uncertainty, there are several answers that are considered as possibly true while there is only one true answer.

1.7 Separation of Uncertainties

Some examples are now in place to serve the illustration of the distinctions made above.

Example 1
There are two dice on the table. One, let us call it A, is being cast continuously. The other, let us call it B, is covered, left untouched and it is uncertain which side faces up. At any given time, the number shown by B and the number that will be shown by A are uncertain, as is their sum. For simplicity, denote these uncertain quantities by A, B and A + B. The uncertainty about the number shown by die A in any (unspecified) cast is aleatoric (since the cast is unspecified there is a population of six simultaneously true answers). It is quantified using the frequentistic interpretation where one simply speaks of "probability". The subjectivistic interpretation, where one speaks of "subjective probability", is used for B the uncertainty of which is epistemic. Since B is constant, i.e. has only one true value, limits of relative frequencies don't make sense. Rather, degrees of belief are held for each of the six numbers on the die to be up. They quantify the state of knowledge for B.

Thus, the question "What is the value of the sum A + B?" is incomplete and causes problems. Firstly, what does the sum refer to? Secondly, if "probability" is used for A and "subjective probability" for B, which is to be used for A + B? Tackling the first problem will also resolve the second. Since A is being cast continuously, the question is ambiguous and needs to be supplemented by a reference unit. Examples of such supplements are ". . .in the next cast of die A" or "in any cast of die A".

As for the *next cast*, does it make any difference as to whether die A is being cast now or has been cast in the past and covered so that one just needs to lift the dice box to find out? In both cases, there is only one true but unknown value for A. The uncertainties of A and B, and consequently of A + B, are, therefore, quantified by subjective probability. Separation of uncertainties is not required.

If the question refers to *any cast*, A does not have only one true value. Rather, a population of values applies such that for any cast, A can be thought of as randomly selected from this population. There is uncertainty as to which value from the population to use in the sum. The variability of the values, within this population, is summarized by their proportions. These proportions are given by the respective limits of relative frequencies. The corresponding probability distribution summarizes the population variability. There is only one true probability distribution of A. Uncertainty of this distribution (for instance, due to uncertainty of whether die A is fair) is quantified by subjective probability. What about the sum A + B in the case where the question refers to any cast? It too has a probability distribution. There is only one true but unknown value B and only one true probability distribution of A. Consequently, A + B has only one true but unknown probability distribution. The uncertainty about the single true distribution of A + B is quantified by subjective probabilities derived from those for alternative values B and possibly for alternative probability distributions of A. One has thus naturally arrived at a separation of uncertainties. Without separation one would say that the population variability of true answers ranges from 2 to 12, which is clearly wrong as the unknown value B is constant and only A varies.

This introductory example seeks to illustrate when and how to separate or when not to separate uncertainties. Does the underlying principle also hold in situations of practical relevance? An answer to this question will be given with the help of the next examples.

Example 2

The experiment X was performed by pouring molten metal into a water pool and by measuring the pressure in the experimental facility at various locations and over time. Subsequently, as part of the validation process, a calculation is performed using a computer model to answer the question: "What was the value of the overall peak pressure increase in the experiment X?" Numerous uncertainties are involved in this calculation.

The experiment X has a unique description and quantities needed for the calculation have single, true, but more or less imprecisely known values, i.e. they are all epistemic uncertainties. This also applies to the functional relationships to be modelled and to the minimum requirements that need to be asked from numerical algorithms in order to still have sufficiently accurate numerical solutions of the underlying equations. Consequently, all uncertainties are quantified by subjective probability, and the uncertainty of the computed peak pressure increase is expressed by the resulting subjective probability distribution. Following the argumentation of the "next cast" case of Example 1, the same applies to a pre-experiment calculation that may be done as a so-called blind validation test of the model.

In the post-experiment case, the overall peak pressure increase is known, subject to measurement error. The corresponding uncertainty is also epistemic and is expressed by a subjective probability distribution since the actual increase has only one true but unknown value. This distribution will differ from the one obtained for the computed peak pressure increase. The reason for this is the difference in the states of knowledge involved. A substantial overlap of both distributions would be one criterion for a desirable outcome of the validation experiment. Another would be an uncertainty range narrow enough to just contain the measured value and all or most of the measurement error uncertainty (Weathers et al. 2009).

The initial temperature of the melt is one of the quantities needed for the computer model application. One may ask: "What about the error in the measurement value of the initial temperature of the melt? Why is the associated uncertainty not to be separated?" Measurement errors typically exhibit stochastic variability that can be summarized by a probability distribution. The initial temperature is an uncertain input of the calculation. Its measurement error has a single, true, but unknown value. Therefore, there is no need for separation. If it can be thought of as randomly sampled according to a given distribution of measurement errors, the distribution details can be used as those of a subjective probability distribution quantifying the respective state of knowledge of the measurement error.

Example 3

A strategy was developed to reduce the potential for vessel failure under described conditions in an industrial facility X. The description is of limited detail.

The question "Will vessel failure be prevented by the strategy, whenever the described conditions apply in facility X?" is answered by a discrete probability distribution assigning probability p to "prevented" and $(1 - p)$ to "not prevented". The value of p quantifies the combined influence of quantities that do assume different values whenever the described conditions are met in facility X, i.e. the aleatoric uncertainties. Their variability under the described conditions is summarized by probability distributions and is responsible for the variability between "prevented" and "not prevented". The uncertainties due to limited knowledge of the distribution functions as well as uncertainty of those quantities that have a single, true, but unknown value, like the parameters of the distributions, are epistemic and lead to a subjective probability distribution for p. An example of an epistemic uncertainty would be the failure rate of a pump to run, if running the pump over a period of time is part of the strategy and the failure rate has the same value whenever the described conditions apply. On the other hand, the time preceding failure of the pump, if not included in the described conditions, would be an example of aleatoric uncertainty, which is to be separated, since this time varies whenever the described conditions apply.

If there are no quantities that do assume different values whenever the described conditions are met in facility X, i.e. no aleatoric uncertainties, then the strategy either always fails or never fails. The corresponding discrete subjective probability distribution tells, for instance, the degree of belief for "prevented", given the present state of knowledge. The situation of no variability is, as an approximation, sometimes

assumed when the aleatoric uncertainty is judged to be negligible compared to the combined influence of the epistemic uncertainties.

Example 4
A hazardous substance (contaminant) was released some time ago, accidentally or continuously over a period of time, and the question is: "What is the value of the dose received by an individual of the exposed population?" There is a single, true, but more or less imprecisely known, local and temporal pattern of contaminant concentrations in the air, on the ground, in water as well as in various foodstuffs. The same applies to the temporal pattern of locations, activities and food consumption habits as well as to the metabolic properties of each of the individuals in the exposed population. However, the question does not refer to a specific individual. Therefore, it has a population of answers that are simultaneously true. For instance, it is uncertain which individual's data to use from the latter pattern. This uncertainty is due to variability among the individuals in the population. The question can be interpreted such that individuals are being drawn at random from the population with equal probability. This makes the variability of the individual's data stochastic and the uncertainty of the aleatoric type. It is to be separated from the epistemic uncertainty due to lack of knowledge of the local and temporal pattern of contaminant concentration in the air, on the ground, in water as well as in all relevant foodstuffs, i.e. it needs to be separated from the uncertainty of all quantities that have single true but unknown values applicable to groups of individuals or to the exposed population in general. Since this latter set contains the local and temporal pattern of concentrations, it also contains all quantities needed to reconstruct this pattern, such as the characteristics of the contaminant release.

The answer to the question is a single, true, but unknown probability distribution summarizing the variability of dose among the individuals of the population. The quantification of the state of knowledge of both patterns mentioned above leads to a subjective probability distribution for any distribution quantile. Such quantiles or percentiles are dose values that are not exceeded by a given percentage of the exposed population. The subjective probability distribution of the percentile tells the decision-maker how well this dose value can be determined, given the present state of knowledge. The same kind of information is available for the percentage of the exposed population with dose values larger than a given limit. Both pieces of information, dose, respectively, percentage on one hand and the uncertainty thereof on the other hand, feed into different decisions.

The question could also be "What is the value of the dose received by each of the N individuals in a given population?" This question has a single true answer namely the true set of N dose values and all uncertainty is epistemic. The uncertainty analysis provides a population of possibly true sets of N dose values each. The variability of the dose values within each of these sets could be summarized by a frequency distribution that, if normalized, could serve as the probabilistic answer to the previous formulation of the question while the population of possibly true probabilistic answers quantitatively expresses the state of knowledge of the true probabilistic answer. In either case, variability, no matter whether deterministic as in

the case of a calculation for N specific individuals, or stochastic, as in the case of a calculation for a random sample of individuals, needs to be separated from epistemic uncertainty.

A further question could be "What is the value of the dose received by individual Z?" This question now refers to a single specific individual and all uncertainty is epistemic. All of the individual's details have single, true, but imprecisely known values. The answer to the question is a single, true, but unknown dose value, and the combined influence of all uncertainties leads to a subjective probability distribution as state of knowledge expression for this dose value.

Example 5
In order to avoid stock depletion of a certain fish species, a fishing management strategy is to be implemented. To provide input for the pertaining decision process, a computer model was developed and it is applied to answer, among others, the following question: "How will the harvestable biomass of the fish species evolve over the next 20 years if no fishing management policy is in place?" Later applications of the model would include alternative management strategies in order to make an optimal choice. The question formulated above is analogous to "the next cast" question of Example 1. It has only one true answer, namely the evolution of the harvestable biomass over the next 20 years, given no management strategy has been implemented. All uncertainties are epistemic. This includes also the uncertainty as to how often and when a certain unpredictable disturbance of the sea current will occur over the next 20 years and how severely it will affect the population of the fish species concerned. Since the number of occurrences, their time and severity are considered to be random variables, one may be inclined to separate them as aleatoric uncertainties. However, once the specified time period is over, it will be known how often, when and how severely the disturbance affected the fish species. Then the single true answer to the question will also be known. For now, however, at the onset of the specified time period, these data are epistemic uncertainties (see also the practical application in Sect. 10.1).

The next six chapters lead through the steps of an uncertainty analysis when separation of uncertainties is not required by the question formulation. The analysis proceeds in six steps. Specifics of an uncertainty analysis in cases requiring separation of uncertainties are discussed in Chap. 9. Chapter 10 presents one practical example each, namely for an analysis that does not require separation as well as for one that does require separation of uncertainties.

References

Aven, T. (2010). Some reflections on uncertainty analysis and management. *Reliability Engineering & System Safety, 95*, 195–201.

Dubbel, H. (Ed.). (1939). *Taschenbuch für den Maschinenbau*. Berlin: Springer.

Glorennec, P. (2006). Analysis and reduction of the uncertainty of the assessment of children's lead exposure around an old mine. *Environmental Research, 100*, 150–158.

Hofer, E. (1996). When to separate uncertainties and when not to separate. *Reliability Engineering and System Safety, 54*, 113–118.

Hoffman, F. O., & Hammonds, J. S. (1994). Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis, 14*(5), 707–712.

Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (1st ed.). JCGM 100 (GUM 1995 with minor corrections).

Koch, F. H., et al. (2009). Evaluating critical uncertainty thresholds in a spatial model of forest pest invasion risk. *Risk Analysis, 29*(9), 1227–1241.

Ragas, A. M. J., et al. (2008). Separation of uncertainty and inter-individual variability in human exposure modeling. *Journal of Exposure Science and Environmental Epidemiology, 19*, 201–212.

Sanchez, A., et al. (2009). Addressing imperfect maintenance modeling uncertainty in unavailability and cost based optimization. *Reliability Engineering and System Safety, 94*, 22–32.

Warren-Hicks, W. J., & Hart, A. (2010). *Application of uncertainty analysis to ecological risks of pesticides*. Boca Raton, FL: CRC Press.

Weathers, J. B., et al. (2009). An exercise in model validation: Comparing univariate statistics and Monte Carlo based multivariate statistics. *Reliability Engineering and System Safety, 94*, 1695–1702.

Chapter 2
STEP 1: Search

2.1 The Scenario Description

The search for potentially important epistemic uncertainties requires a thorough inspection of all stages on the path from the assessment question to the computer model result that is used to answer the question. The inspection starts with the scenario description where a number of simplifications and omissions will most likely have taken place. Their potential for changing the model result needs to be discussed and judged.

For instance, a model used to compute the amount of contaminant that might be transported from an underground waste repository into the upper aquifer tries to represent the main geological features in and around the repository in sufficient detail as part of its scenario description. Despite some information obtained from various boreholes, a large set of possibly true descriptions could be constructed of which a set of representatives needs to be selected and prepared for use in the uncertainty analysis of the computer model application. These representatives will also differ in the number of accidental intrusions into the repository after closure as well as in their times of occurrence and their locations.

A model, trying to answer the question of the point in time at which the population of a certain species of fish will be reduced to below a critical number due to commercial fishing activities, will need to use information from previous harvests and from survey expeditions. These numbers are subject to errors of reporting, detection, measurement and interpretation. Alternative possibly true numbers will need to be used in alternative scenario descriptions.

In the case of the post-experiment calculation, concerning molten metal being poured into a water pool, the amount and the temperature of the melt effectively crossing the water surface may be insufficiently known so as to make it necessary to consider using alternative pairs of values in the calculation.

It may be the case that the model is not flexible enough to accommodate some of the possibly true peculiarities of the scenario. If the model user is also the developer,

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_2

it will be possible to supplement the model accordingly. Otherwise, the possibly applicable features of the scenario description, which cannot be fed into the model, must be documented. They will need to be included in the discussion of the various shortcomings when the model results and their uncertainty analysis are presented to the decision-makers.

2.2 The Conceptual Model

Next one needs to look at the conceptual model. The intentional simplifications and omissions that might, individually or in combination with others, be of importance to the model results need to be scrutinized. The model developers would make excellent partners in this effort. Otherwise, the model user depends on the degree of detail offered by the model documentation. The simplifications and omissions are based on expert judgement with a view to the intended spectrum of model applications. The model user needs to judge whether his application is within this spectrum. He is also left to his own expertise when it comes to the identification of any unintentional simplifications and omissions of interactions between the quantities of the scenario description, phenomena, processes or mechanisms that could have an impact on the model results. Any simplifications and omissions, judged to be of importance, need to be rectified in the model. It may be that this is not possible due to time and cost restrictions or simply because the model user does not have access to the source code. In this case, they need to be documented and included in the list of shortcomings mentioned above for discussion together with the model results. Simplifications and omissions are, by their very nature, not directly subject of the uncertainty analysis as it is clear that their elimination would only improve the result. There is, however, epistemic uncertainty about whether or not the improvement is essential for the model result to be sufficiently accurate. Supplementing the model will most likely be reserved for the elimination of severe shortcomings as it is expensive in two ways. Firstly, there is the effort of modelling, programming and testing and secondly, the supplement may considerably increase the runtime of the model on the computer and may therefore hamper its usability. Model supplements will be outright impossible if the user does not have access to the source code of the model, and the modellers are not willing to provide the required supplement. Documenting the uncertainties that might be associated with a particular simplification or omission may therefore be the only realistic way to go.

The uncertainty as to whether component ageing needs to be included into a model computing the reliability of a technical system might serve as an example. Another example may be the effect that over-fishing the upper-age classes of a population of fish might have on the maturation time. Particularly, if this effect has been observed in another fish population. In the context of health effects of some contaminant in water, it may be uncertain whether one should consider a threshold value below which there is no effect if such a threshold has been observed in animals. A model used to study different strategies in the combat of a viral infection

may contain, as a simplification, the assumption that the medication administered is equally effective in all parts of the body or it may omit the transmission of the virus between body cells.

2.3 The Mathematical Model

The next stage to be inspected is the mathematical model. It is a translation of the conceptual model together with the scenario description into sets of equations and logic expressions. These prescribe how the descriptive variables evolve from initial values while observing any limiting or boundary conditions and disturbances. The question here is whether the equations represent the governing laws in a way that permits an adequate account of the effect of phenomena and mechanisms important for the model result and whether the influence of the disturbances on the evolution of variables is sufficiently accounted for. For instance, the model developers may have, for the sake of simplicity and of saving runtime on the computer, set up the equations under the assumption of symmetry in one dimension of the spatial scenario description while it is uncertain as to whether the effect of some phenomenon could have been better captured without this assumption. It may also be uncertain whether the equations represent the initial and boundary conditions in sufficient detail. For example, the mathematical model may divide a complex geometry into compartments and work only with values averaged over each compartment. The set of equations reflect the number and geometry of the compartments. It may be uncertain whether a refinement of the division at critical locations would change the model result significantly. For example, a model predicting the influence of harvesting practices on a population of fish may divide the population into classes of different age for setting up the equations. This division would need to be chosen such that the age-dependent specifics of spawning, commercial harvesting and loss to predators can be represented in sufficient detail. Again, there may be uncertainty as to whether a finer subdivision of the fish population may render different results.

In addition to equations representing governing laws as well as initial and boundary conditions, the model may contain equations that account for the influence of disturbances. These equations are often approximate relationships derived from experiments. Consequently, there may be different ways of expressing these relationships within the mathematical model.

After the inspection of the set of equations, the choice of parameter values for the model application needs to be questioned. The number of parameters may range from a few dozens to hundreds or even thousands. Parameters are constants of the mathematical model. An example would be a flow rate known to be set at a constant value, while the true flow rate may differ from the set value so that it is uncertain which value to use in the model. Another example would be the constant volume of a complex geometry. Its value would need to be calculated from various measurements that are subject to measurement error. Some parameters are constants of sub-models obtained by fitting the sub-model to observations from tests or other

sources. Different researchers may have performed their own experiments or may have used the same experiment to arrive at their own set of constants through their individual way of judging the quality of fit.

Some parameters are approximate representations of a function, the so-called effective parameter values. The function is, for simplicity, approximated by a constant over the range of application. There will most likely be a range of constant values that could be used, and it may be uncertain which value to use in the specific model application. An example would be the fecundity rate of a species of fish. There is a direct relationship between the size of the female and the number of eggs while there may be no distinction by size within the age classes used by the mathematical model. Further examples are transfer factors of a contaminant to human organs, consumption rates of foodstuffs, permeability within a geological layer, viscosities, specific heat capacities, flammability limits of materials, critical concentration values of dust for the modelling of dust explosions, friction coefficients, deposition velocities of airborne contaminants and so forth.

Some effective parameters may be given in form of a table containing parameter values dependent on variables such as model interim results or on independent variables like time and spatial coordinates of the model equations. The table contains a finite number of pairs of values. Determining the parameter value to be used will require some interpolation between the tabulated values. The tabulated parameter value at a given variable value may be uncertain and so will be the interpolation error and the effect of both on the model result. In fact, any application-specific input value that is to be used by the computer model will give rise to the question "how well is this value actually known and could the uncertainty have significant influence on the model result?" Some input values may be the output from the application of another computer model, a so-called feeder model that feeds into the computer model application to be analysed. Obviously, a quantification of their uncertainty is required so that one may take it into account within the analysis of the computer model application. It would be less of a problem to satisfy this requirement if an uncertainty analysis had been performed for the application of the feeder model.

2.4 The Numerical Model

Finally, one arrives at the numerical model which is the mathematical model turned into a numerical algorithm that provides an approximate solution of the set of equations. In the case of algebraic equations, the solution algorithm will often be iterative using one or more parameters to control the speed of convergence and the accuracy of the approximate solution. In the case of differential equations, approximations of the continuous solution variables will be obtained at discrete grid points of the independent variables. This process may also employ an iteration algorithm. The choice of the grid decides about accuracy and efficiency of the solution algorithm. For instance, in a time-advancement procedure for initial value problems in first-order ordinary differential equations, the accuracy of the advancement from

one time-step to the next will be controlled by parameters, the so-called local accuracy requirements that also control the length of the time-step. The local errors accumulate and it is necessary to know whether the accumulated error is sufficiently attenuated in each step so that the local error plus the accumulated error remain controlled. The numerical model will also contain parameters to control the handling of the solution across discontinuities and to adjust the grid of discrete solution points to the solution behaviour over time. The choice of values for the parameters of the numerical model has immediate effect on the solution accuracy as well as on the runtime of the model on a computer. Therefore, a compromise is often made between accuracy and runtime requirements. It may be uncertain whether other parameter values, leading to more stringent accuracy requirements, would need to be chosen in order to solve the equations of the mathematical model with sufficient accuracy.

2.5 Conclusion

The uncertainties encountered on the way from the scenario description to the numerical model may be categorized as either data or model uncertainties. The first category contains uncertain parameter values and uncertain input data while the second includes uncertainties like:

- How to represent a process by a function of model variables.
- How to incorporate the effect of a disturbance into the model equations.
- The effect of a simplification that is not an omission.
- The effect of omissions (intended or unintended), sometimes called “completeness uncertainty”.

The next chapter discusses ways of quantifying the state of knowledge for epistemic uncertainties of both categories by subjective probability.

Chapter 3
STEP 2: Quantify

3.1 Subjective Probability

Subjective probability is a measure of the degree of belief held for the truth of an answer to a question. It is used in the quantification of uncertainty due to lack of knowledge, also called epistemic uncertainty. The word epistemic stems from the Greek word for knowledge. It indicates that this uncertainty has its origin in the nature and limits of knowledge. In the case of epistemic uncertainty, there are several answers to a question that are considered as possibly true while there is only one true answer. The true answer will either be deterministic or probabilistic, depending on the formulation of the question, i.e. the absence or presence of aleatoric uncertainty. Aleatoric uncertainty is due to random variability so that there are many true answers. Their variability is summarized by a probability distribution (Table 3.1). Both types of answers, i.e. the one true answer or the one true probability distribution, may be subject to epistemic uncertainty. What is known about an epistemic uncertainty (i.e. the respective state of knowledge) is probabilistically expressed by a subjective probability distribution. The state of knowledge may be judged differently by different experts and even by the same experts after some time has gone by and additional information has become available.

Subjective probability serves as a mathematical measure of the state of knowledge. Every now and then attempts are made to apply other concepts. One example is "possibility" from Fuzzy Logic (Ross 1995). Several good reasons speak, however, for the use of the mathematical measure "probability" in its subjectivistic interpretation. It is the input expected for decision-making under uncertainty. Furthermore, all well-established methods and tools from probability calculus and statistics, which are already in use for the quantification of the aleatoric uncertainties by probability in its frequentistic interpretation, can just as well be employed for the treatment of epistemic uncertainties by probability in its subjectivistic interpretation. Among those is the well-known Bayesian method (Bayes 1958; Box and Tiao 1973).

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_3

Table 3.1 Difference between the classical frequentistic and the subjectivistic interpretation of "probability"

Subjectivistic interpretation	Classical frequentistic interpretation
Probability is the degree of belief (held for the truth of an answer to a question)	Probability is the limit of relative frequencies (of a random event)
Expresses the state of knowledge	Summarizes random variability
Quantifies uncertainty due to lack of knowledge (epistemic uncertainty)	Quantifies uncertainty due to random variability (aleatoric uncertainty)

It is used to update once specified subjective probability distributions by observations that subsequently became available.

Subjective probabilities (subsequently denoted by "sw") have to comply with the same rules as probabilities in their frequentistic interpretation. The main rules are:

R1 Subjective probabilities cannot be negative nor can they be larger than 1.

R2 If A_1 is a possibly true answer to question A and B_1 is a possibly true answer to question B, then the subjective probability for both to be true is $\mathrm{sw}(A_1B_1) = \mathrm{sw}(A_1)\mathrm{sw}(B_1|A_1) = \mathrm{sw}(B_1)\mathrm{sw}(\mathrm{A}_1|B_1)$ where the vertical stroke is to be read as "under the condition that what is behind the stroke is true". If the subjective probability for B_1 to be true is the same irrespective of whether A_1 is true or not, then $\mathrm{sw}(B_1|A_1) = \mathrm{sw}(B_1)$ and the subjective probability for both to be true is $\mathrm{sw}(A_1B_1) = \mathrm{sw}(A_1)\mathrm{sw}(B_1)$.

R3 If A_1 is a possibly true answer to question A and B_1 is a possibly true answer to question B, then the subjective probability for at least one of them to be true is $\mathrm{sw}(A_1+\mathrm{B}_1) = \mathrm{sw}(A_1) + \mathrm{sw}(B_1) - \mathrm{sw}(A_1B_1)$.[1] If A_1 and B_1 cannot be true together, then the subjective probability for at least one of them to be true is $\mathrm{sw}(A_1 + B_1) = \mathrm{sw}(A_1) + \mathrm{sw}(B_1)$. Particularly, if $\mathrm{sw}(A_1)$ is the subjective probability for answer A_1 to be true and sw(not A_1) for A_1 to be false, then $\mathrm{sw}(A_1) + \mathrm{sw}(\text{not } A_1) = 1$.

R4 If $A_1, A_2, ..., A_n$ are possibly true answers to question A, not any two or more of them can be true together since there is only one true answer. It follows: $\mathrm{sw}(A_1 + A_2 + ... + A_n) = \mathrm{sw}(A_1) + \mathrm{sw}(A_2) + ... + \mathrm{sw}(A_n)$, for any integer value n. Particularly, if these are the only answers that are possibly true, then $\mathrm{sw}(A_1 + A_2 + ... + A_n) = 1$.

3.2 Data Versus Model Uncertainty

The previous analysis step produced a compilation of epistemic uncertainties that may need to be considered in the uncertainty analysis of results from the application of a computer model. They may be categorized into data and model (or modelling) uncertainties. This differentiation is important for the state of knowledge quantification as can be seen from the following comparison:

[1] Since $\mathrm{sw}(A_1B_1)$ is part of $\mathrm{sw}(A_1)$ and of $\mathrm{sw}(B_1)$, it needs to be subtracted.

Data uncertainty:

- A datum is a constant of a computer model. It may have different numerical values for different applications of the computer model.
- The datum is uncertain if there is lack of knowledge about the true numerical value to be used in the specific application.
- Minimum and maximum of the possibly true numerical values determine, in general, the set of alternatives (completeness in the sense of probability calculus).
- Only one numerical value from the set can be true (mutual exclusiveness in the sense of probability calculus).
- The alternatives are a set with unique total ordering in the sense of set theory (Lipschutz 1964). This property will be shown to be important for the derivation of uncertainty importance measures.

Model uncertainty:

- A model is a computer representation of a process or of a state of affairs.
- The representation is an uncertainty of the computer model application if there is lack of knowledge about the process or the state of affairs or if the representation is only achieved in an approximate or simplified manner. It may even be that the process or state of affairs is totally omitted from the computer model so that there is lack of knowledge about the effect this may have on the result of the model application.
- Frequently, there is not a "minimal" and "maximal" possibly true representation.
- Usually, there is no unique total ordering of the representations.
- Generally, it will be impossible to specify all possibly true representations (no completeness).
- In practice, representations will be taken into consideration that are known to be simplifications or approximations. It will be obvious that none of the representations is actually true. The representations in the set are assumed to be only sufficiently adequate. Several representations may be considered as equally adequate (no exclusiveness like in the case of possibly true representations).

3.3 Ways to Quantify Data Uncertainty

The following categories of data are distinguished:

- Measurable quantities
- The result of functions of measurable quantities
- Values chosen at random according to distributions fitted to measurable quantities
- Sequences of uncertain input data over time, space or interim computer model results.
- Special cases:

- Correction factors to results of models fitted to measurable quantities
- Effective parameter values
- Extrapolations of models fitted to measurable quantities
- Feeder model results
- Literature data
- Interview data

3.3.1 Measurable Quantities as Uncertain Data

Measurable quantities are all those quantities that can, sometimes only in principle, be the subject of a measurement (Joint Committee for Guides in Metrology 2008) states:

- In 3.2.1:
  "In general, a measurement has imperfections that give rise to an error in the measurement result. Traditionally, an error is viewed as having two components, namely, a random component and a systematic component".
- In 3.2.2:
  "Random error presumably arises from unpredictable or stochastic temporal and spatial variations of influence quantities. The effects of such variations, hereafter termed random effects, give rise to variations in repeated observations of the measurand. Although it is not possible to compensate for the random error of a measurement result, it can usually be reduced by increasing the number of observations; its expectation or expected value is zero".
- In 3.2.3:
  "Systematic error, like random error, cannot be eliminated but it too can often be reduced. If a systematic error arises from a recognized effect of an influence quantity on a measurement result, hereafter termed a systematic effect, the effect can be quantified and, if it is significant in size relative to the required accuracy of the measurement, a correction or correction factor can be applied to compensate for the effect. It is assumed that, after correction, the expectation or expected value of the error arising from a systematic effect is zero".
- In 3.3.1:
  "The uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurand. The result of a measurement after correction for recognized systematic effects is still only an estimate of the value of the measurand because of the uncertainty arising from random effects and from imperfect correction of the result for systematic effects".
- In 3.3.2:
  "In practice, there are many possible sources of uncertainty in a measurement, including:

  a) Incomplete definition of the measurand.
  b) Imperfect realization of the definition of the measurand.

c) Non-representative sampling—the sample measured may not represent the defined measurand.
d) Inadequate knowledge of the effects of environmental conditions on the measurement or imperfect measurement of environmental conditions.
e) Personal bias in reading analogue instruments.
f) Finite instrument resolution or discrimination threshold.
g) Inexact values of measurement standards and reference materials.
h) Inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm.
i) Approximations and assumptions incorporated in the measurement method and procedure.
j) Variations in repeated observations of the measurand under apparently identical conditions.

– These sources are not necessarily independent, and some of sources a) to i) may contribute to source j). Of course, an unrecognized systematic effect cannot be taken into account in the evaluation of the uncertainty of the result of a measurement but contributes to its error".
– In 3.1.2:
  "In general, the result of a measurement is only an approximation or an estimate of the value of the measurand and thus is complete only when accompanied by a statement of the uncertainty of that estimate".
– In D.5.2:
  "Uncertainty of measurement is thus an expression of the fact that, for a given measurand and a given result of measurement of it, there is not one value but an infinite number of values dispersed about the result that are consistent with all of the observations and data and one's knowledge of the physical world, and that with varying degrees of credibility can be attributed to the measurand".
– End of quotations from (Joint Committee for Guides in Metrology 2008)

Uncertainty quantification, therefore, has the purpose of providing intervals about the measurement result that encompass, according to a given coverage probability or confidence level, a large fraction of the distribution of values that could reasonably be attributed to the quantity subject to measurement.

The following notation will be used:

z value of the measurand (uncertain datum)
x measurement (result) value (approximation for z)
z^* value that could be attributed to the measurand, for short also called "possibly true" (within the definition given for the measurand) value
x_i value observed in the i-th repetition of the measurement
$\varepsilon_{S,i}$ systematic error in x_i
$\varepsilon_{R,i}$ random error in x_i

Two types of measurement error models are commonly encountered in practice (Armstrong 1998):

In the classical model, there are different measurement results x associated with the same measurand value z.
In the Berkson model (Berkson 1950), there are different measurand values z associated with the same measurement result x.

Examples where the Berkson model applies:

- A specific drug is administered to a number of subjects participating in a medical experiment. The amount to be administered to each subject is set at x. The actual amounts z, however, vary about the value x.
- A computer model uses values of ground contamination at various locations of the site of an abandoned chemical factory. The site has been divided into L subareas and all that is available for use in the model is an estimate of the average contamination values x_l, $l = 1, \ldots, L$. The model requires, however, the actual values $z_{l,\,k}$ at $k = 1, \ldots, K$ locations within each subarea. They vary about the corresponding average values.

3.3.1.1 Uncertainty Due to Additive Classical Measurement Error

The measurement error may be made up of two terms:

A constant additive bias term (systematic error) ε_S (Vasquez and Whiting 2006).
A random additive term (random error) ε_R.

In the case of an additive classical measurement error, the following relationship applies:

$$x = z + \varepsilon = z + \varepsilon_R + \varepsilon_S$$

where x is the measurement result or measurement value, z is the true value of the measurand and ε is the error. ε is made up of the random error ε_R and often includes an unknown constant bias term ε_S.

The following situations are distinguished:

- The measurement values from N repetitions of the measurement of the same measurand are available.
- Only one measurement value is available for each of K measurands.

A procedure is now described, for each of these situations, that permits state of knowledge quantifications for the true value(s) of the measurand(s). If the error terms are multiplicative, the following procedures are applied to the logarithms of true values and measurement values, provided all values are positive.

a) N measurement values for the same measurand

The N measurement values x_n, $n = 1, \ldots, N$, for the uncertain datum z, are subject to the same unknown value ε_S of the bias term and to random error terms $\varepsilon_{R,\,n}$ so that

$$x_n = z + \varepsilon_S + \varepsilon_{R,n}, n = 1, \ldots, N \tag{3.1}$$

The values of the N random error terms are assumed to be a random sample drawn according to the same probability distribution (usually normal) of the random error[2] E_R with mean value 0 and unknown standard deviation σ_R.

A best estimate of z, if $\varepsilon_S = 0$, is obtained from

$$\bar{x} = \frac{1}{N}\sum\nolimits_{n=1}^{N} x_n = \frac{1}{N}\left(Nz + \sum\nolimits_{n=1}^{N} \varepsilon_{R,n}\right) = z + \frac{1}{N}\sum\nolimits_{n=1}^{N} \varepsilon_{R,n} \tag{3.2}$$

where the last term is approximately equal to 0. It is called an unbiased estimate since $\bar{x}$ approaches z with N increasing. The term

$$s_R^2 = \frac{1}{N-1}\sum_{n=1}^{N}\left(x_n - \bar{x}\right)^2 = \frac{1}{N-1}\sum_{n=1}^{N}\left[\varepsilon_{R,n} - \frac{1}{N}\sum_{n=1}^{N}\varepsilon_{R,n}\right]^2 \tag{3.3}$$

is an unbiased estimate of the variance σ_R of the random error E_R.

If one talks of any N repetitions of the measurement of z then their results are aleatoric uncertainties (see Chap. 1). Each of the N measurement values X_n, $n = 1$, …, N is a random variable because of the error terms $E_{R,\,n}$. Since each random variable $E_{R,\,n}$ may be thought of as a sum of independently distributed contributions, the random variables $X_n = z + \varepsilon_S + \mathrm{E}_{R,\,n}$ can be thought of as following a normal distribution with mean value $z + \varepsilon_S$ and standard deviation σ_R, according to the central limit theorem. The distribution of $\bar{X} = \left(\frac{1}{N}\right)\sum_{n=1}^{N} X_n$ is then normal with mean value $z + \varepsilon_S$ and standard deviation $\sigma_R/N^{1/2}$. Furthermore,

$$S_R^2 = \frac{1}{N-1}\sum_{n=1}^{N}\left(X_n - \bar{X}\right)^2$$

is a random variable and s_R^2 (Eq. 3.3) is a realization thereof. Following Cochran's Theorem, the term

$$(N-1)S_R^2/\sigma_R^2 = \sum_{n=1}^{N}\left[\frac{X_n - \bar{X}}{\sigma_R}\right]^2 \tag{3.4}$$

is Chi^2-distributed with $N-1$ degrees of freedom (Heinhold and Gaede 1968). From this one obtains that, with

$$Y = \frac{[\bar{X} - (z + \varepsilon_S)]}{\frac{\sigma_R}{N^{\frac{1}{2}}}} \tag{3.5}$$

following a standard normal distribution, the quotient

[2]Random variables are denoted by capital letters and their realizations by lower case letters.

$$\frac{Y}{\frac{S_R}{\sigma_R}} = \frac{\left[\bar{X} - z - \varepsilon_S\right]}{\frac{S_R}{N^{\frac{1}{2}}} = W} \tag{3.6}$$

follows a Student or t-distribution with $N-1$ degrees of freedom.

If one now talks of N specific repetitions of the measurement, providing N measurement values $x_1, \ldots, x_N$, then

$$w = \left[\bar{x} - z - \varepsilon_S\right] / \left(s_R / N^{1/2}\right)$$

is a realization of the Student distributed random variable with $N - 1$ degrees of freedom. The value of the realization w is, however, unknown due to lack of knowledge (epistemic uncertainties) of z and of ε_S. Since w is sampled at random, the Student or t-distribution can be used to quantify the state of knowledge of w. The t-distribution is symmetric about zero; it follows, therefore, that

$$\begin{aligned} \mathrm{sw}\left(w \leq \mathrm{t}_{(\mathrm{N}-1),\mathrm{Q}\%}\right) &= \mathrm{Q}/100 \\ \mathrm{sw}\left(w > \mathrm{t}_{(\mathrm{N}-1),(100-\mathrm{Q})\%}\right) &= \mathrm{Q}/100 \\ \mathrm{sw}\left(-w \leq -\mathrm{t}_{(\mathrm{N}-1),(100-\mathrm{Q})\%}\right) &= \mathrm{Q}/100 \end{aligned}$$

and

$$\begin{aligned} sw\left(z + \varepsilon_S \leq \bar{x} - t_{(N-1),(100-Q\%)}\left(\frac{s_R}{N^{\frac{1}{2}}}\right)\right) &= sw\left(z + \varepsilon_S \leq \bar{x} + t_{(N-1),(Q\%)}\left(\frac{s_R}{N^{\frac{1}{2}}}\right)\right) \\ &= Q/100. \end{aligned} \tag{3.7}$$

Given a subjective probability distribution that quantifies the state of knowledge of ε_S, a sample of possibly true values z^* for z is obtained as

$$z_J^* = \bar{x} - \varepsilon_{S,j}^* + t_{(N-1),j} s_R N^{-1/2}, j = 1, \ldots, J \tag{3.8}$$

with $t_{(N\,-\,1),\,j}$ sampled according to the t-distribution with $N-1$ degrees of freedom and $\varepsilon_{S,j}^*$ sampled according to the subjective probability distribution quantifying the state of knowledge of the bias term. s_R is obtained from Eq. (3.3) and $\bar{x}$ according to Eq. (3.2).

b) One measurement each for *K* measurands

Computer models often use a set of values z_k, $k = 1, \ldots, K$ sampled from a population with variance σ_Z^2 and mean value μ_Z. Their sample mean and variance are denoted by m_Z and s_Z^2. Available are, however, not the z values but one measurement value each, i.e. x_k, $k = 1, \ldots, K$. Suppose, the measurement values are subject to the unknown systematic error ε_S and to a random error $\varepsilon_{R,k}$ independently sampled according to the same normal distribution for $k = 1, \ldots, K$ with mean

value 0 and unknown standard deviation σ_R. If the error terms are additive then, with m_R and s_R^2 denoting the sample mean and variance of the random error, sample mean and variance of the measurement values $x_k = z_k + \varepsilon_S + \varepsilon_{R,\,k} = z_k + \varepsilon_k$ are[3]

$$m_x = \left(\frac{1}{K}\right)\sum_{k=1}^{K} x_k = \left(\frac{1}{K}\right)\sum_{k=1}^{K}(z_k + \varepsilon_k) = m_z + m_\varepsilon \tag{3.9}$$

$$\begin{aligned} s_x^2 &= \left(\frac{1}{K}\right)\sum_{k=1}^{K} x_k^2 - \left(\left(\frac{1}{K}\right)\sum_{k=1}^{K} x_k\right)^2 = \left(\frac{1}{K}\right)\sum_{k=1}^{K}(z_k+\varepsilon_k)^2 - \left(\left(\frac{1}{K}\right)\sum_{k=1}^{K}(z_k+\varepsilon_k)\right)^2 \\ &= \left(\frac{1}{K}\right)\sum_{k=1}^{K} z_k^2 - \left(\left(\frac{1}{K}\right)\sum_{k=1}^{K} z_k\right)^2 + \left(\frac{1}{K}\right)\sum_{k=1}^{K}(\varepsilon_k)^2 - \left(\left(\frac{1}{K}\right)\sum_{k=1}^{K}(\varepsilon_k)\right)^2 \\ &+ \left(\frac{2}{K}\right)\sum_{k=1}^{K} z_k\varepsilon_k - \left(\frac{2}{K^2}\right)\sum_{k=1}^{K} z_k \sum_{k=1}^{K}\varepsilon_k = s_z^2 + s_\varepsilon^2 + 2r(\boldsymbol{z},\boldsymbol{\varepsilon})s_z s_\varepsilon \end{aligned} \tag{3.10}$$

m_z, s_z^2 and m_ε, s_ε^2 are the sample mean and variance of the K true values and errors, r $(\boldsymbol{z},\boldsymbol{\varepsilon})$ is the sample correlation coefficient of the pairs of values (z_k, ε_k), $k = 1, \ldots, K$ while $\mathbf{z}' = (z_1, \ldots, z_K)$ and $\boldsymbol{\varepsilon}' = (\varepsilon_1, \ldots, \varepsilon_K)$ are the vectors of the true values and errors. Since the populations of Z and E are statistically independent, the population correlation coefficient satisfies $\rho(Z,E) = 0$. The sample correlation coefficient $r(\boldsymbol{z},\boldsymbol{\varepsilon})$ will most likely differ from zero. The smaller the number K of measurement values, the more likely it is that this difference will be significant. As can be seen from Eqs. (3.9) and (3.10), the sample mean and variance of the measurement values differ from those of the true values.

Suppose, the distribution of E is known to have mean value ε_S and standard deviation σ_R. The state of knowledge for each of the errors ε_k may then be quantified, independently for every measurement value x_k, by a subjective probability distribution of the same type with mean value ε_S and standard deviation σ_R. If values a_k, $k = 1, \ldots, K$ are sampled independently according to this subjective probability distribution and are subtracted from the measurement values, one arrives at a set of values

$$\{y_k = x_k - a_k | k = 1, \ldots, K\}. \tag{3.11}$$

Sample mean and variance are

[3] Vectors of random variables are denoted by bold capital letters and their realizations by bold lower case letters.

$$m_y = \left(\frac{1}{K}\right)\sum_{k=1}^{K} y_k = \left(\frac{1}{K}\right)\sum_{k=1}^{K} z_k + \left(\frac{1}{K}\right)\sum_{k=1}^{K} (\varepsilon_k - a_k) = m_z + (m_\varepsilon - m_a) \tag{3.12}$$

$$s_y^2 = s_x^2 + s_a^2 - 2r(\boldsymbol{x},\boldsymbol{a})s_x s_a = s_z^2 + s_\varepsilon^2 + 2r(\boldsymbol{z},\boldsymbol{\varepsilon})s_z s_\varepsilon + s_a^2 - 2r(\boldsymbol{x},\boldsymbol{a})s_x s_a \tag{3.13}$$

where s_x^2, s_a^2 and s_z^2 are the sample variances of the measurement values, of the sampled values that are considered to be possibly true errors and of the true values, and r($\boldsymbol{x}$,$\boldsymbol{a}$) is the sample correlation coefficient of the pairs of values (x_k, a_k), $k = 1$, …, K. As Eq. (3.13) shows, the sample variance of the set of values $\{y_k \mid k = 1, \ldots, K\}$ clearly differs from s_z^2. The set in (3.11) can therefore not be considered a possibly true set of input values suitable for a Monte Carlo simulation run of the uncertainty analysis.

It seems reasonable to require that the values y_k, $k = 1$, …, K should have the same sample mean and variance as the true values z_k in order to qualify as a set of input values for the uncertainty analysis. For the variance, this can be achieved by introducing statistical dependence between the measured values and the a_k, $k = 1$, …, K such that r($\boldsymbol{x}$,$\boldsymbol{a}$) satisfies the requirement

$$s_z^2 = s_z^2 + s_\varepsilon^2 + 2r(\boldsymbol{z},\boldsymbol{\varepsilon})s_z s_\varepsilon + s_a^2 - 2r(\boldsymbol{x},\boldsymbol{a})s_x s_a \tag{3.14}$$

or

$$r(\boldsymbol{x},\boldsymbol{a}) = \left(s_\varepsilon^2 + s_a^2 + 2r(\boldsymbol{z},\boldsymbol{\varepsilon})s_z s_\varepsilon\right)/2s_x s_a. \tag{3.15}$$

Unfortunately, there are some terms on the right-hand side of Eq. (3.15) that are subject to uncertainty. This makes the value required for r($\boldsymbol{x}$,$\boldsymbol{a}$) an uncertain quantity. The sample correlation coefficient r($\boldsymbol{z}$,$\boldsymbol{\varepsilon}$) is unknown and so are the sample standard deviations s_z and s_ε. The sample correlation coefficient is a random variable distributed as follows (Rosner 1995):

$$R(\boldsymbol{z},\boldsymbol{\varepsilon}|K) \sim T_{K-2}/\left(K - 2 + T_{K-2}^2\right)^{1/2} \tag{3.16}$$

where T_{K-2} is Student distributed with $(K - 2)$ degrees of freedom. Consequently, the state of knowledge of the sample correlation coefficient r($\boldsymbol{z}$,$\boldsymbol{\varepsilon}$) in Eq. (3.15) can be quantified by a subjective probability distribution as indicated in Eq. (3.16) where the population distributions of Z and E are assumed to be of the normal type. It is, however, also a useful state of knowledge expression for r($\boldsymbol{z}$,$\boldsymbol{\varepsilon}$) in situations where the distributions of Z and E are continuous but other than normal. If the sample size K satisfies $K \geq 30$, then the standard normal distribution may be used to approximate the Student distribution.

The state of knowledge quantifications for the sample mean m_ε and for the variance s_ε^2 of the true errors follow from the subjective probability distribution for the error values ε_k mentioned above. From Eq. (3.12), we have for the sample mean m_z of the true values

$$m_z = m_y - (m_\varepsilon - m_a). \tag{3.17}$$

Using Eq. (3.10) and assuming $s_x^2 > s_\varepsilon^2$, one obtains, by solving the quadratic equation,

$$s_z = -r(\boldsymbol{z},\boldsymbol{\varepsilon})s_\varepsilon + \left(r^2(\boldsymbol{z},\boldsymbol{\varepsilon})s_\varepsilon^2 + s_x^2 - s_\varepsilon^2\right)^{1/2} \tag{3.18}$$

where only the positive square root renders a solution. From Eq. (3.15), it therefore follows for the desired sample correlation coefficient r($\boldsymbol{x}$,$\boldsymbol{a}$)

$$r(\boldsymbol{x},\boldsymbol{a}) = \frac{s_\varepsilon^2 + s_a^2 + 2r(\boldsymbol{z},\boldsymbol{\varepsilon})\left(-r(\boldsymbol{z},\boldsymbol{\varepsilon})s_\varepsilon + \left(s_\varepsilon^2(r^2(\boldsymbol{z},\boldsymbol{\varepsilon}) - 1) + s_x^2\right)^{\frac{1}{2}}\right)s_\varepsilon}{2s_x s_a}. \tag{3.19}$$

Since there are subjective probability distributions expressing the state of knowledge for each of the imprecisely known terms m_ε, s_ε and r($\boldsymbol{z}$,$\boldsymbol{\varepsilon}$) on the right-hand side of Eqs. (3.17)–(3.19), a subjective probability distribution results for m_z, s_z and for $r(\boldsymbol{x},\boldsymbol{a})$.

In an uncertainty analysis, using Monte Carlo simulation for the propagation of the states of knowledge through the computer model, values a_j, $j = 1, \ldots, K$ are first sampled according to the subjective probability distribution for the errors ε_k. From this sample, the values m_a and s_a are obtained. They are considered as possibly true values for m_ε and s_ε as they are indirectly sampled according to the subjective probability distributions for m_ε and s_ε. Then a value r*($\boldsymbol{z}$,$\boldsymbol{\varepsilon}$) is sampled for r($\boldsymbol{z}$,$\boldsymbol{\varepsilon}$) according to Eq. (3.16) and together with the given value for s_x, a value r*($\boldsymbol{x}$,$\boldsymbol{a}$), indirectly sampled according to the subjective probability distribution for $r(\boldsymbol{x},\boldsymbol{a})$ in Eq. (3.19), is obtained. From these values, and using Eqs. (3.17) and (3.19), the corresponding values m_z^* and s_z^* are also obtained, i.e. indirectly sampled according to the subjective probability distributions for the sample mean m_z and standard deviation s_z of the true values. It is guaranteed that the value r*($\boldsymbol{x}$,$\boldsymbol{a}$) satisfies the necessary condition $-1 \le$ r*($\boldsymbol{x}$,$\boldsymbol{a}$) $\le +1$. The proof is given in (Hofer 2008). For the uncertainty analysis, Eqs. (3.17)–(3.19) therefore take the following form:

$$m_z^* = m_y - (m_a - m_a) = m_y \tag{3.17a}$$

$$s_z^* = -r^*(\boldsymbol{z},\boldsymbol{\varepsilon})s_a + \left(s_a^2\left(r^{*2}(\boldsymbol{z},\boldsymbol{\varepsilon}) - 1\right) + s_x^2\right)^{\frac{1}{2}} \tag{3.18a}$$

$$r^*(\boldsymbol{x},\boldsymbol{a}) = \frac{s_a + r^*(\boldsymbol{z},\boldsymbol{\varepsilon})\left(-r^*(\boldsymbol{z},\boldsymbol{\varepsilon})s_a + \left(s_a^2\left(r^{*2}(\boldsymbol{z},\boldsymbol{\varepsilon}) - 1\right) + s_x^2\right)^{\frac{1}{2}}\right)}{s_x}. \tag{3.19a}$$

A set of input values $\{z_k^* \mid k = 1, \ldots, K\}$, which can be considered as possibly true, is, therefore, obtained by combining the sampled errors a_j, $j = 1, \ldots, K$ with the measurement values x_k, $k = 1, \ldots, K$ to pairs $(x_k, a_{j(k)})$ and by computing

$$z_k^* = x_k - a_{j(k)}, k = 1, \ldots, K. \tag{3.20}$$

Here, the components of the vector $\boldsymbol{a}' = (a_1, \ldots, a_K)$ are permuted such as to have a vector $\tilde{\boldsymbol{a}}\,' = (a_{j(1)},\ldots, a_{j(K)})$ with the sample correlation coefficient $r(\boldsymbol{x}, \tilde{\boldsymbol{a}})$ equal to the value $r^*(\boldsymbol{x},\boldsymbol{a})$ obtained from Eq. (3.19a). The sample standard deviation s_{Z*} of the sample values z_k^* equals then the value s_Z^* from Eq. (3.18a) which is indirectly sampled according to the subjective probability distribution for s_Z. The index $j(k)$ stands for the component index j of $\boldsymbol{a}$ that was combined with the component index k of $\boldsymbol{x}$. To find such a combination, one may transform the x_k and the a_j to standard normal values and apply the method described in (Iman and Conover 1982). This is done in an iterative fashion until the sample correlation coefficient $r(\boldsymbol{x},\tilde{\boldsymbol{a}})$ is close enough to $r^*(\boldsymbol{x},\boldsymbol{a})$.

The method can, therefore, be summarized as follows:

Step 1: Sample a set of values $\{a_j | \, j = 1, \ldots, K\}$. They are considered as possibly true error values since they are sampled according to the subjective probability distribution that quantifies the state of knowledge for the errors ε_k, $k = 1, \ldots, K$.

Step 2: Compute the value s_a. It is considered a possibly true value for s_ε since it is indirectly sampled according to the subjective probability distribution for s_ε.

Step 3: Sample a value $r^*(\boldsymbol{z},\boldsymbol{\varepsilon})$ for $r(\boldsymbol{z},\boldsymbol{\varepsilon})$ according to the subjective probability distribution indicated in expression (3.16).

Step 4: Compute the value s_Z^* from Eq. (3.18a) using the results from Steps 2 and 3.

Step 5: Compute the value $r^*(\boldsymbol{x},\boldsymbol{a})$ from Eq. (3.19a).

Step 6: Find the index permutation $\{j(k) \mid k = 1, \ldots, K\}$, using the value $r^*(\boldsymbol{x},\boldsymbol{a})$ and applying the method in (Iman and Conover 1982) iteratively.

Return to Step 1 if convergence of $r(\boldsymbol{x},\tilde{a})$ to $r^*(\boldsymbol{x},\boldsymbol{a})$ cannot be achieved for the prescribed accuracy requirement and within the prescribed maximum number of iteration cycles.

Step 7: Compute the set of input values $\{z_k^* = x_k - a_{j(k)} \mid k = 1, \ldots, K\}$ to be used in a Monte Carlo simulation run of the uncertainty analysis of the model application.

Depending on their type of distribution, not every value between -1 and $+1$ may be possible for the correlation coefficient of two random variables. The algorithm will return to Step 1 should the value $r^*(\boldsymbol{x},\boldsymbol{a})$, computed in Step 5 of the algorithm, not be achievable within a preset number of iteration cycles. Thereby, it effects a truncation of the subjective probability distribution for $r^*(\boldsymbol{x},\boldsymbol{a})$, if necessary.

If N is the sample size of the Monte Carlo simulation of the uncertainty analysis, then these seven steps are repeated N-times giving N sets of input values $\{z_k^* \mid k = 1, \ldots, K\}$, each being considered as possibly true. The sample mean and standard deviation of each set are indirectly sampled according to the subjective probability distributions that quantify the state of knowledge for the sample mean m_z and standard deviation s_z of the true values, and the values $a_{j(k)}$ subtracted from the

measured values x_k are sampled according to the subjective probability distribution that quantifies the state of knowledge of the true errors ε_k.

3.3.1.2 Uncertainty Due to Multiplicative Classical Measurement Error

In the case of a multiplicative classical measurement error, the measurement value x is the product of the true value z and of the error factor $(1 + \varepsilon)$:

$$x = z(1 + \varepsilon) = z(1 + \varepsilon_R + \varepsilon_S).$$

Let $\{x_k | \, k = 1, \ldots, K\}$ be again the set of measurement values and let $\{\varepsilon_k | \, k = 1, \ldots, K\}$ be the set of error factors in the equation $x_k = z_k(1 + \varepsilon_k)$. It is reasonable to require that the sampling procedure produces a set $\{z_k^* | \, k = 1, \ldots, K\}$ with the same geometric mean and standard deviation as the true values. If X, Z and E can only assume positive values, then the equations derived for the additive classical measurement error can be applied to the natural logarithms and $\ln(x_k) = \ln(z_k) + \ln(1 + \varepsilon_k)$. The values

$$\ln\left(z_k^*\right) = \ln\left(x_k\right) - \ln\left(a_{j(k)}\right), k = 1, \ldots, K$$

are then a set of natural logarithms that can be considered as possibly true, with the sample mean m_{lnz*} and variance s_{lnz*}^2 indirectly sampled according to the subjective probability distributions for m_{lnz} and s_{lnz}^2. This approach provides input values

$$z_k^* = x_k / a_{j(k)}, k = 1, \ldots, K$$

with $\exp(m_{lnz*})$ and $\exp(s_{lnz*})$ sampled according to the subjective probability distribution for the geometric mean $\exp(m_{lnz})$ and for the standard deviation $\exp(s_{lnz})$ of the true values.

3.3.1.3 Uncertainty Due to Additive Berkson Measurement Error

In instances where the value x, taken from:

- A construction plan
- An experimental design
- A medical prescription
- A technical specification
- A compilation of averages
- etc.,

is considered to be the measurement value, the corresponding true value z will be scattered around x. The corresponding error is of the Berkson type and

$$z = x + \varepsilon_S + \varepsilon_R. \tag{3.21}$$

The notation is the same as for the classical error above. Again, subjective probability distributions are needed to quantify the state of knowledge of ε_S (if $\varepsilon_S \neq 0$ cannot be excluded) and of ε_R. These distributions may be based on expert judgment. The random error variable E_R will most likely follow a normal distribution with mean value 0 and standard deviation σ_R.

A possibly true value for z is then obtained as

$$z^* = x + \varepsilon_S^* + \upsilon\sigma_R^* \tag{3.22}$$

where ε_S^* is sampled according to the subjective probability distribution, σ_R^* is the standard deviation of the subjective probability distribution specified for E_R by expert judgment and v is sampled according to the standard normal distribution.

A set of J possibly true values $z_j^*, j = 1, \ldots, J$ for uncertainty analysis by Monte Carlo simulation may be obtained as

$$z_j^* = x + \varepsilon_{S,j}^* + v_j\sigma_R^*, j = 1, \ldots, J$$

where x is the given average or prescribed or specified value.

3.3.1.4 Uncertainty Due to Multiplicative Berkson Measurement Error

In the case of a multiplicative Berkson measurement error, the true value z is the product of the measurement value x and of the error factor $(1 + \varepsilon)$:

$$z = x(1 + \varepsilon) = x(1 + \varepsilon_R + \varepsilon_S).$$

As in the case of the additive Berkson error, subjective probability distributions are needed to quantify the state of knowledge of ε_S (if $\varepsilon_S \neq 0$ cannot be excluded) and of ε_R. These distributions may be based on expert judgment. The random error variable E_R will most likely follow a normal distribution with mean value 0 and standard deviation σ_R.

A possibly true value for z is then obtained as

$$z^* = x\left(1 + \varepsilon_S^* + v\sigma_R^*\right)$$

where ε_S^* is sampled according to the subjective probability distribution, σ_R^* is the standard deviation of the subjective probability distribution specified for E_R by expert judgment and v is sampled according to the standard normal distribution.

A set of J possibly true values $z_j^*, j = 1, \ldots, J$ for uncertainty analysis by Monte Carlo simulation may be obtained as

$$z_j^* = x\left(1 + \varepsilon_{S,j}^* + v_j\sigma_R^*\right), j = 1, \ldots, J$$

where x is the given average or prescribed or specified value.

Modelling this type of error for a set of values z_k, $k = 1, \ldots, K$ does not lead to the variance inflation observed in subsection 1 for the classical error type. Reason is that x is not a random variable that could depend on E_R.

3.3.2 Results of Functions of Measurable Quantities

Computer models often use input that is the value of a function of one or more quantities that are in principle measurable (for short "measurables"). The function values are either provided by the model user or are obtained from so-called feeder models that feed input into the computer model application. For instance, if $y = \mathrm{f}(x_1, \ldots, x_K)$ is obtained from K measurables, their uncertainty is propagated through the function f thereby leading to the uncertainty of y. The state of knowledge of the measurables is expressed by subjective probability distributions each (as discussed above and in Sect. 3.6) and a subjective probability distribution for y follows in a mathematically consistent way. The method of variance propagation works for simple functions f. Even then it is straightforward only if there is no state of knowledge dependence (see Sect. 3.5) between the x_k, $k = 1, \ldots, K$. In what follows, it is, therefore, assumed that the uncertainty propagation through the function f is done via Monte Carlo simulation (see Chap. 4) whereby state of knowledge dependence may be treated conveniently to arrive at J possibly true values for y

$$y_j^* = \mathrm{f}\left(x_{1,j}^*, \ldots, x_{K,j}^*\right), \quad \mathrm{j} = 1, \ldots, J$$

where $\{x_{k,j}^* \mid k = 1, \ldots, K\}$ is a possibly true set of values of the K measurables. If some or all of the x_k enter L functions $\mathrm{f}_1, \ldots, \mathrm{f}_\mathrm{L}$, then there is state of knowledge dependence among the results y_l, $l = 1, \ldots, L$ which is automatically accounted for in the Monte Carlo simulation since the $y_{l,\,j}$ are computed using the same set of values $\{x_{k,j}^* \mid k = 1, \ldots, K\}$.

3.3.3 Distributions Fitted to Measurable Quantities

All that is known about an uncertain datum may be the fact that it was chosen at random according to a probability distribution that summarizes the variability within a population. Consequently, its state of knowledge is expressed by using the population distribution as subjective probability distribution. Often, very little is known about this distribution. Some values drawn at random according to the population distribution may be known and, once a decision is made about the distribution type, may be used in a Bayesian analysis (see Sect. 9.2) to derive a joint subjective probability distribution for the distribution parameters. The state of knowledge for the uncertain datum may now be expressed by a random sample of

size J of possibly true values by first choosing a set of parameter values for the population distribution according to their joint subjective probability distribution and then choosing a possibly true value of the uncertain datum according to the population distribution with the set of parameter values just obtained and by repeating this process J times.

If the computer model application uses sets of values from the population (as uncertain data), it must be made sure that all values in the set are obtained using the same set of parameter values for the population distribution.

3.3.4 Sequences of Uncertain Input Data

The uncertain input data, given as a discrete function of an interim model result or of independent variables like time and/or space, may be state of knowledge independent. They may, however, also be state of knowledge dependent. This may, for instance, be due to some integral constraint that has to be satisfied by their set of values over all points of time and/or space or over all interim result values. In the first case, each input datum could be treated as a separate data uncertainty with the state of knowledge obtained as discussed in Sects. 3.6.1 and 3.6.3. In the case of dependence, however, the data may be treated simultaneously for all values of the independent variable as one "single model" uncertainty, using the most appropriate option from among the possibilities of state of knowledge elicitation considered in Sect. 3.6.2.

In some instances, this "single model" is the output from another computer model (feeder model) that provides input for the computer model application. In this situation, it will be appropriate to expect that the application of the feeder model has been subjected to an uncertainty analysis by Monte Carlo simulation providing a random sample of "single models". This sample is then an adequate expression of the state of knowledge. The random sample may be considered as a set of alternative model formulations of equal subjective probability and may be used as discussed in Sect. 3.4.1.

3.3.5 Special Cases

– Adjustment constant

A model was fitted to measurables under certain (i.e. experimental) conditions. It may be intended to use this model over the same range of values of the independent variables but under different conditions. The assumption is that a suitably chosen adjustment constant (multiplicative and/or additive) to the model output is capable of accounting for the change in conditions. There is uncertainty about the value of the constant that would best achieve this goal. The corresponding state of knowledge

will be based on expert judgment and needs to be expressed by a subjective probability distribution. To incorporate this state of knowledge into the state of knowledge for the value of the dependent variable obtained in the fitting process, one would multiply each value obtained for the dependent variable by a value for the adjustment constant chosen at random according to the specified subjective probability distribution. If the computer model application makes multiple uses of results of the fitted model under identical conditions, it must be made sure that the same value chosen for the adjustment constant is applied.

– Effective parameter value

This parameter value is a constant serving as an approximation to a dependent variable over a suitably small domain in the space spanned by the independent variables. Usually, it is supposed to be the average value of the dependent variable over this domain. In principle, this average could be obtained from sufficiently many measurements or calculations (if there were a model) at random points within the domain, and its state of knowledge could be quantified as shown in Sect. 3.3.1. In practice, however, there are often none or too few measurements. Lacking a model that could provide the value of the dependent variable at any point within the considered domain, the state of knowledge for the effective parameter value will then be largely based on expert judgment.

– Uncertainty factor

If a model, fitted to measurement values, is to be used beyond the range of measured values of the independent variables used in the fitting process then there is additional uncertainty about the validity of the model result. This additional uncertainty could be accounted for by an uncertainty factor. The state of knowledge for this factor would most likely be based on expert judgment. Use of the factor is just as for the adjustment constant mentioned above. If the computer model application makes multiple uses of results of the fitted model over the same range, it will be necessary to make sure that the same value chosen for the uncertainty factor is applied.

– Feeder model results as uncertain data

Input to the computer model may come from an external model (see Sect. 3.3.2). Uncertainty analysis of its output by Monte Carlo simulation provides sets of model output values. The uncertainty analysis of the results of the computer model application may sample at random from these sets or use them one to one if the number of model runs of both Monte Carlo simulations is the same. If the computer model application makes multiple calls to the feeder model (e.g. for various interim results), then it will be necessary to run the uncertainty analysis of the feeder model simultaneously with that of the computer model application. This way it is made sure that the feeder model results from all calls are obtained with the same values of the feeder model uncertainties.

– Literature data

Literature data are often just single values with no indication as to how well this datum is known. In this case, expert judgment would be needed to arrive at a subjective probability distribution as state of knowledge expression.

In some instances, the datum is published together with its underlying database and with the numerical procedure of its calculation. In this case, it may be possible to arrive at a state of knowledge quantification by one of the procedures described in this subchapter.

Some literature data are reported together with either a $(+/-k\sigma)$ range ($k = 1, 2$ or 3) or a symmetrical $u\%$ subjective probability range. Expert judgment will be required to decide about the type of subjective probability distribution, and the uncertainty information given may be used to derive the values of the distribution parameters.

In the ideal case, not only the best estimate value of the datum but also a subjective probability distribution quantifying its state of knowledge is given in the literature.

– Interview data

It may be assumed that the interviewed individuals have sufficient expertise with respect to the required data. The interview should then follow the elicitation process outlined in Sect. 3.6.1. Data concerning events far in the past may be prone to large recall errors. Their state of knowledge quantifications would have to account for this. Experts familiar with this error phenomenon may have to participate in the interviews.

3.4 Ways to Quantify Model Uncertainty

Because of the differences between data and model (or modelling) uncertainties (see Sect. 3.2), the influence of data uncertainties is often analyzed in a methodologically closed form while the influence of model uncertainties is only investigated roughly by what is often called a sensitivity study. In the simplest case, this study uses the best estimate values for the uncertain data and runs the computer model for separately selected variations of each uncertain model. On the other extreme, a complete data uncertainty analysis is performed for each of a set of selected model variations. Both ways of treating model uncertainty do not permit a comprehensive uncertainty statement that includes the combined influence of data and model uncertainties. Both approaches will often be impossible to realize in practice considering the many model uncertainties of computer models that are of practical relevance. This section shows ways of how to account for all data and all model uncertainties simultaneously and thus to permit a comprehensive uncertainty quantification of the result of the computer model application. Contrary to data uncertainties, model uncertainties are a very inhomogeneous group. Consequently, the selection of an approach

for state of knowledge quantification requires some preparatory investigations. It will be necessary to ask questions like:

- Are phenomenological uncertainties involved?
- Is there a model for the representation of the phenomenon, process or state of affairs within the computer model? Or is the state of knowledge sufficiently advanced to develop a model? Or is the representation still subject to scientific controversy and there are only work hypotheses?
- Do several models exist in addition to the one already implemented in the computer model?
- Did the model(s) undergo a validation process, specifically over the relevant domain of the intended application?
- Are there any significant differences between the models with regard to model structure, discretization strategies, database used for model development and validation process?

The choice from the approaches mentioned below will be guided by the answers to these questions. In what follows, the various possibly adequate computer representations of a process or state of affairs are called model alternatives or alternative model formulations.

3.4.1 Sets of Alternative Model Formulations

There may exist a set $\{r_i \mid i = 1, \ldots, I\}$ of different computer representations of the process or state of affairs in question. The representation implemented in the computer model is an element of this set. It will most likely not be possible to exclude all but one as inadequate for the purpose of the computer model application. Consequently, subjective probabilities need to be assigned to each element in the set. For this task, there is an ideal and a general case.

In the ideal case, the I elements in the set satisfy the following criteria:

- Each element is possibly the true representation.
- There are no other representations that are possibly true (completeness).
- If the element with index i is the true representation, then none of the other elements can be true (mutual exclusiveness).

With $\mathrm{sw}(r_i)$ the subjective probability (degree of belief) for element i to be the true representation, it follows

$$\sum_{i=1}^{I} \mathrm{sw}(r_i) = 1.$$

In the general case, however, the following applies:

- Each element of the set is possibly an adequate representation, i.e. it possibly comes sufficiently close to the truth. It is known, however, that none is actually true.
- There are infinitely many more representations that could be claimed to be possibly adequate (no completeness). However, the I elements of the set cover their population well enough in the sense that they produce a spectrum of model results that can be considered as sufficiently close to the spectrum of model results obtained by the full population.
- If the element with index i is an adequate representation, it cannot be excluded that other elements of the set are also adequate (no mutual exclusiveness).

In these situations, it will be necessary to change the interpretation of the subjective probability to "degree of belief for element i to come sufficiently close to the truth" where the meaning of "sufficiently" is often not explained in mathematical terms. If $\mathrm{sw}(r_i)$ is the subjective probability for element i to come sufficiently close to the true representation, then $\Sigma_{i=1}^{I}\mathrm{sw}(r_i)$ will most likely be larger than 1 as mutual exclusiveness is no longer given. The probabilities will then need to be normalized so that their sum is 1.

In some situations, it will be possible to interpret $\mathrm{sw}(r_i)$ as the subjective probability for element i, from a sufficiently covering set (completeness), to come closest to the true representation (be best among the set of elements such that mutual exclusiveness applies in this respect) so that $\Sigma_{i=1}^{I}\mathrm{sw}(r_i) = 1$.

The interpretation of the subjective probability has now changed from "degree of belief for element i to be true" to the weaker "degree of belief for element i to be best among a set of adequate representations where none is actually true".

These changes of the interpretation of "subjective probability", should they have been made use of for some model uncertainties, would need to be mentioned in the caveats accompanying the analysis results.

In practice, the set of model alternatives will have to be kept rather small since each element of the set needs to be implemented as a supplement to the computer model and will need to be made accessible through the input file. The implemented set of model alternatives will quantify the state of knowledge only approximately since each model in the set is most likely not a true but only a possibly adequate representation of the process or state of affairs in question, and the set is not completely but only approximately covering the population of possibly adequate models. Each element of the set is given an index value, and the elements of the set of index values are now the "possibly true values" of a new data uncertainty that takes the place of the model uncertainty in the input file. A run of the computer model uses the model alternative that corresponds to the index value that is given in the input file for the new uncertain datum representing the model uncertainty. Each model alternative may have its own set of uncertain parameters with values that are specific for the particular computer model application. Their state of knowledge needs to be quantified and expressed by subjective probability if they are potentially important contributors to uncertainty.

For many model uncertainties, it is quite natural to express the state of knowledge by subjective probabilities for a set of model alternatives. Some of these model uncertainties are due to:

– The choice of the degree of complexity

  The existence of a complex model does not in general render a simpler model obsolete. Complex models usually increase the runtime on the computer and often require additional assumptions and introduce new data uncertainties. It is uncertain whether the complex model is necessary for the intended application of the computer model or whether the implemented simpler model provides already sufficiently accurate results. If the implemented model is a complex one, it may be of interest for future applications of the computer model whether a simpler model would have been sufficient. The subjective probability assigned to differing degrees of complexity $i = 1, \ldots, I$ is to be interpreted as the degree of belief that the model with index i represents the minimum complexity needed in order to have sufficiently accurate results. The analysis may show that the uncertainty about the degree of complexity required does not contribute significantly to the uncertainty of the results. Or the analysis shows that the difference in complexity does lead to significantly different model results. It may then sometimes be clear in retrospect that the more complex models would need to be used. This kind of uncertainty differs therefore from all others, as there may now be reason to base the uncertainty statement for the model result on only that subsample of values, provided by the uncertainty analysis for the computer model result, that was obtained with model alternatives of sufficient complexity.

  For example, it may not always be obvious whether 3D modelling of certain parts of a system is required in order to improve the computer model results over those obtained from the use of a 2D model exploiting symmetries or even a "lumped parameter" model where parameter values and values of descriptive variables are averaged over spatial compartments. Another example may be the uncertainty whether the modelling of the behaviour of airborne contaminant particles in a vessel requires the detailed modelling of the water film on the vessel walls or whether global consideration of this film is sufficient.

– The selection of the scenario

  The uncertainty about how to model the scenario lends itself directly to a quantification of the state of knowledge by a set of different possibly true or adequate scenario descriptions. For instance, the application of a computer model predicting the evolution of an investment portfolio over the next year requires the consideration of a sequence of possible advantageous but also of disadvantageous events that might occur during this time. This sequence will be uncertain with respect to the type of events and their timing. A set of different representations of the sequence may be a natural choice for the quantification of the respective state of knowledge.

  A computer model application that is to answer the question of profitability of the exploitation of a certain oil reservoir requires a sufficiently accurate description of the geology as well as a forecast of the evolution of the oil price over the

time period in question. Again, a set of alternative possibly adequate descriptions may be a natural choice for the quantification of the respective state of knowledge.

In other instances, the model alternatives may differ by their assumptions about the unknown spatial or temporal distribution of airborne contaminants, or of the population density around a potentially hazardous industrial site or of the agricultural usage of an area in the coming years or far back in the past.

A computer model application that is to answer the question of how the global climate will develop over the coming decades will need to consider a variety of possible evolutions of the worldwide population and of its industrial and agricultural activities.

– The selection of the relevant phenomena

  The uncertainty of whether or not a certain phenomenon is relevant for the question that is to be answered by the computer model application would be an example of the ideal case of representing model uncertainty by a set of model alternatives. There are only two alternatives namely "the phenomenon is relevant", which means the computer model runs in a version that does include a model of this phenomenon, and "the phenomenon is not relevant", i.e. it runs in a version without a model of the phenomenon. The new uncertain datum in the input file has, therefore, only two possibly true values. An example would be the uncertainty whether a computer model application that is to evaluate the reliability of a technical system needs to consider component ageing or not. This type of uncertainty is also called a "zero, one" uncertainty.

– Functions fitted to observational data

  The model alternatives may be derived from different sets of observations. Or they may be derived from the same set of observations but different researchers fitted different functions. Each researcher may have possibly applied different criteria to judge the fit. It may be uncertain which of the alternatives is to be preferred for the intended application of the computer model.

– Discretizations, categorizations, groupings

  The model alternatives may differ by assumptions of how one should discretize a complex technical system, categorize a large set of actions or group the members of a population in order to be sufficiently accurate within the budget available for the application of the computer model. For instance, a computer model application that is to answer the question of how a system of pressurized piping and vessels will react to a sudden loss of pressure will work with a discretized representation of the system and it may be uncertain how best to discretize. A computer model that is applied to answer the question of how a population of an endangered species of animals will evolve over time may use a grouping into age classes. It may be uncertain how the grouping needs to be chosen in order to capture all relevant age dependent effects with sufficient accuracy. A suitably chosen set of discretizations or groupings will be used in the uncertainty analysis together with assigned subjective probabilities. The probabilities express the state of knowledge about which element in the set is thought of providing the minimum of detail required in order to have sufficiently

accurate results of the computer model application. If the elements in the set represent increasing levels of refinement, then the remarks made above in connection with differing degrees of complexity will apply.

– Representatives of populations

  Representatives need to be selected from subpopulations, categories, groups, etc. for computational efficiency since it will not be possible to consider all of their members individually in the application of the computer model. Clearly, there are degrees of freedom in the choice of representatives. Each choice may lead to different results, and it will be uncertain by how much these results will differ. The state of knowledge will, therefore, need to be expressed by a suitably chosen variety of different sets of representatives. For instance, the population of events of sudden loss of pressure mentioned above will need to be divided into categories characterized by opening diameter, initial pressure, roughness of the flow path and so forth. A representative needs to be chosen from each category for the computer model application. The uncertainty analysis may need to consider various sets of representatives.

3.4.2 Two Extreme Models

It will not always be feasible to express a model uncertainty by a set of model alternatives. In this case, two extreme models may be developed that delimit the population of possibly true computer representations of the process or state of affairs concerned. None of the two models is considered to be possibly adequate or even true since it is known that they always produce results from beyond the extreme upper and lower end of the range of possibly true results. The model uncertainty may then be represented by an additional uncertain parameter s and the computer model application uses

$$m = sm_u + (1 - s)m_o$$

m	is a possibly true model result
m_u, m_o	is the pair of results obtained from the extreme models
s	is an uncertain datum.

The expert expresses his state of knowledge for the position of the true result relative to m_u and m_o by a subjective probability distribution for s over the interval (0, 1). In case the computer model needs to make multiple uses of the uncertain model during its application, the question arises whether s should always have the same value or whether the value of s should vary. In the latter case, the expert needs to decide whether the subjective probability distribution should also be different in these instances. He would also need to judge whether there is state of knowledge dependence (see Sect. 3.5) among the various values of s to be used during the computer model application.

3.4.3 Corrections to the Result from the Preferred Model

In many instances, the uncertainty about the difference between the true computer representation and the representation by the preferred model can be expressed by an uncertain correction to the results from the preferred model. The state of knowledge for the needed correction will be expressed by a subjective probability distribution. The preference for a specific model may be founded in a qualitative judgment on model features like structure of the model, discretization strategy used (if any), developmental database, applicability (ease of use), availability (it may be the model presently implemented). Other factors may be the relevance and scope of its validation process or the simple fact that there is no alternative to the model presently implemented. The most commonly used correction terms are additive and/or multiplicative, i.e. in the form

$$m = am_p + b$$

m is used in the uncertainty analysis as a possibly true model result
m_p is the output from the preferred model
a, b are uncertain data.

The state of knowledge for the values of a and b needed in order to have a model result that can be considered as possibly true will have to be expressed by a bivariate subjective probability distribution. This distribution will be based on validation experience, if applicable. If the model is applied at various instances within the computer model application, then it will be necessary to judge whether there is state of knowledge dependence between the pairs of values needed in those instances. It will also be required to judge whether the state of knowledge varies from instance to instance (Siu et al. 1992; Park et al. 2010) and therefore has to be expressed by different subjective probability distributions for the values of the uncertain data a and b. Even if a set of model alternatives is used to express the state of knowledge, it may be necessary to assign correction terms to each. With respect to the elicitation of the state of knowledge for the data a and b, see Sect. 3.6.1.

Figure 3.1 compares the measurement values from validation experiments to the corresponding model results. Least squares linear regression of the measurement value on the computed value may be suitable in order to obtain estimates for the correction parameters. If the measurement errors are negligible, the regression coefficients may be used directly as the correction parameters. Sometimes, the number of validation experiments is rather small. In this case, the confidence interval information from the regression analysis will be needed to derive state of knowledge quantifications for the correction terms.

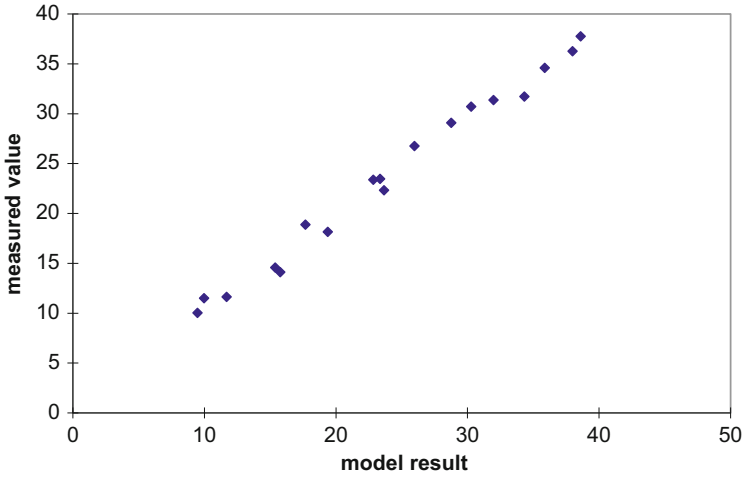


Fig. 3.1 Comparison of measurement values from validation experiments to the corresponding model results

3.4.4 Issues

Issues are processes or states of affair that have not yet been subjected to successful modelling attempts. They may even be subject to scientific controversy and all that is available are hypotheses to work with. Frequently, the respective state of knowledge is not sufficiently advanced to develop a model. In this case, the state of knowledge is directly elicited for the quantities in question. To this end, the range of application (of the non-existing model) is divided into a set of representative cases, the so-called case structure, and the state of knowledge is elicited from the expert for the value of the quantity of interest in each of the cases of this structure. The expert quantifies his state of knowledge by a subjective probability distribution for the value in question. As before, state of knowledge dependence between the cases will need to be quantified. Often the subjective probability distributions are obtained from a decomposition of the issue into a sequence of sub-issues that may be more amenable to state of knowledge quantification. The state of knowledge for the actual issue result is then obtained by propagation of the states of knowledge for the sub-issues through this decomposition. Different experts may choose different issue decompositions.

3.4.5 Some Final Remarks

There are no methodological limits to the consideration of model (or modelling) uncertainty within the uncertainty analysis of the results from a computer model application. There may, however, be practical limits. For instance, it will not always

be affordable to have all possibly adequate model alternatives implemented in the computer model. Furthermore, within the budget available for the application of the computer model, the number of uncertain data and models for which the state of knowledge can be expressed by subjective probability distributions will be limited. Rather it will be attempted to concentrate the resources on those that are considered potentially important contributors to the uncertainty.

The identification of model alternatives and the specification of subjective probabilities for these alternatives require expert judgment. The elicitation of this judgment needs to follow a structured approach to reduce the influence of various sources of bias (Vick 2002), to generate adequate documentation and thus to have sufficient traceability of the analysis results (see Sect. 3.7).

As has been shown above, approaches do exist that permit the inclusion of the effect of model uncertainties into a combined quantitative uncertainty statement as is necessary for decision-making and for comparisons of computer model results with safety standards and limit values. Some of these approaches can be followed without significantly increasing the actual analysis effort. Among the additional benefits of a combined treatment of model and data uncertainties are answers to questions concerning the relative importance of phenomenological uncertainties, the relevance of uncertainties in numerical solution processes, the decision as to whether the state of knowledge needs to be improved primarily on the side of parameters and input data (data gathering) or on the modelling side (model development).

3.4.6 Completeness Uncertainty

Uncertainty about the completeness of the computer model is actually a model uncertainty. Accordingly, the uncertainty contribution from intentional model simplifications, omissions of phenomena, the simplifying use of representatives instead of all members of a category, class or group is already treated in the respective sections above.

Completeness uncertainty that is due to unintentional simplifications and omissions can naturally not be subject of the uncertainty analysis. Quantitative conclusions could possibly be drawn from learning processes in some fields where computer models have been applied for quite a while.

It can also not be the subject of the uncertainty analysis to consider the influence of programming and input errors (like erroneous encoding of the model, typing the wrong input number, using the wrong physical unit, choosing the wrong model option, etc.) In some cases, the uncertainty importance analysis (see Chap. 6), performed in combination with the uncertainty analysis, may hint at such errors (see Sect. 6.5) as practical experience from a number of uncertainty analyses has shown.

3.5 Ways to Quantify State of Knowledge Dependence

The epistemic uncertainties have been categorized into uncertain data and modelling uncertainties in Sect. 3.2. In Sect. 3.4 the latter have been represented by uncertain data. The uncertainty analysis investigates the computer model application for different values of these data. They may, therefore, be called uncertain parameters of the computer model application. Consequently, this section uses the term "uncertain parameter" for simplicity and without loss of generality. In other subchapters, it will, for the sake of convenience, be preferred to use the term "uncertain data" which, as has been explained above, amounts to the same.

The state of knowledge of each uncertain parameter P_m, $m = 1, \ldots, M$ is expressed by a subjective probability distribution $\mathrm{F_m}(p_m)$ with density function $\mathrm{f_m}(p_m)$. These distribution and density functions quantify the unconditional state of knowledge, i.e. the state of knowledge that is not conditioned on any specific value of any of the other uncertain parameters. These individual subjective probability distributions are called the marginal subjective probability distributions and their density functions are the marginal subjective probability density functions. The joint subjective probability density for any combination of possibly true values of the M uncertain parameters is the product of the corresponding M marginal density values provided the M uncertain parameters are state of knowledge independent, i.e. the state of knowledge of none of the parameters depends on the value thought to be true for anyone of the other $M - 1$ parameters. If any two of the parameters have contributors (not necessarily additive) to their uncertainty in common, their states of knowledge are dependent and the joint density is no longer the product of the individual densities.

If the uncertainty analysis of a computer model application uses sets of parameter values drawn independently according to their marginal distributions, the inherent assumption is that any value combination selected according to the marginal distributions makes sense and its subjective probability density equals the product of the corresponding individual density values, i.e. the states of knowledge are independent. However, quite often the assumption of state of knowledge independence does not adequately reflect the available information. For example, if x_i and x_j are measurement values of the unknown true values z_i and z_j, required in a computer model application, ε_S is a positive measurement bias and $(\varepsilon_{R,i}, \varepsilon_{R,j})$ is the pair of independent additive random measurement errors so that

$$z_i = x_i - \varepsilon_S - \varepsilon_{R,i} \quad \text{and} \quad z_j = x_j - \varepsilon_S - \varepsilon_{R,j}$$

then z_i and z_j are uncertain parameters (say: P_i and P_j) since the bias term and the random measurement errors are unknown. The marginal subjective probability distributions express the state of knowledge independently for P_i and for P_j. The uncertainty of P_i and of P_j is, however, partly due to a common contributor, the measurement bias. Knowledge of the bias improves the state of knowledge of P_i and of P_j, i.e. the uncertain bias term makes P_i and P_j state of knowledge dependent.

Another example:
Let the total intake of a dairy cow be made up of four different categories of dry matter. Their fractions of the total intake are represented in a computer model application by the uncertain parameters P_1, P_2, P_3, P_4. The following uniform subjective probability distributions express the state of knowledge for each fraction independently:

State of knowledge of	is quantified by a uniform distribution over the interval
P_1	[0.1, 0.3]
P_2	[0.1, 0.5]
P_3	[0.3, 0.7]
P_4	[0.05, 0.25]

i.e. values below or above the interval enclosed by the brackets are considered as not possibly true while all others are thought to be equally possibly true.

Choosing a set of values $\{p_1, p_2, p_3, p_4\}$ independently according to these four marginal distributions would most likely violate the condition that fractions of the total intake have to add up to 1. This condition introduces state of knowledge dependence since

$$P_i = 1 \quad - \sum\nolimits_{k \neq i} P_k,$$

i.e. the state of knowledge of any pair of fractions has not only the remaining fractions as common uncertainty contributor but also the value of the other parameter in the pair.

3.5.1 How to Identify State of Knowledge Dependence?

Any two or more output quantities Y of the computer model will be state of knowledge dependent provided they have at least one uncertain parameter or model formulation in common. The same applies to any pair of interim results Z of the computer model.

While modelling ends at the level of the uncertain parameters, $P_1, \ldots, P_M$, they frequently share uncertainty contributors Q as shown in Fig. 3.2. Any pair of uncertain parameters sharing uncertainty contributors is state of knowledge dependent and so are any interim model results that share any of the M uncertain parameters, and so forth.

Identification of state of knowledge dependence requires a systematic search for any information that could lead to the abandoning of the assumption of state of knowledge independence. Consequently, in order to identify reasons for state of knowledge dependence, the following question is to be asked: "Does the parameter pair (P_i, P_j) have any contributors to their uncertainty in common?" or "Would the state of knowledge for P_i change if more became known about P_j or vice versa?"

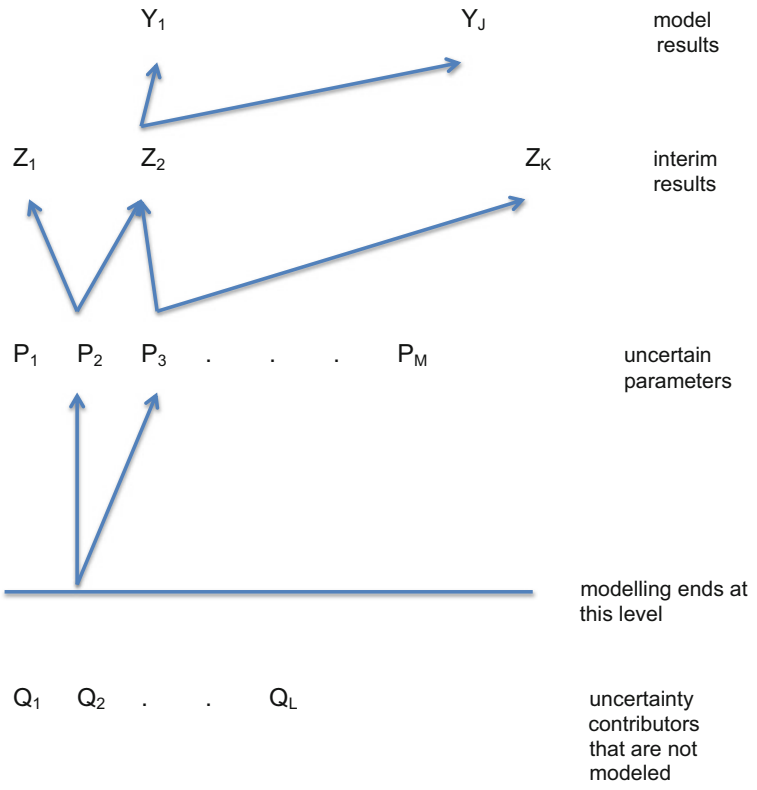


Fig. 3.2 Influence diagram of uncertainty contributions

With the uncertainty contributors $Q_1, Q_2, \ldots, Q_K$ assumed to be state of knowledge independent, one may specifically ask:

- Are there state of knowledge independent uncertainty contributors Q_1, Q_2, Q_3 such that
  $P_i = h_i(Q_1, Q_2)$ and $P_j = h_j(Q_1, Q_3)$?
  The relationships h_i and h_j may only be vaguely known.
  Examples for P_i and P_j:
  - The concentration of a contaminant in two different species of fish living in the same environment.
  - The effective porosity of two adjacent spatial compartments in the same geological formation.
  - Measurements of concentrations of a contaminant in two samples measured with the same device, following the same procedure.

- Are there relationships h_i, h_j such that
  $P_i = h_i(Q_2)$
  $P_j = h_j(P_i, Q_3)$
  where Q_2 and Q_3 are state of knowledge independent uncertainty contributors?
  Examples for P_i and P_j:
  - Food intake and body mass of a specific individual.
  - Body mass and skin surface.
  - Mass median diameter of particles and total particle mass in the atmosphere.
  - Contaminant concentration in prey and predator.
  - Biomass and interception factor.
- Are there relationships h_i and h_j such that
  $P_i = h_i(Q_2)$
  $P_j = h_j(P_i)$
  h_j may not exactly but predominantly be a function in the argument P_i alone?
- Are there relationships h_i and h_j such that
  $P_i = h_i(Q_k, k = 1, \ldots, i), i = 1, \ldots, I$
  $P_j = h_j(P_1, \ldots, P_I)$
  $P_j = 1 - \sum_{i=1}^{I} P_i, 0 \leq P_i, P_j \leq 1$?
  Examples for P_i and P_j:
  - Fractions of intake of dry matter (in a dairy cow diet) from each of an exhaustive set of non-overlapping categories i, $i = 1, ..., I$.
  - Fraction of predator diet consisting of each of an exhaustive set of non-overlapping categories of prey i, $i = 1, \ldots, I$.
- Is stochastic (statistical) dependence involved?
  Statistical dependence of stochastic variables X_1 and X_2 is due to shared contributors to their stochastic variability. If a pair of values (x_1, x_2) is required as input to a computer model application while all that is known about these values is that the pair was drawn at random according to the joint probability distribution quantifying the stochastic variability of (X_1, X_2), then (x_1, x_2) is a pair of state of knowledge dependent uncertain parameters (say: P_i and P_j) of the model application. The state of knowledge dependence of P_i and P_j is due to shared contributors to their uncertainty namely the shared contributors to the stochastic variability of X_1 and X_2.

With the help of some analysis, it is usually possible to identify the uncertainty contributors shared by two or more uncertain parameters. For many pairs of parameters, it may be immediately clear that they do not share contributors to their uncertainty. For many of the remaining pairs shared contributors may safely be judged as negligible. Several of the still remaining pairs of uncertain parameters may

share major uncertainty contributors, but the influence of these parameters on the uncertainty of the computer model result may be judged as minor, i.e. they are potentially important but will definitely not rank among the main uncertainties. State of knowledge dependences will only be of importance if both uncertain parameters of the state of knowledge dependent pair are main contributors to the uncertainty of the model result due to the shared uncertainty contributors. Often this possibility cannot be safely excluded prior to the uncertainty analysis. Neglect of state of knowledge dependence may be uncritical (Smith et al. 1992) but may also cause substantial over- or underestimation of the model result uncertainty and may even lead to a shift of the state of knowledge expression for the model result to safe or unsafe values (if compared to safety limits).

The following example illustrates the effect of significant state of knowledge dependence between two uncertain parameters. Uncertainty analysis using Monte Carlo simulation provided the empirical subjective probability distribution for the model result shown in Fig. 3.3a. The two main uncertainty contributors P_i and P_j are state of knowledge dependent. The dependence is quantified by a correlation coefficient of 0.8. This coefficient says that a shared contributor to uncertainty leads to a positive almost linear relationship between the values thought to be true for P_i and P_j. Consequently, there is reason to think that large (small) values of P_j tend to be true if a large (small) value is true for P_i (and vice versa). The little square box on the abscissa indicates the model result value $R = 118$ that is not exceeded by at least 95% of the possibly true values at a confidence level of at least 95%, i.e. an upper (95%, 95%) tolerance confidence limit (see Chap. 5) of the model result. Figure 3.3b shows the same model result for the same marginal subjective probability distributions of the uncertain parameters but this time P_i and P_j are known not to share an uncertainty contributor, i.e. their states of knowledge are independent and the upper (95%, 95%) tolerance confidence limit of the model result is now $R = 103$. Finally, in Fig. 3.3c, P_i and P_j are state of knowledge dependent; however, this time the dependence is quantified by a correlation coefficient of -0.8. This coefficient says that a shared contributor to uncertainty leads to a negative almost linear relationship between the values thought to be true for P_i and P_j. Consequently, there is reason to think that large (small) values of P_j tend to be true if a small (large) value is true for P_i (and vice versa). The upper (95%, 95%) tolerance confidence limit of the model result is now $R = 80$.

3.5.2 *How to Express State of Knowledge Dependence Quantitatively?*

Saying that the uncertain parameters P_i and P_j are state of knowledge dependent means that the state of knowledge of P_j depends on the value thought to be true for P_i and vice versa. State of knowledge dependence is quantitatively expressed by a joint subjective probability density function. Its marginal densities are the individual state

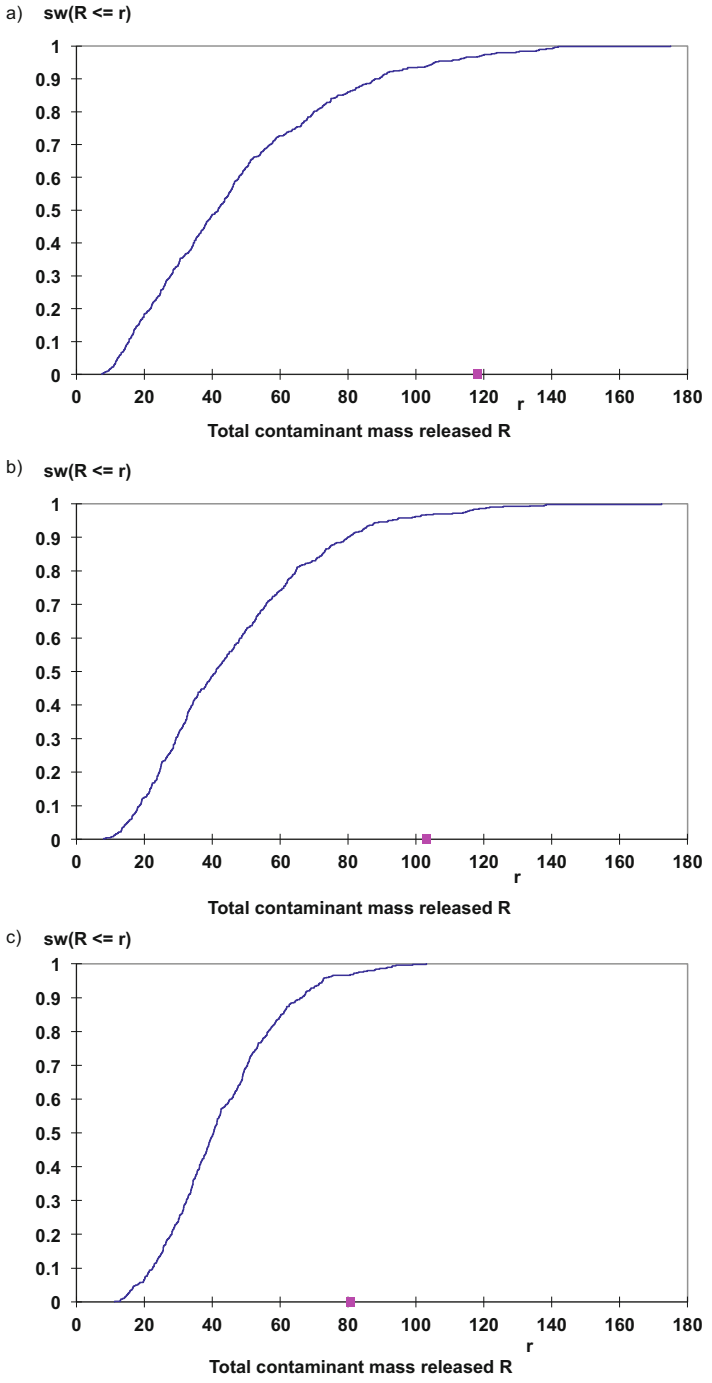


Fig. 3.3 (**a**–**c**) Three empirical (e.g. obtained by Monte Carlo simulation) subjective probability (sw) distributions quantifying the uncertainty for the same computer model result. The states of

of knowledge expressions for P_i and P_j. The value of the joint density function at the point $P_i = p_i$ and $P_j = p_j$ is given by $f_{ij}(p_i, p_j)$. It can be written as

$$f_{ij}(p_i, p_j) = f_i(p_i) f_{j|i}(p_j | P_i = p_i) = f_j(p_j) f_{i|j}(p_i | P_j = p_j) \tag{3.23}$$

$f_i(p_i)$ is the value of the marginal subjective probability density function for P_i at the point $P_i = p_i$, and $f_{j|i}(p_j | P_i = p_i)$ is the value of the conditional subjective probability density function for P_j at the point $P_j = p_j$, given $P_i = p_i$. Quantifying state of knowledge dependence for two uncertain parameters therefore requires either the joint density function or the marginal density function for one parameter and the conditional density function for the other parameter.

The subjective probability for $P_j \leq p_j$, given $P_i = p_i$, follows as

$$\mathrm{sw}(P_j \leq p_j | P_i = p_i) = \int_{p_{j,min}}^{p_j} f_{j|i}\left(p'_j | P_i = p_i\right) \mathrm{d}p'_j$$

while in the case of state of knowledge independence

$$\mathrm{sw}(P_j \leq p_j | P_i = p_i) = \mathrm{sw}(P_j \leq p_j) = \int_{p_{j,min}}^{p_j} f_j\left(p'_j\right) \mathrm{d}p'_j.$$

The joint subjective probability density of three state of knowledge dependent uncertain parameters P_i, P_j and P_k may be written as

$$f_{ijk}(p_i, p_j, p_k) = f_i(p_i) f_{j|i}(p_j | P_i = p_i) f_{k|ij}(p_k | P_i = p_i, P_j = p_j) \tag{3.24}$$

$f_{k|ij}(p_k | P_i = p_i, P_j = p_j)$ is the state of knowledge expression for P_k given p_i and p_j are the values thought to be true for P_i and P_j, respectively. In the case of pairwise state of knowledge independence, the joint density value is simply the product of the marginal density values.

3.5.2.1 Conditional Subjective Probability Density Functions

Specifying a conditional subjective probability density function for P_j to each and every value thought to be true for P_i will be too much of a challenge in most practical situations. It will often be sufficient to divide the value range of P_i into a small

Fig. 3.3 (continued) knowledge of the uncertain parameters are expressed by the same marginal subjective probability distributions, but the states of knowledge of the two main contributors to model result uncertainty are: (**a**) dependent, with the dependence quantified by a correlation coefficient of 0.8; (**b**) independent; (**c**) dependent, with the dependence quantified by a correlation coefficient of – 0.8. A (95%, 95%) one-sided upper statistical tolerance limit (or tolerance confidence limit, see Chap. 5) for the model result is indicated on the abscissa. This limit would most likely be the value to be compared to a release limit laid down in a regulatory standard

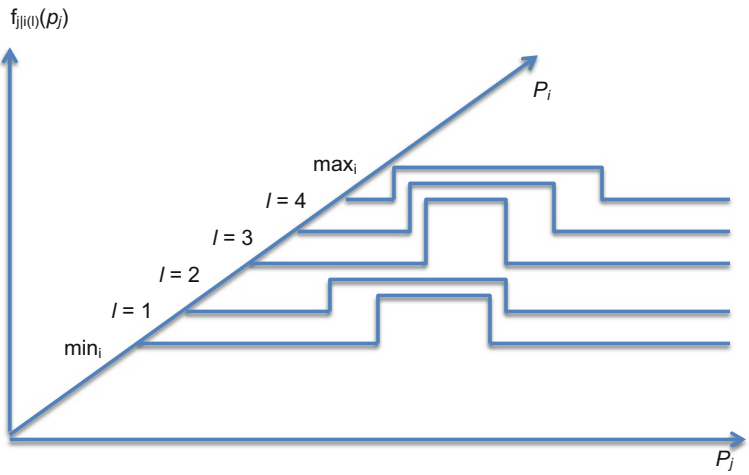


Fig. 3.4 Schematic illustration of state of knowledge dependence between P_i and P_j expressed by five conditional subjective probability density functions for P_j to given intervals for the value of P_i

number (L) of exhaustive non-overlapping intervals (see Fig. 3.4) and to specify the conditional subjective probability density function for P_j given the value thought to be true for P_i is from the respective interval.

The corresponding approximate expressions for the joint subjective probability density function f_{ij} of P_i and P_j are

$$f_{ij}(p_i, p_j) \approx f_{j|i(l)}(p_j)\, f_i(p_i) \quad \text{for } p_i \in i_l, l = 1, \ldots, L \tag{3.25}$$

$f_{j|i(l)}(p_j)$ is the specified conditional density function for P_j given the value of P_i is from the interval i_l.

Figures 3.5 and 3.6 show random samples obtained according to this simplified quantitative expression of state of knowledge dependence while Fig. 3.4 is a schematic illustration of the conditional density functions with the intervals for P_i delimited as follows:

$min_i < b_1 < b_2 < \ldots < b_{L-1} < max_i = b_L$ and $[min_i, max_i]$ is the interval of possibly true values for P_i.

This way it is possible to account for varying degrees of linear, nonlinear but monotone and even non-monotone relationships that are judged to exist between the values thought to be true for P_i and P_j. This population expression for state of knowledge dependence is versatile enough to be of practical use also in situations where the state of knowledge of P_j depends on the value thought to be true for an interim result of the computer model application. It is also suitable for modelling the state of knowledge dependence between successive elements of a sequence of uncertain data over time or any other independent variable. For instance, if time is the independent variable then P_i corresponds to the value of the uncertain parameter P at time $t - 1$ and P_j corresponds to the value of P at time t.

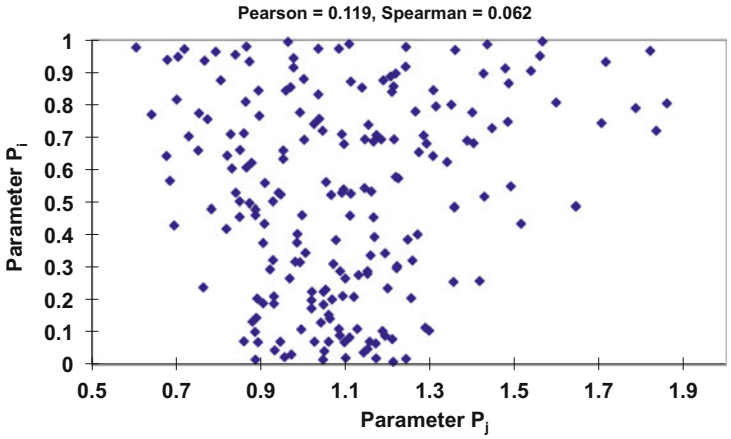


Fig. 3.5 The effect of state of knowledge dependence quantified by five conditional subjective probability distributions for P_j with a (0, 1) uniform marginal subjective probability distribution specified for P_i. The value range of P_i is divided into five intervals of equal subjective probability

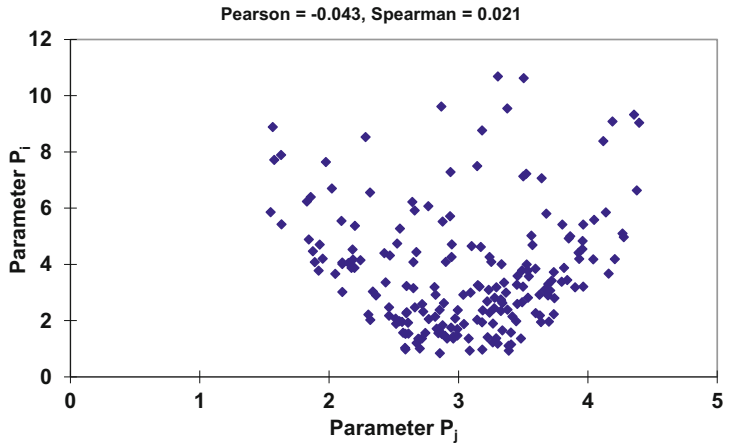


Fig. 3.6 The effect of state of knowledge dependence quantified by five conditional uniform subjective probability distributions for P_j with a (1.125, 0.557) lognormal marginal subjective probability distribution specified for P_i. The value range of P_i is divided into five intervals of equal subjective probability

The quantification of state of knowledge dependence by conditional density functions specified for intervals of the conditioning uncertain parameter(s) is, for practical reasons, limited to small numbers of state of knowledge dependent uncertain parameters. For instance, if $L = 3$ is chosen and the number K of pairwise state of knowledge dependent uncertain parameters is 4 then $\sum_{k=1}^{K-1} L^k = 39$ conditional

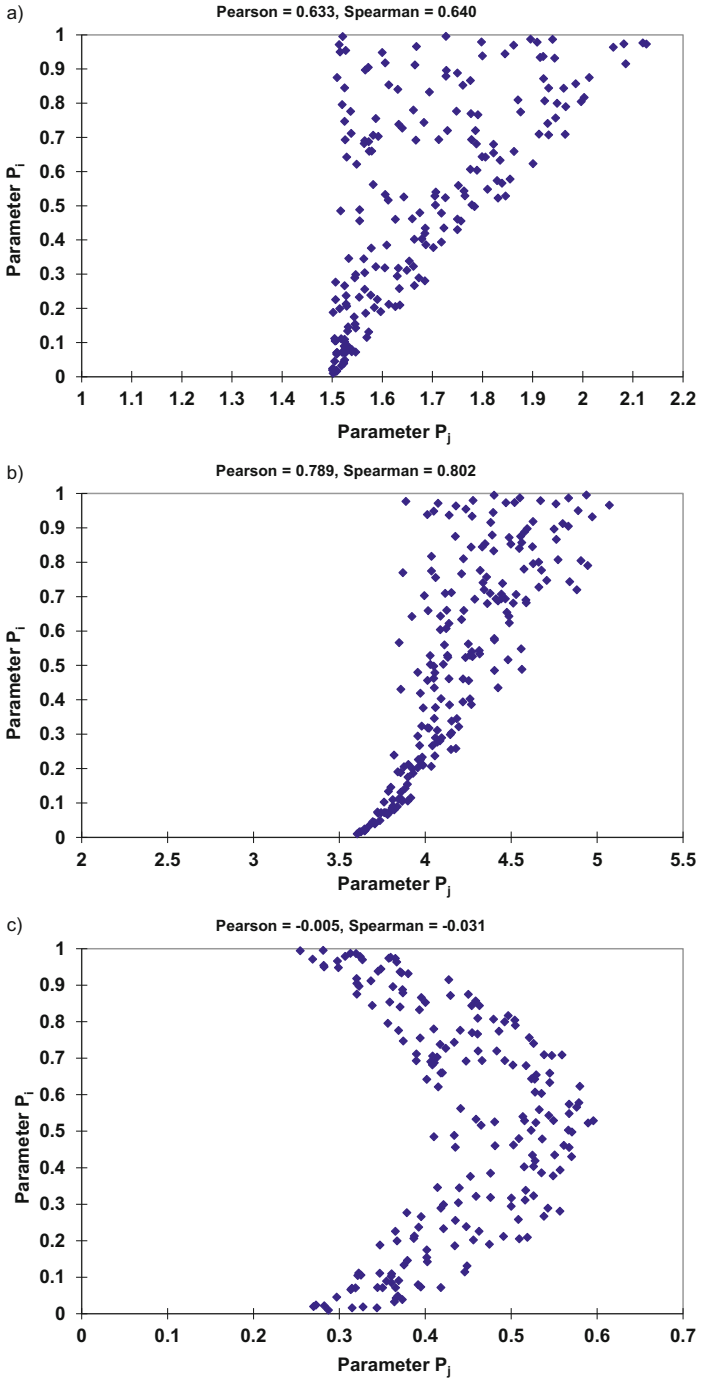


Fig. 3.7 (**a**–**c**) Three examples of state of knowledge dependence quantified by conditional subjective probability distributions for P_j with their parameters given as a continuous function of

subjective probability density functions would need to be specified, namely 3 for P_2 (given P_1), 9 for P_3 (given P_2 and P_1) and 27 for P_4 (given P_1, P_2, P_3).

Specifying the state of knowledge by a finite number of conditional subjective probability distributions for P_j may be sufficient in a number of situations. However, in some cases, it will be necessary to define the parameters of the conditional distribution as continuous functions of the value thought to be true for P_i. Figure 3.7a–c shows random samples drawn according to three examples of state of knowledge dependence. The dependence is quantified by conditional subjective probability density functions for P_j with their distribution parameters given as continuous functions of the value thought to be true for P_i. The state of knowledge of P_i is expressed by a [0, 1] uniform subjective probability distribution in these three examples.

Figure 3.7a–c serve to illustrate the flexibility of state of knowledge dependence modelling offered by conditional distributions, particularly if the distribution parameters of P_j are continuous functions of P_i.

The quantification of state of knowledge dependence by conditional density functions, with their distribution parameters given as functions of the conditioning uncertain parameter(s), has practical limits. The number of arguments of the functions that are to be specified for the distribution parameters grows with every uncertain parameter that is included in the conditioning set. If each conditional density function has two distribution parameters and K is the number of pairwise state of knowledge dependent uncertain parameters, then $K - 1$ conditional density functions are required and therefore $2(K - 1)$ functions for their parameters with a minimum of 1 argument value and a maximum of $K - 1$ argument values (i.e. one value each of the uncertain parameter(s) in the conditioning set of the conditional density function).

3.5.2.2 Restricting Inequalities

Quite frequently, the value thought to be true for P_j has to comply with specific constraints that are formulated as a function of the value thought to be true for P_i and possibly additional uncertain parameters. For instance, in the example of Fig. 3.8, the range of possibly true values of P_j is delimited by $g_j(p_i)$ where p_i is the value thought to be true for P_i. The conditional subjective probability density function for P_j can be specified such that it complies with this constraint.

In many practical situations, however, both marginal subjective probability distributions (for P_i and for P_j) are given together with constraints such as

Fig. 3.7 (continued) the value of P_i and the state of knowledge of P_i quantified by a [0, 1] uniform subjective probability distribution

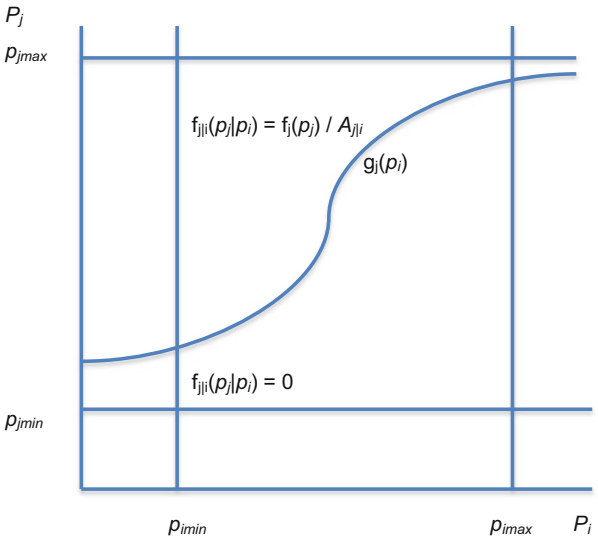


Fig. 3.8 Construction of a joint subjective probability density function for P_i and P_j that applies the specified marginal distributions and complies with the constraint $P_j >(\leq) g_j(P_i)$. It is obvious that, due to the restriction, the resulting marginal distribution for P_j may have to differ from the specified marginal distribution

$$P_j \geq g_j(P_i) \text{ and } g_j(p_i) \leq p_j \leq p_{j,max} \tag{3.26}$$

or

$$P_j \leq g_j(P_i) \text{ and } g_j(p_i) \geq p_j \geq p_{j,min}$$

for all values p_i from $[p_{i,min}, p_{i,max}]$ instead of the marginal density function for P_i and the conditional density function for P_j.

These restricting inequalities, with the function g_j specified in addition to the marginal distributions for P_i and for P_j, introduce state of knowledge dependence. The given information may be used to define a joint subjective probability density function in the following way:

$$f_{ij}(p_i, p_j) = f_i(p_i) f_{j|i}(p_j|p_i) = f_i(p_i) f_j(p_j) / A_{j|i} \tag{3.27}$$

$$A_{j|i} = \int_{g_j(p_i)}^{p_{jmax}} f_j(p_j) dp_j = 1 - F_j\left(g_j(p_i)\right) \quad \text{for} \quad p_j > g_j(p_i)$$

$$A_{j|i} = \int_{p_{jmin}}^{g_j(p_i)} f_j(p_j) dp_j = F_j\left(g_j(p_i)\right) \quad \text{for} \quad p_j \leq g_j(p_i)$$

Random sampling according to the joint density function is explained in Chap. 4. Figure 3.9 shows a random sample that complies with the constraint $P_j \leq g_j(P_i) = P_i$.

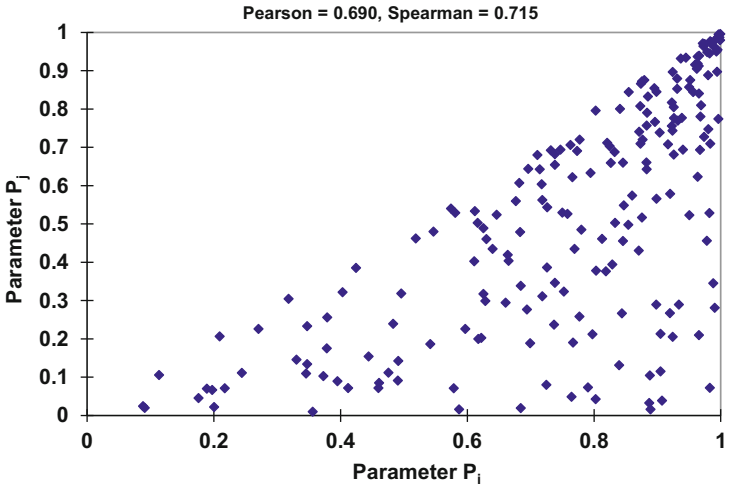


Fig. 3.9 The effect of state of knowledge dependence due to the requirement that the value thought to be true for P_j always has to be smaller or equal to the value thought to be true for P_i

3.5.2.3 Known Functional Relationships

Sometimes, there is a known functional relationship between P_j and P_i as well as other uncertain parameters P_k, k = 1,. . .,K

$$P_j = \mathrm{g_j}(P_i, P_1, P_2, \ldots, P_K)$$

that is not part of the computer model and the model is not supposed to be amended by this relationship or it cannot be amended because the source code is not available (see Figs. 3.10 and 3.11a–c). If $\mathrm{g_j}$ is uncertain, it may need to be treated as a model uncertainty. Subjective probability distributions are specified only for the so-called free uncertain parameters P_i, P_1, P_2, …, P_K. Some or all of the parameters in the argument list of $\mathrm{g_j}$ may be state of knowledge dependent.

If K uncertain parameters P_1, P_2, …, P_K are uncertain fractions of a total C and $\Sigma_{k=1}^{K} P_k = 1$ while $0 \le P_k \le 1$, $k = 1, \ldots, K$ so that $C = \Sigma_{k=1}^{K} CP_k$, then they are state of knowledge dependent due to this constraint for the sum. This type of dependence is modelled with the help of $K - 1$ state of knowledge independent uncertain parameters, namely the conditional fractions Q_1, Q_2, …, $Q_{K\text{-}1}$. $Q_k\left(C - \Sigma_{j=1}^{k-1} CP_j\right)$ equals P_kC, i.e. Q_k is the fraction under the condition that P_1C, …, $P_{k-1}C$ have already been subtracted from the total C. This requires to express the state of knowledge by a subjective probability density function for each of the data Q_k, $k = 1, \ldots, K-1$ instead of a density function for each of the P_k. The following relationships apply

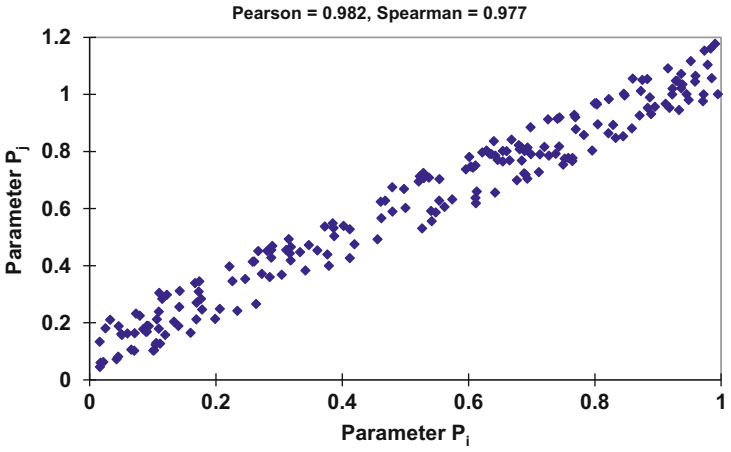


Fig. 3.10 The effect of state of knowledge dependence due to the functional relationship $P_j = P_i + P_k$ with the state of knowledge of P_i quantified by a uniform subjective probability distribution over [0, 1] and that of P_k by a uniform subjective probability distribution over [0, 0.2]. P_i and P_k are state of knowledge independent

$$
\begin{aligned}
&P_1 = Q_1 \\
&P_2 = (1 - Q_1)Q_2 \\
&\cdot \\
&\cdot \\
&\cdot \\
&P_{K-1} = (1 - Q_1) \ldots (1 - Q_{K-2})Q_{K-1} \\
&P_K = 1 - \sum_{k=1}^{K-1} P_k
\end{aligned}
\tag{3.28}
$$

with $\sum_{k=1}^{K} P_k = 1$; $0 \leq P_k \leq 1$, $k = 1, \ldots, K$; $0 \leq Q_k \leq 1$, $k = 1, \ldots, K-1$.

Examples:

- Fractions of the total dry matter intake in a dairy cow diet consisting of K different kinds of dry matter.
- Fractions of the total prey intake in a predator diet consisting of K species of prey.
- Branch point probabilities of an event tree with K branches.

Figure 3.12a, b illustrates the state of knowledge dependence between P_1 and P_2 as well as between P_1 and P_3 that is due to the functional relationship $P_1 + P_2 + P_3 = 1$; $0 \leq P_k \leq 1$, $k = 1, 2, 3$ with a [0.3, 0.5] and a [0.2, 0.4] uniform subjective probability distribution specified as state of knowledge expressions for the state of knowledge independent parameters Q_1 and Q_2.

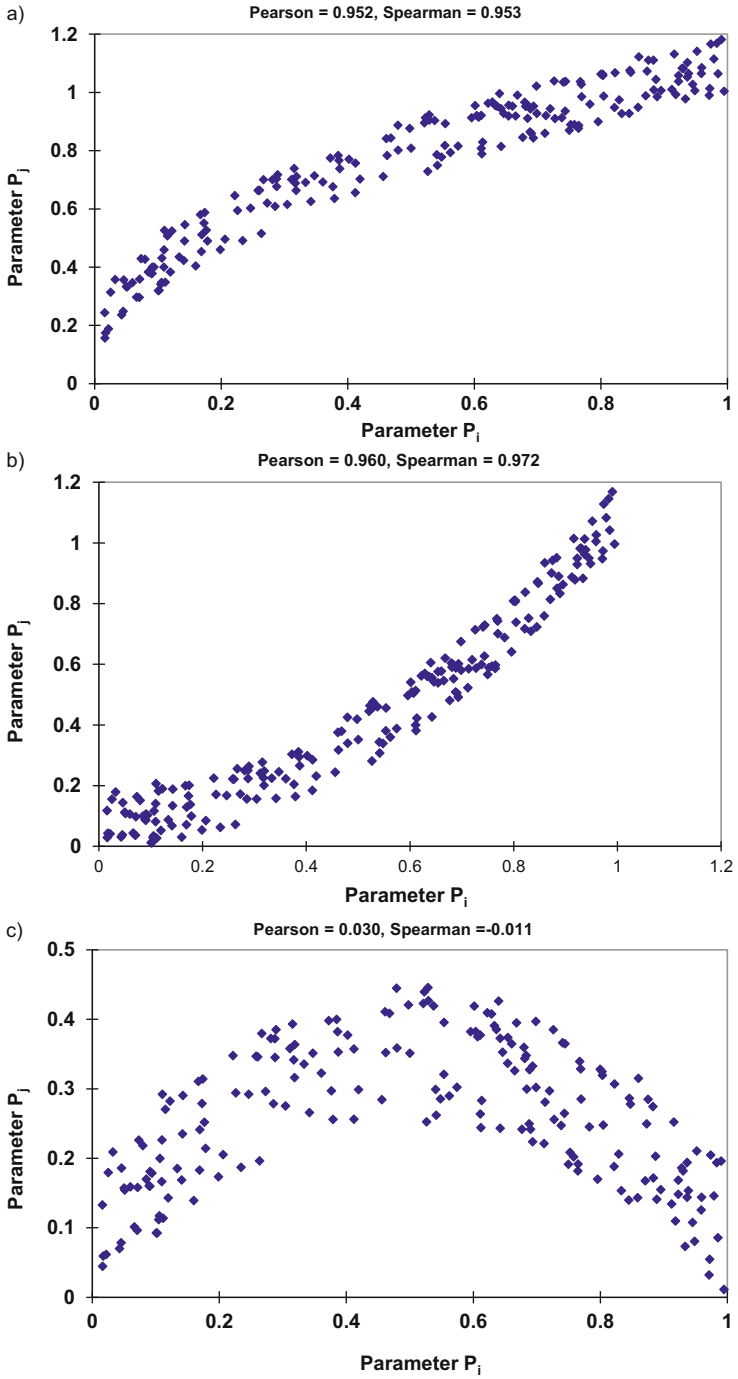


Fig. 3.11 (**a**–**c**) The effect of state of knowledge dependence due to the functional relationships. (**a**) $P_j = (P_i)^{1/2} + P_k$, (**b**) $P_j = (P_i)^2 + P_k$, (**c**) $P_j = (P_i)(1 - P_i) + P_k$ with the states of knowledge of P_i and of P_k specified as for Fig. 3.10

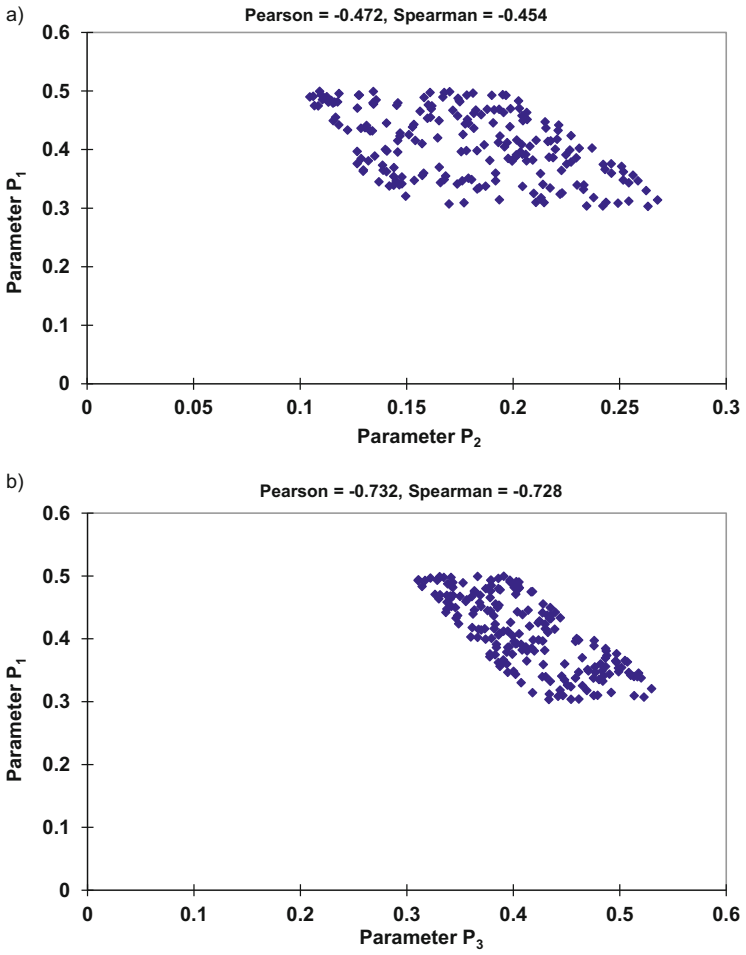


Fig. 3.12 (**a**, **b**) The effect of state of knowledge dependence that results for the pairs (P_1, P_2) and (P_1, P_3) from the functional relationship $P_1 + P_2 + P_3 = 1$ with $0 \leq P_k \leq 1$, $k = 1, 2, 3$

Often state of knowledge dependence is judged to exist between P_j and P_i while only the marginal subjective probability densities are given and neither a constraint nor a functional relationship is specified. In this case, the dependence is frequently quantified by a measure of association like Pearson's or Spearman's correlation coefficient.

3.5.2.4 State of Knowledge Dependence Expressed by Pearson's Correlation Coefficient

Any functional relationship $P_j = g_j(P_i)$ between uncertain parameters P_j and P_i causes state of knowledge dependence. The function g_j may have additional

uncertain data Q_k in its argument list, and it may also be a model uncertainty introducing state of knowledge dependence of P_j not only with the arguments of g_j but also with its functional form. Suspecting a functional relationship but lacking knowledge of its functional form, an approximate linear relationship

$$P_j \approx a_i + b_i P_i = \widehat{P}_j \tag{3.29}$$

is frequently assumed in order to arrive at a quantitative expression for the state of knowledge dependence. Subtracting the mean values in Eq. (3.29) gives

$$W_i = P_i - \mathrm{E}\{P_i\} \quad \text{and} \quad W_j = P_j - \mathrm{E}\{P_j\}$$

and therefore

$$W_j + \mathrm{E}\{P_j\} \approx a_i + b_i(W_i + \mathrm{E}\{P_i\}).$$

Setting $a_i = \mathrm{E}\{P_j\} - b_i\mathrm{E}\{P_i\}$ leads to $W_j \approx b_i W_i$.

b_i is chosen such that the variance $\mathrm{Var}\{W_j - b_i W_i\}$ of the approximation error is minimal. The solution for b_i is obtained from the requirement that

$$\left(\frac{\mathrm{d}}{\mathrm{d}b_i}\right) \int\int (w_j - b_i w_i)^2 \mathrm{f_{ij}}(p_i, p_j)\mathrm{d}p_i\mathrm{d}p_j = 0$$

or

$$\int\int (w_j - b_i w_i) w_i \mathrm{f_{ij}}(p_i, p_j)\mathrm{d}p_i\mathrm{d}p_j = 0$$

so that

$$b_i = \left[\int\int (w_j w_i)\mathrm{f_{ij}}(p_i, p_j)\mathrm{d}p_i\mathrm{d}p_j\right] / Var\{W_i\} = \mathrm{Cov}\{P_i, P_j\}/\mathrm{Var}\{P_i\} \tag{3.30}$$

where Cov{.,.} denotes the covariance operator.

On the other hand, from

$$P_i \approx a_j + b_j P_j$$

follows

$$b_j = \left[\int\int (w_j w_i)\mathrm{f_{ij}}(p_i, p_j)\mathrm{d}p_i\mathrm{d}p_j\right] / Var\{W_j\} = \mathrm{Cov}\{P_i, P_j\}/\mathrm{Var}\{P_j\}$$

Since Pearson's correlation coefficient (also called "ordinary correlation coefficient" or simply "correlation coefficient") is defined as (mean values and variances assumed to exist)

$$\begin{aligned}\rho_{ij} &= \rho(P_i, P_j) = \mathrm{E}\{(P_i - \mathrm{E}\{P_i\})(P_j - \mathrm{E}\{P_j\})\}/(\mathrm{Var}\{P_i\}\mathrm{Var}\{P_j\})^{\frac{1}{2}} \\ &= \mathrm{Cov}\{P_i, P_j\}/(\mathrm{Var}\{P_i\}\mathrm{Var}\{P_j\})^{1/2} \\ &= (\mathrm{E}\{P_iP_j\} - \mathrm{E}\{P_i\}\mathrm{E}\{P_j\})/(\mathrm{Var}\{P_i\}\mathrm{Var}\{P_j\})^{\frac{1}{2}},\end{aligned} \tag{3.31}$$

it follows for the correlation coefficient

$$\rho_{i,j} = \mathrm{sign}(\mathrm{Cov}\{P_i, P_j\})(b_ib_j)^{1/2}$$

i.e. ρ_{ij} is the geometric mean value of the product of the least squares linear regression coefficients b_i and b_j from the regression (see Chap. 6) of P_j on P_i and from the regression of P_i on P_j.

$$\rho_{ij}^2 = \frac{(\mathrm{Cov}\{P_i, P_j\})^2}{\mathrm{Var}\{P_i\}\mathrm{Var}\{P_j\}} = b_i^2\frac{Var\{P_i\}}{Var\{P_j\}} = b_i^{*2}$$

where $b_i{}^*$ is the standardized regression coefficient of the least squares linear regression of P_j on P_i.

The extent of the assumed linear relationship is expressed by $|\rho_{ij}|$ while the sign of ρ_{ij} gives the direction of it (i.e. positive, if a change in the value thought to be true for P_i tends to change the value thought to be true for P_j in the same direction and negative if it tends to change it in the opposite direction). Consequently, ρ_{ij} is an expression of the extent and direction of an approximate (in the least squares sense) linear relationship between the values thought to be true for P_j and for P_i. ρ_{ij}^2 quantifies the fraction of the uncertainty of P_j (as measured by the variance of its subjective probability distribution) that can be explained by the approximate linear relationship with P_i

$$\frac{Var\{\widehat{P}_j\}}{Var\{P_j\}} = \frac{b_i^2Var\{P_i\}}{Var\{P_j\}} = \frac{(\mathrm{Cov}\{P_i, P_j\})^2}{\mathrm{Var}\{P_i\}\mathrm{Var}\{P_j\}} = \rho_{ij}^2. \tag{3.32}$$

From Eqs. (3.29) and (3.32) follows that $\rho(.,.)$ can only assume values from $[-1, +1]$ and from Eq. (3.31) it is evident that, with $b = \mathrm{sign}(b)|b|$ and $b^2 = |b|^2$,

$\rho(P_j, a + bP_i) = \mathrm{sign}(b)\rho(P_i, P_j)$ $(a,b \in \boldsymbol{R})$ where $\boldsymbol{R}$ is the set of all real numbers and

$\rho(P_i, P_j) = 1$ if $P_j = a + bP_i$ i.e. the linear relationship is exact and positive.

Furthermore, if

$$P_i' = a + bP_i, \quad P_j' = c + dP_j, \quad a, b, c, d \in \mathcal{R},$$

then

$$\rho(P_i', P_j') = \mathrm{sign}(bd)\rho(P_i, P_j)$$

since

$$\rho\left(P_i', P_j'\right) = E\{(bP_i + a - bE\{P_i\} - a)(dP_j + c - dE\{P_j\} - c)\} / [Var\{bP_i + a\}Var\{dP_j + c\}]^{\frac{1}{2}} = sign(bd)E\{(P_i - E\{P_i\})(P_j - E\{P_j\})\} / [Var\{P_i\}Var\{P_j\}]^{1/2}$$

Correlation coefficients are frequently used as quantitative measures of state of knowledge dependence. Instead of a set of conditional distributions, there is only one number to be provided namely ρ_{ij}. The two marginal subjective probability distributions and the covariance (as defined by the correlation coefficient) will, however, not suffice to fully determine the joint subjective probability distribution for P_i and P_j, not even in the case where both marginal distributions are of the normal type[4] unless it is assumed that the joint distribution is normal. In the general case, however, there is room for interpretation left, i.e. there is a choice of joint distributions that have the same specified marginal distributions and the same specified value of the correlation coefficient. If the subjective probability distribution for the computer model result is sensitive to characteristics of the joint distribution of P_i and P_j, other than those expressed by the marginal distributions and by the covariance, then a joint subjective probability distribution with minimum additional information should be chosen (see Sect. 3.5.2.7 on copulas).

The following properties need to be kept in mind when specifying values for ρ_{ij} as quantitative expressions of state of knowledge dependence:

a) $\rho(.,.)$ can only assume values from the interval [–1, +1].
b) If P_i and P_j are judged to be state of knowledge independent, then $\rho_{ij} = 0$ is the appropriate choice since

$$\mathrm{Cov}\{P_i, P_j\} = \mathrm{E}\{(P_i - \mathrm{E}\{P_i\})(P_j - \mathrm{E}\{P_j\})\} = \mathrm{E}\{P_i - \mathrm{E}\{P_i\}\}\mathrm{E}\{P_j - \mathrm{E}\{P_j\}\} = 0$$

However, specifying $\rho_{ij} = 0$ does not necessarily imply state of knowledge independence. It only says that $\mathrm{Cov}\{P_i,P_j\} = 0$. The specification of the marginal distributions and of the covariance does not uniquely determine the joint subjective probability distribution (see above for exceptions). $\rho_{ij} = 0$ is a necessary but not a sufficient condition for state of knowledge independence.

c) If the value that is thought to be true for P_j is suspected to be a linear increasing (or decreasing) function of the value that is thought to be true for P_i then $\rho_{ij} = +1$ (–1) needs to be specified.
d) Values other than 0, +1 or −1 specified for ρ_{ij} are a quantitative judgment of the extent of linear relationship that is suspected to exist between the values thought to be true for P_i and for P_j. The larger the chosen absolute value of ρ_{ij}, the more the relationship is judged to be linear, i.e. the larger is the fraction of the variance

[4] If the joint distribution is normal, then both marginals are normal but the joint distribution is not necessarily normal if both marginals are normal.

(as measure of uncertainty) of P_j that is explained by a least squares approximation of P_j in P_i (and vice versa).

If, for example, it is assumed that

$$P_i = Q_1 + Q_2$$
$$P_j = Q_1 + Q_3$$

with Q_1, Q_2, Q_3 pairwise state of knowledge independent uncertainty contributors and Q_1 representing the shared contributor to uncertainty, it follows

$$\rho_{ij} = \rho\left(P_i, P_j\right) = \mathrm{Var}\{Q_1\}/(\mathrm{Var}\{Q_1 + Q_2\}\mathrm{Var}\{Q_1 + Q_3\})^{1/2} \tag{3.33}$$

while for

$$P_i = -Q_1 + Q_2$$
$$P_j = Q_1 + Q_3$$
$$\rho_{ij} = -\mathrm{Var}\{Q_1\}/(\mathrm{Var}\{Q_1 + Q_2\}\mathrm{Var}\{Q_1 + Q_3\})^{\frac{1}{2}}.$$

ρ_{ij} will be the closer to +1 (–1) the smaller $\mathrm{Var}\{Q_2\}$ and $\mathrm{Var}\{Q_3\}$ are relative to $\mathrm{Var}\{Q_1\}$. In particular, if the subjective probability distributions for Q_1, Q_2 and Q_3 are of equal variance, then

$$\rho_{ij} = 0.5\ (-0.5).$$

The type of distributions specified for Q_1, Q_2 and Q_3 does not matter.
For

$$P_i = Q_1 + Q_2 + Q_3$$
$$P_j = Q_1 + Q_2 + Q_4$$
$$\rho_{ij} = 2/3$$

given that Q_1, Q_2, Q_3, Q_4 are pairwise state of knowledge independent and of equal variance.
If k_c is the number of uncertainty contributors shared by P_i and P_j while k_{si} is the number of uncertainty contributors specific to P_i and k_{sj} is the number of those specific to P_j and the contributors are pairwise state of knowledge independent and their subjective probability distributions are of equal variance then, irrespective of their types of subjective probability distribution,

$$\rho_{ij} = c_{ij}/\left[\left(k_c + k_{si}\right)\left(k_c + k_{sj}\right)\right]^{1/2} \tag{3.34}$$

$$c_{ij} = \sum_{m=1}^{k_c} c_m$$

with $c_m = -1$ if the shared contributor is subtracted for either P_i or P_j only and +1 otherwise

For example, if all shared contributors are of positive sign then

k_c	$k_{si} = k_{sj}$	ρ_{ij}
0	1	0
1	1	1/2
1	2	1/3
2	1	2/3
1	3	1/4
1	0	1

e) Depending on the distributions specified as state of knowledge expressions for P_i and P_j not every value from [−1, +1] can be specified as value of the correlation coefficient ρ_{ij}. This follows from solving Eq. (3.35) for ρ_Z (it has to satisfy $-1 \leq \rho_Z \leq +1$) (Krzykacz 1993). The cumulative distribution functions F_i and F_j are assumed to be strictly monotone.

$$\begin{aligned}\rho_{ij} &= \mathrm{Cov}\{P_i, P_j\} / \left(\mathrm{Var}\{P_i\}\mathrm{Var}\{P_j\}\right)^{\frac{1}{2}} = \left(\mathrm{E}\{P_i P_j\} - \mathrm{E}\{P_i\}\mathrm{E}\{P_j\}\right) \\ &\quad / \left(\mathrm{Var}\{P_i\}\mathrm{Var}\{P_j\}\right)^{\frac{1}{2}} \\ &= \left[\int\int F_i^{-1}(\phi(z_i)) F_j^{-1}\left(\phi(z_j)\right) \mathrm{f}_{\mathrm{ij}}(z_i, z_j; \rho_Z) dz_i dz_j - \mathrm{E}(P_i)\mathrm{E}\{P_j\}\right] \\ &\quad / \left(\mathrm{Var}\{P_i\}\mathrm{Var}\{P_j\}\right)^{\frac{1}{2}}\end{aligned} \tag{3.35}$$

$\Phi(z_i)$ is the cumulative standard normal probability at z_i, $\mathrm{F}^{-1}(\Phi(z_i))$ is the value p_i of P_i with cumulative subjective probability $\Phi(z_i)$, $\mathrm{f}_{\mathrm{ij}}(z_i,z_j; \rho_Z)$ is the bivariate normal probability density at (z_i,z_j) and ρ_Z is its correlation coefficient.

If normal distributions are specified as state of knowledge expressions for P_i and for P_j, any value from [−1, +1] can be specified for the correlation coefficient ρ_{ij}. For instance, if subjective probability distributions of the lognormal type are specified for P_i and P_j with distribution parameters (μ_i, σ_i) and (μ_j, σ_j), then the specified value ρ_{ij} has to satisfy the following inequality

$$-1 \leq \rho_Z = \ln\left\{1 + \rho_{ij}\left[\left(\exp\left(\sigma_i^2\right) - 1\right)\left(\exp\left(\sigma_j^2\right) - 1\right)\right]^{1/2}\right\} / \sigma_i \sigma_j \leq +1.$$

The bivariate normal distribution with the normal distributions of $\ln P_i$ and $\ln P_j$ as marginal distributions and with $\rho(\ln P_i, \ln P_j) = \rho_Z$ is after transformation to P_i and P_j a bivariate distribution with the given lognormal distributions of P_i and P_j as marginal distributions and with the given correlation coefficient ρ_{ij}. It follows from Eq. (3.35) that, depending on σ_i and σ_j of the two lognormal distributions, ρ_{ij} cannot have every value from the interval [−1, +1].

The following relationships between ρ_{ij} and ρ_Z for additional combinations of marginal distributions are also given in (Krzykacz 1993):

Both marginal distributions are arbitrary uniform distributions:

$$\rho_Z = 2 \sin\left(\frac{\rho_{ij}\pi}{6}\right)$$

i.e. every value from [−1, +1] can be specified for ρ_{ij}.
One marginal distribution is an arbitrary normal distribution and one is an arbitrary uniform distribution:

$$\rho_Z = \rho_{ij}\left(\frac{\pi}{3}\right)^{1/2}$$

An arbitrary normal distribution is specified for P_i and an arbitrary lognormal distribution with parameters (μ_j, σ_j) for P_j:

$$\rho_Z = \rho_{ij}\left(\exp\left(\sigma_j^2\right) - 1\right)^{\frac{1}{2}} / \sigma_j$$

f) If there is state of knowledge dependence of P_i with two or more uncertain parameters (P_j, P_k, etc.), an additional requirement is to be observed also in the case of normal distributions, namely: The symmetric matrix of specified correlation coefficients

$$\boldsymbol{R} = \begin{bmatrix} 1 & \rho_{ij} & \rho_{ik} \\ \rho_{ji} & 1 & \rho_{jk} \\ \rho_{ki} & \rho_{kj} & 1 \end{bmatrix}$$

must be positive definite in order to be a correlation matrix.
In general: If $\boldsymbol{R}$ is a $K \times K$ symmetric matrix of specified correlation coefficients, then the product $\boldsymbol{a}'\boldsymbol{R}\boldsymbol{a}$, where $\boldsymbol{a}' = (a_1, \ldots, a_K)$ is a row vector with not all of its components equal to zero, must be greater than zero for $\boldsymbol{R}$ to be a correlation matrix. Reason is that any linear combination of K state of knowledge dependent uncertain parameters with coefficients $(a_1, \ldots, a_K)$ has positive variance. The variance equals

$$\begin{aligned} \mathrm{Var}\left\{\sum_{k=1}^{K} a_k P_k\right\} &= \sum_{k=1}^{K} a_k^2 \mathrm{Var}\{P_k\} + 2\sum_{k=1}^{K} \\ &\quad \times \sum_{j=k+1}^{K} a_j a_k \mathrm{Cov}\{P_j P_k\} \\ &= \sum_{k=1}^{K} \sum_{j=1}^{K} a_j a_k \rho_{jk} \left[\mathrm{Var}\{P_j\}\mathrm{Var}\{P_k\}\right]^{1/2} = \boldsymbol{d}'\boldsymbol{R}\boldsymbol{d} \end{aligned} \quad (3.36)$$

$\boldsymbol{d}'$ is the row vector $(a_1(\mathrm{Var}\{P_1\})^{1/2}, \ldots, a_K(\mathrm{Var}\{P_K\})^{1/2})$.

The likelihood that a matrix of correlation values, specified by expert judgment, will be positive definite, decreases rapidly with growing number K of state of knowledge dependent uncertain parameters.

Figures 3.13, 3.14, 3.15, 3.16, 3.17 and 3.18 show pairs of values sampled at random for the uncertain parameters P_i and P_j according to a joint subjective probability distribution that complies with the specified marginal distributions and with the correlation coefficient provided as a measure of state of knowledge dependence. Each dot has the value sampled for P_i as its ordinate and the value sampled for P_j as its abscissa. The figures are intended to give an impression of the effect that:

- Different correlation coefficients may have on a random sample of N pairs of values given the same marginal subjective probability distributions.
- Different marginal subjective probability distributions may have on a random sample of N pairs of values given the same value for the correlation coefficient.
- The same value for Pearson's and Spearman's (see Sect. 3.5.2.5) correlation coefficient may have on a random sample of N pairs of values given the same marginal subjective probability distributions.

In an uncertainty analysis using Monte Carlo simulation, each of the sampled pairs of values will be used in a replication of the model application (see Chap. 4). The correlation coefficients are specified as population expressions of state of knowledge dependence. The figures show that the sample values of the correlation coefficients (given in the headline of each Figure) differ more or less from the specified population values due to the sampling error. If the sample is required to have exactly the specified correlation, then the latter would need to be defined as a

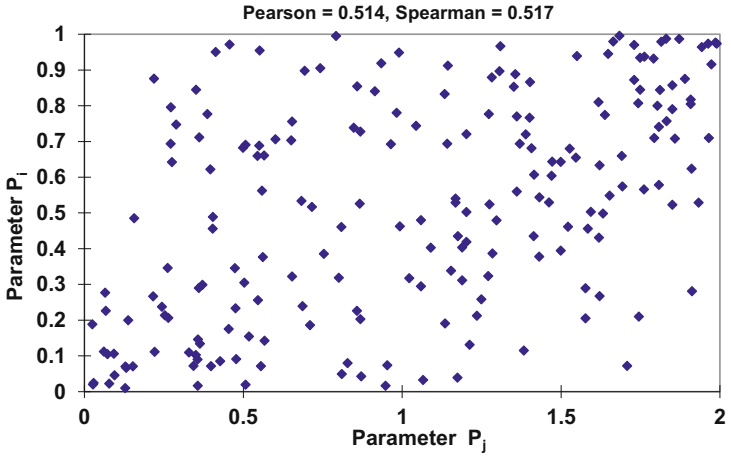


Fig. 3.13 The effect of state of knowledge dependence quantified by Pearson's $\rho_{ij} = 0.5$ for uncertain parameters P_i and P_j with a (0, 1) and a (0, 2) uniform subjective probability distribution as state of knowledge expressions

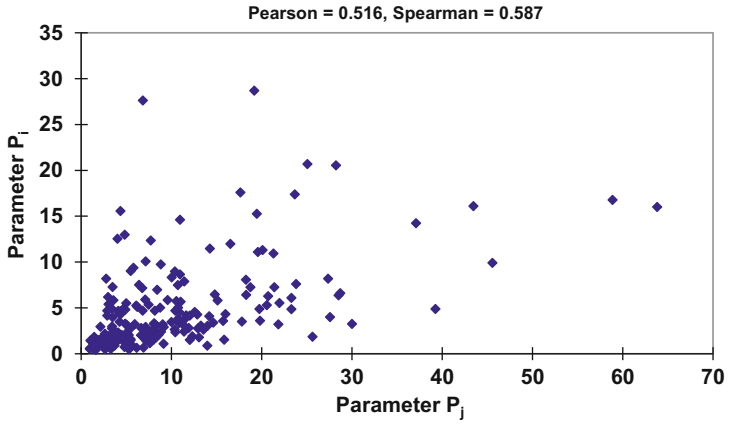


Fig. 3.14 The effect of state of knowledge dependence quantified by Pearson's $\rho_{ij} = 0.5$ for uncertain parameters P_i and P_j with a (1.11, 0.86) and a (1.91, 0.86) lognormal subjective probability distribution as state of knowledge expressions

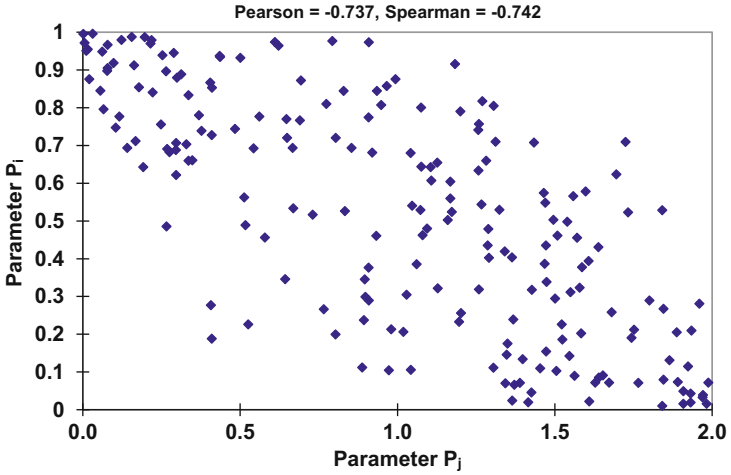


Fig. 3.15 The effect of state of knowledge dependence quantified by Pearson's $\rho_{ij} = -0.7$ for uncertain parameters P_i and P_j with a (0, 1) and a (0, 2) uniform subjective probability distribution as state of knowledge expressions

sample expression of state of knowledge dependence and a sampling procedure such as the one presented in Sect. 4.4.1.14 would need to be followed.

This Figure is blank as the value specified for

Pearson's correlation coefficient is inadmissible for the two marginal

distributions

(i.e. equation 3.35 has no solution for ρ_{ij} *= - 0.7 with* ρ_z *within* [-1, +1]*) .*

Fig. 3.16 The effect of state of knowledge dependence quantified by Pearson's $\rho_{ij} = -0.7$ for uncertain parameters P_i and P_j with a (1.11, 0.86) and a (1.91, 0.86) lognormal subjective probability distribution as state of knowledge expressions

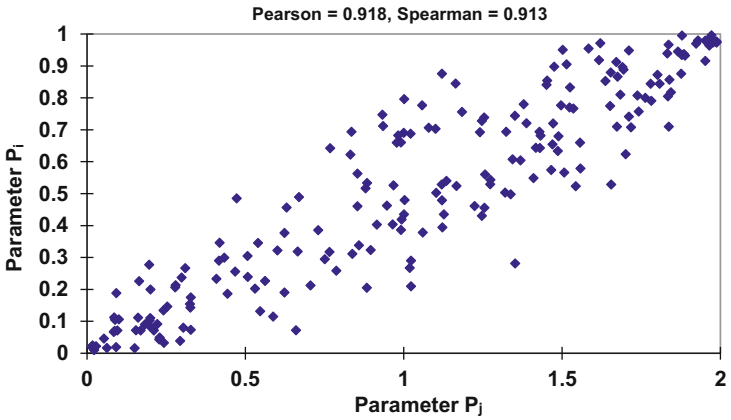


Fig. 3.17 The effect of state of knowledge dependence quantified by Pearson's $\rho_{ij} = 0.9$ for uncertain parameters P_i and P_j with a (0, 1) and (0, 2) uniform subjective probability distribution as state of knowledge expressions

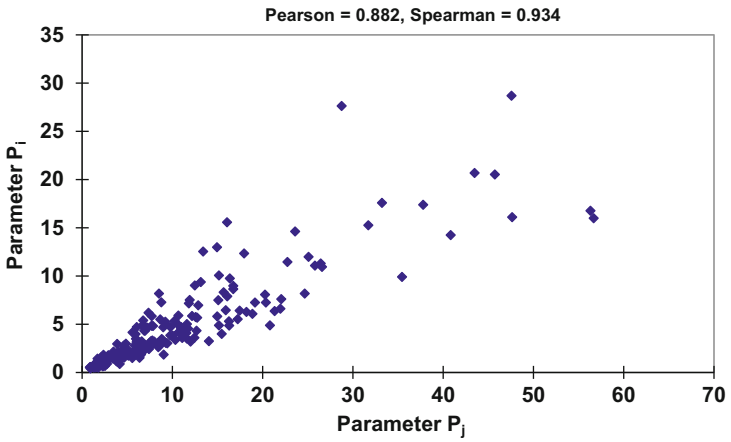


Fig. 3.18 The effect of state of knowledge dependence quantified by Pearson's $\rho_{ij} = 0.9$ for uncertain parameters P_i and P_j with a (1.11, 0.86) and a (1.91, 0.86) lognormal subjective probability distribution as state of knowledge expressions

Summary:
While it is tempting to express state of knowledge dependence by a single number, such as the correlation coefficient, there are limitations to the usefulness of this approach when using Pearson's correlation coefficient ρ.

1. Pearson's correlation coefficient quantifies the pairwise state of knowledge dependence only in so far as it can be expressed by a least squares linear (approximate) relationship of the value thought to be true for one uncertain parameter with the value thought to be true for the other parameter. There are many joint subjective probability distributions with the same marginal distributions and the same value of the correlation coefficient, i.e. the joint subjective probability distribution is not uniquely specified by the marginal distributions and the correlation coefficient except in the case of a bivariate normal distribution.
2. How would one justify the specified value of the correlation coefficient? In other words: Which reasons could one give for the assumed extent of linear relationship between the two states of knowledge?

   A representation of the uncertain parameters by sums of shared and unshared state of knowledge independent uncertainty contributors of equal variance leads to a correlation coefficient that would then be justified by this representation (see Eq. 3.34).

   Quite frequently, a distribution that summarizes stochastic variability in a population is used as quantitative expression of the state of knowledge for an epistemic uncertainty. Given the probability distributions of two stochastic variables X_i and X_j and assuming that the unknown true values of the uncertain parameters P_i and P_j can be thought of as sampled at random from their population of values, then it is quite natural to use the probability distributions of X_i and X_j as subjective probability distributions (i.e. as state of knowledge expressions) for P_i and P_j unless additional knowledge suggests otherwise. Any population expression for the stochastic dependence between X_i and X_j becomes then the population expression for the state of knowledge dependence between P_i and P_j.
3. The range of correlation coefficient values that may be specified for P_i and P_j depends on the types of marginal distributions chosen to quantify the state of knowledge of each, i.e. the specified correlation coefficient must satisfy Eq. (3.35) for a value ρ_Z from $[-1, +1]$. If both marginal distributions are of the normal type, then any value from $[-1, +1]$ is admissible as correlation coefficient for P_i and P_j.
4. If there is state of knowledge dependence among several uncertain parameters and correlation coefficients are specified for all affected pairwise combinations, then the coefficients are elements of a symmetric matrix **R**

$$\boldsymbol{R} = \begin{bmatrix} 1 & r_{12} & \ldots & r_{1M} \\ r_{21} & 1 & \ldots & r_{2M} \\ . & & & \\ . & & & \\ . & & & \\ r_{M1} & \ldots & & 1 \end{bmatrix}$$

This matrix is a correlation matrix only if it is positive definite. The elements of the correlation matrix $\boldsymbol{R}$ are the variances and co-variances of the standardized uncertain parameters $P_j^* = \left(P_j - \mathrm{E}\{P_J\}\right)/\left(\mathrm{Var}\{P_j\}\right)^{1/2}$, $j = 1, \ldots, M$. Any linear combination of the P_j^* with coefficients a_j, $j = 1, \ldots, M$ (not all $a_j = 0$) has positive variance, namely

$Var\left\{\sum_{j=1}^{M} a_j P_j^*\right\} = \boldsymbol{a}'\boldsymbol{R}\boldsymbol{a} > 0, \boldsymbol{a} > \boldsymbol{0}$, where $\boldsymbol{a}$ is the column vector of coefficients of the linear combination. This property of the correlation matrix $\boldsymbol{R}$ is called positive definiteness. The correlation coefficients quantifying state of knowledge dependence are often specified by expert judgment. The chance of specifying the $M(M-1)/2$ elements ρ_{ij}, $j > i$, of the symmetric matrix $\boldsymbol{R}$ such that the matrix is positive definite decreases rapidly with increasing M. If the matrix $\boldsymbol{R}$ is not positive definite some of the non-zero elements must be corrected and/or some of the zero elements may need to be non-zero. There is unfortunately no guidance as to which elements of $\boldsymbol{R}$ need to be corrected and by how much unless additional information is provided. Section 3.5.2.8 presents a method of expressing state of knowledge dependence quantitatively without the direct specification of correlation coefficients. The matrix R that results from the application of this method is guaranteed to be positive definite.

3.5.2.5 Spearman's (Rank) Correlation Coefficient

Using Spearman's instead of Pearson's correlation coefficient for state of knowledge dependence quantification improves on point 2 and eliminates the limitation mentioned under point 3 of the summary given above for Pearson's correlation coefficient.

Let the state of knowledge of the uncertain parameter P_j be expressed by a subjective probability density function $\mathrm{f_j}$ with cumulative subjective probability function $\mathrm{F_j}$. If $\mathrm{F_j}$ is continuous and strictly monotone on $[p_{j,min}, p_{j,max}]$, then the subjective probability $G_j = \mathrm{F_j}(P_j)$ varies according to a uniform distribution on [0,1]. To show this, let $\mathrm{v_j}(g_j)$ be the density function of G_j then

$$\mathrm{v_j}(g_j) = \mathrm{f_j}\left(\mathrm{F_j^{-1}}(g_j)\right)\mathrm{dF_j^{-1}}(g_j)/\mathrm{d}g_j \qquad \text{with } 0 \le g_j \le 1$$

and

$$\begin{aligned}\int_0^{g_j} \mathrm{v_j}\left(g_j'\right)\mathrm{d}g_j' &= \int_{p_{j,min}}^{p_j} \mathrm{f_j}\left(\mathrm{F_j^{-1}}\left(g_j'\right)\right)\mathrm{dF_j^{-1}}\left(g_j'\right) \qquad \text{with} \quad p_j = F_j^{-1}(g_j) \\ &= \mathrm{F_j}\left(\mathrm{F_j^{-1}}(g_j)\right) - \mathrm{F_j}(p_{j,min}) = g_j\end{aligned} \tag{3.37}$$

G_j is the probability integral transformed P_j (see Fig. 3.19),
i.e. $G_j = \mathrm{F_j}(P_j)$ and $P_j = \mathrm{F_j}^{-1}(G_j)$.
With

$$G_i = \mathrm{F_i}(P_i)$$
$$G_j = \mathrm{F_j}(P_j)$$

Spearman's correlation coefficient $\rho^{\mathrm{S}}(P_i, P_j)$ is defined as

$$\rho_{ij}^{S} = \rho^{\mathrm{S}}(P_i, P_j) = \rho(G_i, G_j) \tag{3.38}$$

where ρ denotes Pearson's correlation coefficient. In other words: Spearman's correlation coefficient equals Pearson's correlation coefficient of the probability integral transformed pair of uncertain parameters.

Since G_i and G_j are distributed according to a uniform distribution over [0, 1]

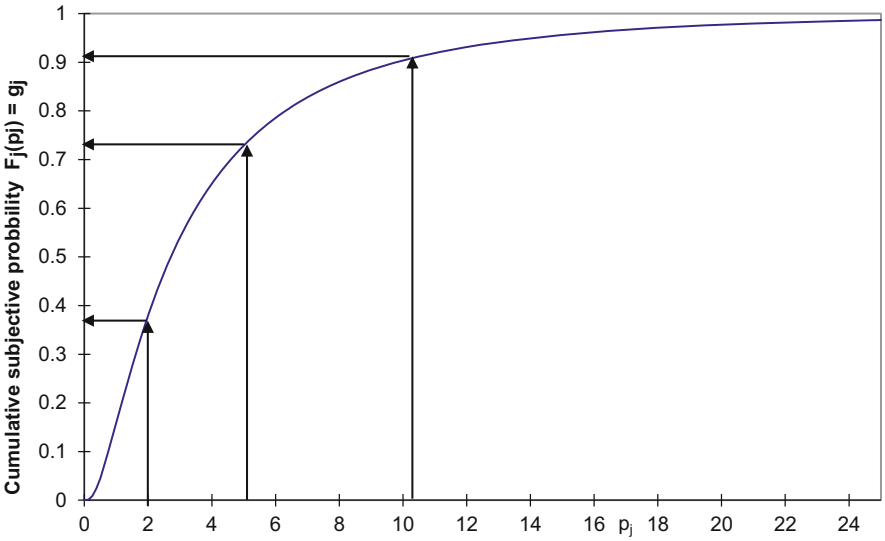


Fig. 3.19 Illustration of the probability integral transformation of P_j

$$\rho^{S}(P_i, P_j) = \rho(G_i, G_j) = 12E\{(G_i - 0.5)(G_j - 0.5)\}$$

1/12 is the variance of the uniform distribution over [0, 1].

If the $u\%$ quantile of the subjective probability distribution for P_j is thought to be the true value of P_j whenever the $u\%$ (or $(100 - u)\%$) quantile of the subjective probability distribution for P_i is thought to be the true value of P_i, then Spearman's correlation coefficient is an appropriate measure of the corresponding state of knowledge dependence and its value is +1 (−1). $\rho^{S}(P_i, P_j) = +1\ (-1)$ indicates complete positive (negative) state of knowledge dependence.

G_i and G_j are symmetrically distributed on [0,1] and therefore $\rho(G_i,G_j)$ $(= \rho^{S}(P_i, P_j))$ can assume any value from [−1, +1] irrespective of the type of marginal distribution function specified for the uncertain parameters.

The following properties need to be observed when specifying values for $\rho^{S}(.,.)$:

a) $\rho^{S}(.,.)$ can only assume values from the interval [−1, +1].
b) If the pair of uncertain data P_i and P_j are judged to be state of knowledge independent, then $\rho(G_i, G_j) = 0$ and therefore $\rho_{ij}^{S} = 0$ is the appropriate choice.

   However, specifying $\rho_{ij}^{S} = 0$ does not necessarily imply state of knowledge independence. It only says that

$$\mathrm{Cov}\{G_i, G_j\} = \mathrm{E}\{(\mathrm{F_i}(P_i) - 0.5)(\mathrm{F_j}(P_j) - 0.5)\} = 0.$$

   $\rho_{ij}^{S} = 0$ is a necessary but not a sufficient condition for state of knowledge independence.

c) If the value thought to be true for P_j is suspected to be a strictly monotone increasing (or decreasing) function of the value thought to be true for P_i, then $\rho_{ij}^{S} = +1\ (-1)$.
d) Values other than 0, +1 or −1, specified for ρ_{ij}^{S}, are quantitative judgments about the extent to which a strictly monotone relationship is suspected to exist between the values thought to be true for P_i and P_j.
e) An estimate r_{ij}^{S} of ρ_{ij}^{S} is obtained from a random sample $\{(p_{i,n}, p_{j,n}))|n = 1, \ldots, N\}$ as

$$r_{ij}^{S} = \sum_{n=1}^{N} [\mathrm{F_i}(p_{i,n}) - 0.5][\mathrm{F_j}(p_{j,n}) - 0.5] \Big/ \left\{ \left(\sum_{n=1}^{N} [\mathrm{F_i}(p_{i,n}) - 0.5]^2 \right) \left(\sum_{n=1}^{N} [\mathrm{F_j}(p_{j,n}) - 0.5]^2 \right) \right\}^{1/2}$$

and observing that the standard deviation of a [0,1] uniform distribution is $(1/12)^{1/2}$

$$r_{ij}^{S} = \left[\left(\frac{1}{N}\right)\sum_{n=1}^{N} \mathrm{F_i}(p_{i,n})\mathrm{F_j}(p_{j,n}) - 0.25\right]12.$$

Following the least squares linear approximation interpretation given for Pearson's ρ, Spearman's ρ_{ij}^{S} quantifies the suspected state of knowledge dependence between P_j and P_i in so far as the dependence can be expressed by a least squares linear approximation of $\mathrm{F_j}(P_j)$ in $\mathrm{F_i}(P_i)$.

However, if only the set $\{(p_{i,n}, p_{j,n}) \mid n = 1, \ldots, N\}$ of possibly true pairs of values of P_i and P_j is known but not the probability integral transforms of the individual values, then rank transformed values are used as an empirical approximation to the probability integral transformed values times N. The rank transformation arranges the sample values in each of the two sets of random values $\{(p_{i,n}) \mid n = 1, \ldots, N\}$ and $\{(p_{j,n}) \mid n = 1, \ldots, N\}$ in ascending order and assigns an ordinal number (the rank among the N sample values) to each.

Ranks are integer numbers with rank 1 given to the smallest value and rank N given to the largest value in the sample. Exception: Equal sample values receive the arithmetic mean of the ranks they would have been given had they been only slightly different.

Example:

n	p_i	$\mathrm{rk}(p_i)$	p_j	$\mathrm{rk}(p_j)$
1	0.76	4	12.16	10
2	0.01	1	8.05	6
3	1.34	7	6.03	4
4	2.46	10	9.16	7
5	0.05	2	11.73	9
6	0.56	3	5.95	3
7	1.03	6	7.03	5
8	2.05	9	10.25	8
9	1.76	8	1.68	1
10	0.95	5	2.35	2

r_{ij}^{S} is then computed as

$$r_{ij}^{S} = \left[\sum_{n=1}^{N} \mathrm{rk}(p_{i,n})\mathrm{rk}(p_{j,n}) - N\left(\frac{N+1}{2}\right)^2\right] / \left\{\left[\sum_{n=1}^{N} \left(\mathrm{rk}(p_{i,n})\right)^2 - N\left(\frac{N+1}{2}\right)^2\right]^{\frac{1}{2}}\left[\sum_{n=1}^{N} \left(\mathrm{rk}(p_{j,n})\right)^2 - N\left(\frac{N+1}{2}\right)^2\right]^{\frac{1}{2}}\right\}$$

where $\mathrm{rk}(p_{i,\,n})$ and $\mathrm{rk}(p_{j,\,n})$ are the ranks of the n-th sampled values of P_i and P_j among the bivariate sample of size N and $(N+1)/2$ is the arithmetic mean value of the N ranks.

For the case of no ties (Spearman 1904)

$$r_{ij}^{S} = \left\{ \sum_{n=1}^{N} \left[\mathrm{rk}\left(p_{i,n}\right) - (N+1)/2 \right] \left[\mathrm{rk}\left(p_{j,n}\right) - (N+1)/2 \right] \right\} / \left[N\left(N^2 - 1\right)/12 \right].$$

An equivalent but computationally easier form (Conover 1980) is

$$r_{ij}^{S} = 1 - \left[6 \sum_{n=1}^{N} \left(\mathrm{rk}\left(p_{i,n}\right) - \mathrm{rk}\left(p_{j,n}\right) \right)^2 \right] / \left(N\left(N^2 - 1\right) \right).$$

f) From Eq. (3.38), it is clear that a symmetric matrix of specified pairwise Spearman correlation values is a correlation matrix only if it is positive definite.

In an uncertainty analysis, using Monte Carlo simulation, each of the sampled pairs of values will be used in a replication of the computer model application (see Chap. 4). The Spearman's (rank) correlation coefficients are specified as population expressions of state of knowledge dependence. Figures 3.20, 3.21 and 3.22 show that the sample values of the correlation coefficients (given in the headline of each Figure) differ more or less from the specified population values due to the sampling error. If the sample is required to have exactly the specified rank correlation, then the latter would need to be defined as a sample expression of state of knowledge dependence and a sampling procedure, such as the one presented in Sect. 4.4.1.14, would need to be followed.

For uniform distributions $F_i(P_i)$ and $F_j(P_j)$

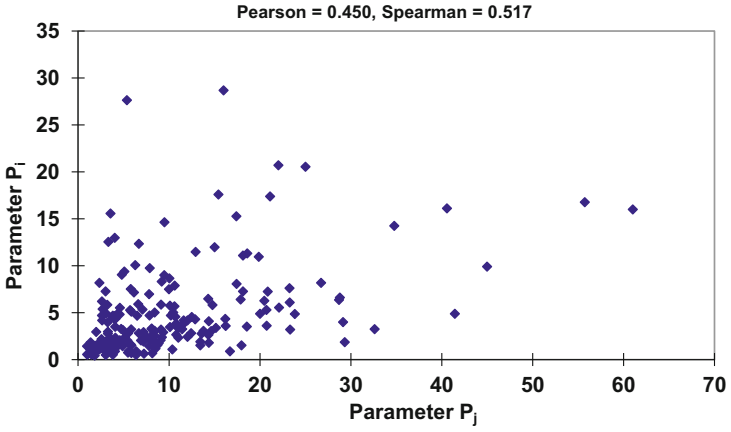


Fig. 3.20 The effect of state of knowledge dependence quantified by Spearman's $\rho_{ij}^{S} = 0.5$ for uncertain parameters P_i and P_j with a (1.11, 0.86) and a (1.91, 0.86) lognormal subjective probability distribution as state of knowledge expressions

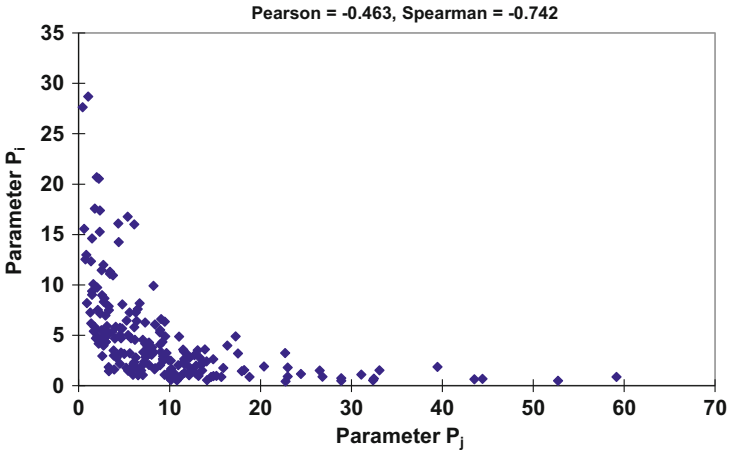


Fig. 3.21 The effect of state of knowledge dependence quantified by Spearman's $\rho_{ij}^{S} = -0.7$ for uncertain parameters P_i and P_j with a (1.11, 0.86) and a (1.91, 0.86) lognormal subjective probability distribution as state of knowledge expressions

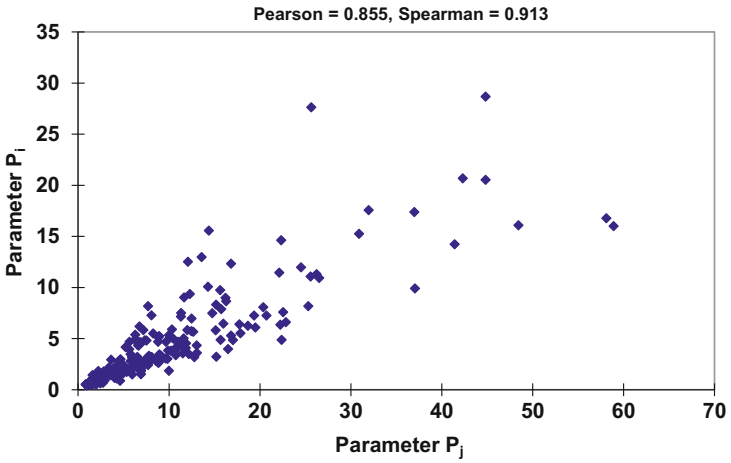


Fig. 3.22 The effect of state of knowledge dependence quantified by Spearman's $\rho_{ij}^{S} = 0.9$ for uncertain parameters P_i and P_j with a (1.11, 0.86) and a (1.91, 0.86) lognormal subjective probability distribution as state of knowledge expressions

$$G_i = \left(P_i - p_{i,min}\right) / \left(p_{i,max} - p_{i,min}\right)$$
$$G_j = \left(P_j - p_{j,min}\right) / \left(p_{j,max} - p_{j,min}\right).$$

It has been shown in Sect. 3.5.2.4 that Pearson's correlation coefficient is unchanged by linear transformations and therefore

$$\rho(P_i, P_j) = \rho(G_i, G_j) = \rho(\mathrm{F_i}(P_i), \mathrm{F_j}(P_j)) = \rho^{\mathrm{S}}(P_i, P_j).$$

Consequently, the scatterplots for cases with uniform marginal distributions for P_i and P_j and ρ_{ij}^S specified are essentially the same as shown for the same value of Pearson's ρ_{ij} in Sect. 3.5.2.4.

Comparison of Pearson's ρ and Spearman's ρ^S

1. Spearman's correlation coefficient quantifies the pairwise state of knowledge dependence only in so far as it can be expressed by a least squares linear approximation of the probability integral transform of one uncertain parameter in the probability integral transform of the other parameter. Since the linear approximation is operating with probability integral transforms instead of the parameter values (as in the case of Pearson's ρ), it quantifies not only the extent of a linear relationship but also of nonlinear strictly monotone relationships.

   There are many joint subjective probability distributions with the same marginal distributions and the same value of Spearman's correlation coefficient. The marginal and Spearman's ρ^S do not uniquely specify the joint subjective probability distribution just as is the case with Pearson's ρ.
2. How would one justify the specified value of the correlation coefficient ρ^S? In other words: Which reasons could one give for the assumed extent of linear relationship between the two states of knowledge? Instead of specifying a value for Spearman's ρ_{ij}^S directly, one could try to give the degree of belief for the true value of P_j to lie below the median value $\mathrm{F_j}(p_{j,\ 50\%}) = 0.5$ of the subjective probability distribution specified for P_j, under the condition that the true value of P_i lies below the median value $p_{i,\ 50\%}$ of the subjective probability distribution specified for P_i. The value given for $\mathrm{sw}(P_j < p_{j,\ 50\%} \mid P_i < p_{i,\ 50\%})$ may then be used to find the corresponding value of ρ_{ij}^S using the relationships given in Sect. 3.6.3.
3. Different to Pearson's ρ, the range of values that may be specified for Spearman's ρ^S is always $[-1, +1]$ irrespective of the type of the subjective probability distribution specified for each of the uncertain parameters.
4. Since $\rho^{\mathrm{S}}(P_i, P_j) = \rho(\mathrm{F_i}(P_i), \mathrm{F_j}(P_j))$, the matrix of specified Spearman's correlation coefficients needs to be positive definite just as the matrix of specified Pearson's correlation coefficients, and all remarks made under point 4 of the summary of Sect. 3.5.2.4 apply here accordingly.
5. If the joint subjective probability distribution specified for P_i and P_j is normal, then the following relationship between Pearson's ρ and Spearman's ρ^S applies (Kruskal 1958)

$$\rho(P_i, P_j) = 2 \sin\left[\left(\frac{\pi}{6}\right)\rho^{\mathrm{S}}(P_i, P_j)\right] \tag{3.39}$$

where the argument of the sinus function is in radians (multiply by $360/2\pi$ to have degrees).

6. Spearman's ρ^S is invariant under nonlinear strictly monotone transformations of P_i and/or P_j as these transformations do not change the cumulative subjective probabilities.

3.5.2.6 Unknown Functional Relationship

State of knowledge dependence needs to be quantified also in the case of an unknown functional relationship h_j that is judged to exist between the values thought to be true for P_j and P_i (see Fig. 3.23)

$$P_j = h_j(P_i).$$

If h_j is strictly monotone increasing (decreasing), then $\rho^S(P_i, P_j) = +1$ (–1) and the same applies to $\rho(G_i, G_j)$ since $G_j = G_i$ ($G_j = 1 - G_i$). Specifying that the unknown functional relationship h_j is strictly monotone increasing or decreasing indicates complete state of knowledge dependence and suffices therefore as a population expression for state of knowledge dependence quantification.

For a strictly monotone increasing function $P_j = h_j(P_i)$, the following applies

$$F_j(p_j) = F_j(h_j(p_i)) = F_i(p_i) \text{ and therefore } \rho^S(P_i, P_j) = 1.$$

In the case of a strictly monotone decreasing function $P_j = h_j(P_i)$, it follows

$$F_j(p_j) = F_j(h_j(p_i)) = 1 - F_i(p_i) \text{ and therefore } \rho^S(P_i, P_j) = -1.$$

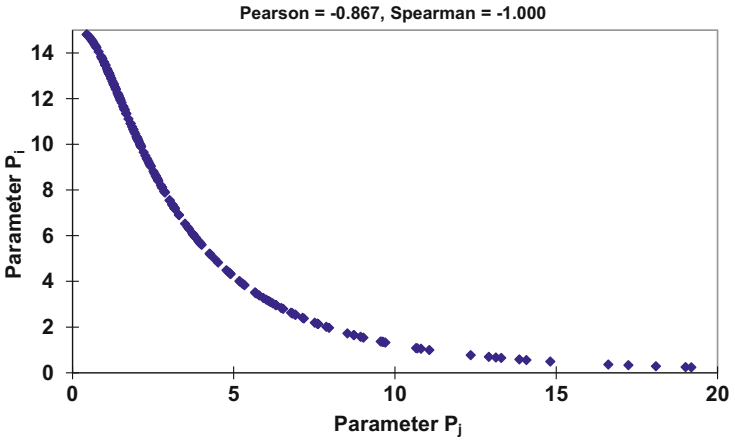


Fig. 3.23 The effect of negative complete state of knowledge dependence (i.e. due to a suspected strictly monotone decreasing relationship) specified for two uncertain parameters with their state of knowledge expressed by a (0, 15) uniform subjective probability distribution for P_i and a (1.11, 0.86) lognormal subjective probability distribution for P_j

3.5.2.7 Copulas

Section 3.5.2.5 introduced the probability integral transformation for uncertain parameters. The use of probability integral transformed values leads directly to the concept of copulas.

If the state of knowledge of the uncertain parameters P_i and P_j is quantified by continuous marginal subjective probability distributions with the joint subjective probability distribution F_{ij}, then the probability integral transformed parameters G_i and G_j follow a joint subjective probability distribution C_{ij} over the unit square with $g_i = F_i(p_i)$ and $g_j = F_j(p_j)$. C_{ij} together with the continuous marginal distributions F_i and F_j fully determines the joint distribution for P_i and P_j since

$$C_{ij}(g_i, g_j) = C_{ij}(F_i(p_i), F_j(p_j)) = F_{ij}(p_i, p_j). \tag{3.40}$$

C_{ij} is called the copula distribution that joins P_i and P_j. It contains all the information about the state of knowledge dependence of P_i and P_j needed in order to know their joint distribution.

The density function of the copula is connected to the joint density of P_i and P_j via

$$\begin{aligned} f_{ij}(p_i, p_j) &= \partial^2 F_{ij}(p_i, p_j)/\partial p_i \partial p_j \\ &= \left[\partial^2 C_{ij}(F_i(p_i), F_j(p_j))/\partial F_i(p_i), \partial F_j(p_j)\right][dF_i(p_i)/dp_i]\left[dF_j(p_j)/dp_j\right] \\ &= c_{ij}(F_i(p_i), F_j(p_j)) f_i(p_i) f_j(p_j) \end{aligned}$$

and therefore

$$c_{ij}(F_i(p_i), F_j(p_j)) = f_{ij}(p_i, p_j)/f_i(p_i) f_j(p_j). \tag{3.41}$$

Equation (3.41) shows that the copula density is, at every point (p_i, p_j), the factor by which the joint density in the case of state of knowledge independence is to be multiplied to obtain the joint density in the case of state of knowledge dependence. From this it follows for the conditional density

$$f_{j|i}(p_j|p_i) = f_{ij}(p_i, p_j)/f_i(p_i) = c_{ij}(F_i(p_i), F_j(p_j)) f_j(p_j). \tag{3.42}$$

Any multivariate subjective probability distribution

$$F_{1\ldots K}(p_1, \ldots, p_K) = \mathrm{sw}(P_1 \le p_1, \ldots, P_K \le p_K)$$

with continuous marginals

$$F_k(p_k) = g_k = \mathrm{sw}(P_k \le p_k)$$

can be written as $C_{1\ldots K}(g_1, \ldots, g_K)$ with $C_{1\ldots K}$ as a unique copula (Sklar 1959; Nelsen 2006). However, there are many families of copulas (Nelsen 2006) that can be used to join $P_1, \ldots, P_K$ as long as only their marginal distributions and their

pairwise (rank) correlation coefficients are given (i.e. the K-variate subjective probability distribution is not uniquely specified).

In uncertainty analysis, the available (expert elicited) information often only consists of the marginal distributions and pairwise Pearson's ρ or Spearman's ρ^S. In this case, the normal copula distribution $C_{1\ldots K}(g_1,\ldots,g_K) = \phi(\phi^{-1}(g_1),\ldots,\phi^{-1}(g_K)); \boldsymbol{R_Z})$ is a popular choice. It is the multivariate normal distribution with correlation matrix $\boldsymbol{R_Z}$, and ϕ^{-1} is the inverse of the standard normal distribution function.

With $z_i = \phi^{-1}(F_i(p_i))$, $z_j = \phi^{-1}(F_j(p_j))$ and $\rho^S(P_i,P_j)$given, $\rho_Z = \rho(Z_i,Z_j)$ is obtained from

$$\rho_Z = 2\sin\left[\left(\frac{\pi}{6}\right)\rho^S(P_i,P_j)\right] \text{ since } \rho^S(Z_i,Z_j) = \rho^S(P_i,P_j).$$

$$\boldsymbol{R_Z} = \begin{bmatrix} 1 & \rho_Z \\ \rho_Z & 1 \end{bmatrix}$$

and the density function of the normal copula distribution

$$C_{ij}(F_i(p_i),F_j(p_j)) = \phi(\phi^{-1}(F_i(p_i)),\phi^{-1}(F_j(p_j)));\boldsymbol{R_Z}) = \phi(z_i,z_j;\boldsymbol{R_Z})$$

is obtained as

$$c_{ij}(F_i(p_i),F_j(p_j);\boldsymbol{R_Z}) = \left[\frac{\partial^2\phi(z_i,z_j;\boldsymbol{R_Z})}{\partial z_i \partial z_j}\right]\left[\frac{dz_i dz_j}{dF_i(p_i)dF_j(p_j)}\right]$$

which gives with

$$\frac{dF(p)}{dz} = \frac{d\phi(z)}{dz} = \varphi(z) = (2\pi)^{-\frac{1}{2}}\exp\left(-\frac{z^2}{2}\right)$$

$$c_{ij}(F_i(p_i),F_j(p_j);\boldsymbol{R_Z}) = \left[(2\pi)^{-1}|\boldsymbol{R_Z}|^{-\frac{1}{2}}\exp\left(\frac{-\boldsymbol{z'R_Z}^{-1}\boldsymbol{z}}{2}\right)\right]/\left[\left(2\pi\right)^{-1}\exp\left(\frac{-\boldsymbol{z'z}}{2}\right)\right]$$

$$= |\boldsymbol{R_Z}|^{-\frac{1}{2}}\exp\left(\frac{-\boldsymbol{z'}(\boldsymbol{R_Z}^{-1}-\boldsymbol{I})\boldsymbol{z}}{2}\right). \tag{3.43}$$

With $dF(p)/dp = f(p)$

$$\begin{aligned} \mathrm{f_{ij}}(p_i, p_j) &= \mathrm{c_{ij}}\left(\mathrm{F_i}(p_i), \mathrm{F_j}(p_j); \boldsymbol{R_Z}\right) \mathrm{f_i}(p_i) \mathrm{f_j}(p_j) \\ &= |\boldsymbol{R_Z}|^{-\frac{1}{2}} \exp\left(\frac{-\boldsymbol{z}'\left(\boldsymbol{R_Z}^{-1} - \boldsymbol{I}\right)\boldsymbol{z}}{2}\right) \mathrm{f_i}(p_i) \mathrm{f_j}(p_j) \end{aligned} \tag{3.44}$$

and

$$\mathrm{f_{j|i}}(p_j|p_i) = |\boldsymbol{R_Z}|^{-\frac{1}{2}} \exp\left(\frac{-\boldsymbol{z}'\left(\boldsymbol{R_Z}^{-1} - \boldsymbol{I}\right)\boldsymbol{z}}{2}\right) \mathrm{f_j}(p_j).$$

If the state of knowledge of (P_i, P_j) may be expressed by a joint normal distribution, then the marginal distributions and $\rho_Z = \rho(Z_i, Z_j)$ are sufficient to uniquely specify the joint density function as

$$\mathrm{f_{ij}}(p_i, p_j) = \mathrm{c_{ij}}\left(\mathrm{F_i}(p_i), \mathrm{F_j}(p_j); \boldsymbol{R_Z}\right) \left[\frac{\mathrm{dF_i}(p_i)\mathrm{dF_j}(p_j)}{\mathrm{d}z_i \mathrm{d}z_j}\right] \left[\frac{\mathrm{d}z_i \mathrm{d}z_j}{\mathrm{d}p_i \mathrm{d}p_j}\right]$$

with

$$\boldsymbol{R_Z}^{-1} = \begin{bmatrix} 1/(1-\rho_Z^2) & -\rho_Z/(1-\rho_Z^2) \\ -\rho_Z/(1-\rho_Z^2) & 1/(1-\rho_Z^2) \end{bmatrix}$$

$$\boldsymbol{z}' = \left[\frac{p_i - \mu_i}{\sigma_i}, \frac{p_j - \mu_j}{\sigma_j}\right], |\boldsymbol{R_Z}| = 1 - \rho_Z^2$$

and since

$$\frac{\mathrm{dF}(p)}{\mathrm{d}z} = (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{z^2}{2}\right)$$

and $\left(\frac{\mathrm{d}z}{\mathrm{d}p}\right) = 1/\sigma$

$$\mathrm{f_{ij}}(p_i, p_j) = \left(2\pi\sigma_i\sigma_j\right)^{-1} |\boldsymbol{R_Z}|^{-\frac{1}{2}} \exp\left(\frac{-\boldsymbol{z}'\boldsymbol{R_Z}^{-1}\boldsymbol{z}}{2}\right)$$

If $\mathrm{F_{ij}}(P_i, P_j)$ is a bivariate normal distribution function with Pearson's correlation coefficient ρ_Z then

$$\mathrm{C_{ij}}(g_i, g_j) = \phi\left(\phi^{-1}(g_i), \phi^{-1}(g_j); \boldsymbol{R_Z}\right)$$

is the unique copula with $\phi^{-1}(g_i) = z_i$, $\phi^{-1}(g_j) = z_j$ and $p_i = z_i\sigma_i + \mu_i = F_i^{-1}(\phi(z_i))$, $p_j = z_j\sigma_j + \mu_j = F_j^{-1}(\phi(z_j))$, $(g_i, g_j) \in (0,1)^2$.

If $F_{ij}(P_i, P_j)$ is a non-normal bivariate distribution function with continuous marginals and correlation coefficient $\rho(P_i, P_j)$, then with

$$z_i = \phi^{-1}(F_i(p_i)), z_j = \phi^{-1}(F_j(p_j))$$
$$C_{ij}(g_i, g_j) = \phi(z_i, z_j; \boldsymbol{R_Z}) = \phi(\phi^{-1}(g_i), \phi^{-1}(g_j); \boldsymbol{R_Z})$$

is a copula of P_i and P_j with correlation coefficient $\rho_Z = \rho(Z_i, Z_j)$. The correlation coefficient ρ_Z needs to be chosen such that

$$F_{ij}(F_i^{-1}(\phi(z_i)), F_j^{-1}(\phi(z_j)))$$

has correlation $\rho(P_i, P_j)$.

The value of ρ_Z is obtained either analytically, using Eq. (3.35), or via iteration. Otherwise, if $\rho^S(P_i, P_j)$ is provided instead of $\rho(P_i, P_j)$, the correlation coefficient ρ_Z needs to be (Kruskal 1958)

$$\rho_Z = 2\sin\left[\left(\frac{\pi}{6}\right)\rho^S(P_i, P_j)\right].$$

Spearman's correlation coefficient $\rho^S(P_i, P_j)$ equals $\rho^S(F_i(P_i), F_j(P_j))$ which in turn equals Pearson's correlation coefficient $\rho(F_i(P_i), F_j(P_j))$ while $\rho(F_i(P_i), F_j(P_j)) \neq \rho(P_i, P_j)$ except for the case where the state of knowledge of both uncertain parameters is expressed by a uniform distribution. On the other hand, $\rho^S(F_i(P_i), F_j(P_j))$ equals $\rho^S(Z_i, Z_j)$ and the latter can be transformed into $\rho(Z_i, Z_j)$ using Eq. (3.39).

Obviously, $C_{ij}(g_i, g_j) = \phi(z_i, z_j; \boldsymbol{R_Z})$ is not the only copula that may be used if instead of F_{ij} only the marginal distributions F_i and F_j are specified together with $\rho^S(P_i, P_j)$. Any bivariate distribution on the unit square with uniform marginal distributions and correlation coefficient $\rho(F_i(P_i), F_j(P_j)) = \rho^S(P_i, P_j)$ is a copula distribution for F_{ij} under these conditions. If only the marginal distributions and the (rank) correlation coefficient are provided as state of knowledge expression for P_i and P_j then, for the purpose of uncertainty analysis, the copula should be chosen such that its information contents, relative to the case of state of knowledge independence, are minimal.

In Kurowicka and Cooke (2006), the relative information of the minimum information copula is given for various rank correlation values and it is compared to that of four other copulas. Among the four copulas compared, Frank's copula is shown to have the smallest increase of the relative information over the minimum information copula. In Schirmacher and Schirmacher (2008), five copulas are compared for different correlation coefficients, namely the Clayton, Frank, Galambos, Gumbel and bivariate normal copulas. The comparison of Frank's copula and the bivariate normal copula via scatterplots seems to suggest that the increase of

the relative information contents over the minimum information copula is even lower for the bivariate normal than for Frank's copula.

The fact, that rank correlations are not changed by the probability integral transformation as well as by the distribution function inversions, suggests (among other reasons given at the end of Sect. 3.5.2.5) to specify Spearman's ρ^S instead of Pearson's ρ when quantifying state of knowledge dependence. In case Pearson's ρ was specified, an iterative approach is often required to find a correlation coefficient ρ_Z of the copula such that after reversion of the probability integral transformation the resulting correlation agrees with the correlation value originally specified for state of knowledge dependence quantification.

If the state of knowledge dependence between P_i and P_j was not quantified by correlation coefficient but by conditional distributions for P_j with parameters that are a function of the value given for P_i, then the density function of the corresponding copula follows from Eq. (3.42) as

$$\mathrm{c_{ij}}\left(\mathrm{F_i}(p_i), F_j(p_j)\right) = \frac{\mathrm{f_{j|i}}(p_j|p_i)}{\mathrm{f_j}(p_j)}.$$

A multivariate joint density function can be specified as the product of the marginal densities and an expression using pairwise copulas with given (conditional) rank correlation coefficients.

A bivariate density function can, for instance, be decomposed as follows:

$$f_{ij}(p_i, p_j) = f_i(p_i)\mathrm{f_{j|i}}(p_j|p_i)$$

and in the trivariate case

$$f_{ijk}(p_i, p_j, p_k) = f_i(p_i)\mathrm{f_{j|i}}(p_j|p_i)f_{k|ij}(p_k|p_i, p_j)$$

and so forth to higher dimensions.

Using copulas, one gets

$$\mathrm{f_{j|i}}(p_j|p_i) = \mathrm{c_{ij}}\left(\mathrm{F_i}(p_i), \mathrm{F_j}(p_j)\right)\mathrm{f_j}(p_j)$$

and choosing p_j for conditioning

$$\begin{aligned} f_{k|ij}(p_k|p_i, p_j) &= c_{ki|j}\left(F_{i|j}(p_i|p_j), F_{k|j}(p_k|p_j)\right)f_{k|j}(p_k|p_j) \\ f_{ki|j}(p_k, p_i|p_j) &= c_{ki|j}\left(F_{i|j}(p_i|p_j), F_{k|j}(p_k|p_j)\right)f_{k|j}(p_k|p_j)f_{i|j}(p_i|p_j) \\ f_{k|ij}(p_k|p_i, p_j) &= f_{ki|j}(p_k, p_i|p_j)/f_{i|j}(p_i|p_j) \end{aligned} \tag{3.45}$$

which leads to

$$f_{ijk}(p_i,p_j,p_k) = f_i(p_i)\mathrm{c_{ij}}\left(\mathrm{F_i}(p_i),\mathrm{F_j}(p_j)\right)\mathrm{f_j}(p_j)c_{ki|j}\left(F_{i|j}(p_i|p_j),F_{k|j}(p_k|p_j)\right) \times \mathrm{c_{jk}}\left(\mathrm{F_j}(p_j),\mathrm{F_k}(p_k)\right)\mathrm{f_k}(p_k) \tag{3.46}$$

where one copula has to be evaluated at values of conditional distributions. If p_i had been chosen for conditioning in $f_{k\,|\,ij}$, then the corresponding copula densities would have been $c_{kj\,|\,i}$ and c_{ik} instead of $c_{ki\,|\,j}$ and $\mathrm{c_{jk}}$. The decomposition of a quadruple-variate density function involves already six copulas, three of which are to be evaluated at values of conditional distributions.

Graphical means like dependence trees and vines on these trees are in use in order to visualize the decomposition of the multivariate joint density function into copulas and marginal density functions. The trees are made up of nodes that are connected by edges. For instance, in the context of uncertainty analysis, the tree in Fig. 3.24 shows that P_2, P_3 and P_4 each have an uncertainty contributor in common with $\mathrm{P_1}$ so that there is state of knowledge dependence with P_1.

If this common contributor is the same for P_2, P_3, P_4, then these uncertain parameters have an uncertainty contributor in common and are therefore pairwise state of knowledge dependent. If the only uncertainty contributor that P_2, P_3, P_4 have in common is the one they share with P_1, then they are conditionally independent given P_1. If P_2, P_3, P_4 each share a different uncertainty contributor with P_1, then they are state of knowledge dependent with P_1 but are pairwise state of knowledge independent.

In the second degree tree of Fig. 3.25, the edges of the original tree are now nodes connected by conditional edges. There is a choice: The conditional edge 2,4|1 could have been chosen instead of 2,3|1. Finally, there is the third degree tree shown in Fig. 3.26.

From this follows the decomposition of the quadruple-variate density function into

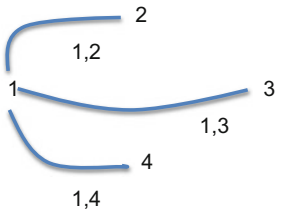


Fig. 3.24 Dependence tree

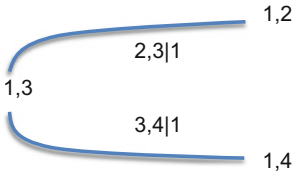


Fig. 3.25 Second degree tree to the dependence tree of Fig. 3.24

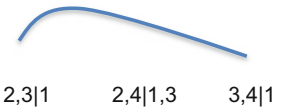


Fig. 3.26 Third degree tree to the dependence tree of Fig. 3.24

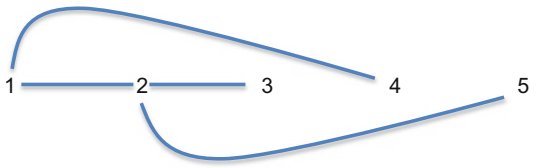


Fig. 3.27 Dependence tree

$$
\begin{aligned}
\mathrm{f}_{1234}(p_1,p_2,p_3,p_4) = \quad & \mathrm{f}_1(p_1) \\
& \mathrm{c}_{12}(\mathrm{F}_2(p_2),\mathrm{F}_1(p_1))\,\mathrm{f}_2(p_2) \\
& \mathrm{c}_{13}(\mathrm{F}_3(p_3),\mathrm{F}_1(p_1))\,\mathrm{f}_3(p_3) \\
& \mathrm{c}_{14}(\mathrm{F}_4(p_4),\mathrm{F}_1(p_1))\,\mathrm{f}_4(p_4) \\
& \mathrm{c}_{23|1}\left(\mathrm{F}_{2|1}(p_2|p_1),\mathrm{F}_{3|1}(p_3|p_1)\right) \\
& \mathrm{c}_{24|1}\left(\mathrm{F}_{2|1}(p_2|p_1),\mathrm{F}_{4|1}(p_4|p_1)\right) \\
& \mathrm{c}_{34|12}\left(\left(\mathrm{F}_{3|12}(p_3|p_1,p_2),\mathrm{F}_{4|12}(p_4|p_1,p_2)\right)\right).
\end{aligned}
$$

The edges of the second and third degree trees are called the vines.

Any dependence tree can be represented by entering the primary edges into a K by K array where K is the number of state of knowledge dependent uncertain parameters, i.e. the nodes of the tree. A tree has branches and no loops. Several branches may start in the same node but only one branch can end in a node. The branches are the primary edges. The start number of the edge serves as the row number of its entry in the array and the end number serves as the column number. The array is symmetric; therefore, it is sufficient to have entries only in the upper triangle (excluding the diagonal), given the start numbers are always chosen lower than the end numbers of the edges. There is only one entry in each column. This leads to the array in Table 3.2.

In order to find the additional parameter pairs that need to have a copula assigned to, one first builds all those secondary entries that can be formed from primary ones with one number in common. This gives the array in Table 3.3:

The tertiary entries, shown in Table 3.4, are built from the secondary ones that have two numbers in common.

Finally, Table 3.5 shows the array supplemented by the entry 45|123 which was built from the tertiary entries with three numbers in common.

The array in Table 3.5 shows the pair copulas needed to construct the multivariate density function according to which a random sample of parameter values would need to be taken in Step 3 of the uncertainty analysis (see Chap. 4).

The (de)composition of the 5-variate density function is therefore

Table 3.2 The edges of the tree in Fig. 3.27 are the primary array entries

	12		14	
		23		25

Table 3.3 The array in Table 3.2 supplemented by the secondary entries (i.e. the edges of the second degree tree)

	12	13|2	14	15|2
		23	24|1	25

Table 3.4 The array in Table 3.3 supplemented by the tertiary entries (i.e. the edges of the third degree tree)

	12	13|2	14	15|2
		23	24|1	25
			34|12	35|12

Table 3.5 The array in Table 3.4 supplemented by the edge of the fourth degree tree

	12	13|2	14	15|2
		23	24|1	25
			34|12	35|12
				45|123

$$
\begin{aligned}
\mathrm{f}_{12345}(p_1,p_2,p_3,p_4,p_5) = \quad & \mathrm{f}_1(p_1)\mathrm{f}_2(p_2)\mathrm{f}_3(p_3)\mathrm{f}_4(p_4)\mathrm{f}_5(p_5) \\
& \mathrm{c}_{12}(\mathrm{F}_1(p_1),\mathrm{F}_2(p_2)) \\
& \mathrm{c}_{13|2}\big(\mathrm{F}_{1|2}(p_1|p_2),\mathrm{F}_{3|2}(p_3|p_2)\big) \\
& \mathrm{c}_{14}(\mathrm{F}_1(p_1),\mathrm{F}_4(p_4)) \\
& \mathrm{c}_{15|2}\big(\mathrm{F}_{1|2}(p_1|p_2),\mathrm{F}_{5|2}(p_5|p_2)\big) \\
& \mathrm{c}_{23}(\mathrm{F}_2(p_2),\mathrm{F}_3(p_3)) \\
& \mathrm{c}_{24|1}\big(\mathrm{F}_{2|1}(p_2|p_1),\mathrm{F}_{4|1}(p_4|p_1)\big) \\
& \mathrm{c}_{25}(\mathrm{F}_2(p_2),\mathrm{F}_5(p_5)) \\
& \quad \mathrm{c}_{34|12}\big(\big(\mathrm{F}_{3|12}(p_3|p_1,p_2),\mathrm{F}_{4|12}(p_4|p_1,p_2)\big) \\
& \quad \mathrm{c}_{35|12}\big(\big(\mathrm{F}_{3|12}(p_3|p_1,p_2),\mathrm{F}_{5|12}(p_5|p_1,p_2)\big) \\
& \quad \mathrm{c}_{45|123}\big(\big(\mathrm{F}_{4|123}(p_4|p_1,p_2,p_3),\mathrm{F}_{5|123}(p_5|p_1,p_2,p_3)\big).
\end{aligned}
\tag{3.47}
$$

$\mathrm{f}_{12345} = \mathrm{f}_1\ \mathrm{f}_{2|1}\ \mathrm{f}_{3|12}\ \mathrm{f}_{4|123}\ \mathrm{f}_{5|1234}$
with
$\mathrm{c}_{12}\ \mathrm{f}_{12}\ /\ (\mathrm{f}_1\mathrm{f}_2) = \mathrm{f}_{2|1}\ /\ \mathrm{f}_2 \Rightarrow \mathbf{f_{2|1}} = \mathrm{c}_{12}\ \mathrm{f}_2$
$\mathrm{c}_{14} = \mathrm{f}_{14}\ /\ (\mathrm{f}_1\mathrm{f}_4) = \mathrm{f}_{4|1}\ /\ \mathrm{f}_4 \Rightarrow \mathbf{f_{4|1}} = \mathrm{c}_{14}\ \mathrm{f}_4$
$\mathrm{c}_{23} = \mathrm{f}_{23}\ /\ (\mathrm{f}_2\mathrm{f}_3) = \mathrm{f}_{3|2}\ /\ \mathrm{f}_3 \Rightarrow \mathbf{f_{3|2}} = \mathrm{c}_{23}\ \mathrm{f}_3$
$\mathrm{c}_{24|1} = \mathrm{f}_{24|1}\ /\ (\mathrm{f}_{2|1}\mathrm{f}_{4|1}) = \mathrm{f}_{4|12}\ /\ \mathrm{f}_{4|1} \Rightarrow \mathbf{f_{4,12}} = \mathrm{c}_{24|1}\ \mathbf{f_{4|1}}$
$\mathrm{c}_{13|2} = \mathrm{f}_{13|2}\ /\ (\mathrm{f}_{1|2}\mathrm{f}_{3|2}) = \mathrm{f}_{3|12}\ /\ \mathrm{f}_{3|2} \Rightarrow \mathbf{f_{3|12}} = \mathrm{c}_{13|2}\ \mathbf{f_{3|2}}$
$\mathrm{c}_{34|12} = \mathrm{f}_{34|12}\ /\ (\mathrm{f}_{3|12}\mathrm{f}_{4|12}) = \mathrm{f}_{4|123}\ /\ \mathrm{f}_{4|12} \Rightarrow \mathbf{f_{4|123}} = \mathrm{c}_{34|12}\ \mathbf{f_{4|12}}$
$\mathrm{c}_{25} = \mathrm{f}_{25}\ /\ (\mathrm{f}_2\mathrm{f}_5) = \mathrm{f}_{5|2}\ /\ \mathrm{f}_5 \Rightarrow \mathbf{f_{5|2}} = \mathrm{c}_{25}\ \mathrm{f}_5$
$\mathrm{c}_{15|2} = \mathrm{f}_{15|2}\ /\ (\mathrm{f}_{1|2}\mathrm{f}_{5|2}) = \mathrm{f}_{5|12}\ /\ \mathrm{f}_{5|2} \Rightarrow \mathbf{f_{5|12}} = \mathrm{c}_{15|2}\ \mathbf{f_{5|2}}$
$\mathrm{c}_{35|12} = \mathrm{f}_{35|12}\ /\ (\mathrm{f}_{3|12}\mathrm{f}_{5|12}) = \mathrm{f}_{5|123}\ /\ \mathrm{f}_{5|12} \Rightarrow \mathbf{f_{5|123}} = \mathrm{c}_{35|12}\ \mathbf{f_{5|12}}$
$\mathrm{c}_{45|123} = \mathrm{f}_{45|123}\ /\ (\mathrm{f}_{4|123}\mathrm{f}_{5|123}) = \mathrm{f}_{5|1234}\ /\ \mathrm{f}_{5|123} \Rightarrow \mathbf{f_{5|1234}} = \mathrm{c}_{45|123}\ \mathbf{f_{5|123}}$

The final array for a so-called C-vine tree (see Fig. 3.28) with five uncertain parameters looks as shown in Table 3.6, where all primary entries are in the top row

The (de)composition of the corresponding 5-variate density function is therefore

$$
\begin{aligned}
\mathrm{f}_{12345}(p_1,p_2,p_3,p_4,p_5) = \quad & \mathrm{f}_1(p_1)\mathrm{f}_2(p_2)\mathrm{f}_3(p_3)\mathrm{f}_4(p_4)\mathrm{f}_5(p_5) \\
& \mathrm{c}_{12}(\mathrm{F}_1(p_1),\mathrm{F}_2(p_2)) \\
& \mathrm{c}_{13}(\mathrm{F}_1(p_1),\mathrm{F}_3(p_3)) \\
& \mathrm{c}_{14}(\mathrm{F}_1(p_1),\mathrm{F}_4(p_4)) \\
& \mathrm{c}_{15}(\mathrm{F}_1(p_1),\mathrm{F}_5(p_5)) \\
& \mathrm{c}_{23|1}\big(\mathrm{F}_{2|1}(p_2|p_1),\mathrm{F}_{3|1}(p_3|p_1)\big) \\
& \mathrm{c}_{24|1}\big(\mathrm{F}_{2|1}(p_2|p_1),\mathrm{F}_{4|1}(p_4|p_1)\big) \\
& \mathrm{c}_{25|1}\big(\mathrm{F}_{2|1}(p_2|p_1),\mathrm{F}_{5|1}(p_5|p_1)\big) \\
& \quad \mathrm{c}_{34|12}\big(\big(\mathrm{F}_{3|12}(p_3|p_1,p_2),\mathrm{F}_{4|12}(p_4|p_1,p_2)\big) \\
& \quad \mathrm{c}_{35|12}\big(\big(\mathrm{F}_{3|12}(p_3|p_1,p_2),\mathrm{F}_{5|12}(p_5|p_1,p_2)\big) \\
& \quad \mathrm{c}_{45|123}\big(\big(\mathrm{F}_{4|123}(p_4|p_1,p_2,p_3),\mathrm{F}_{5|123}(p_5|p_1,p_2,p_3)\big).
\end{aligned}
\tag{3.48}
$$

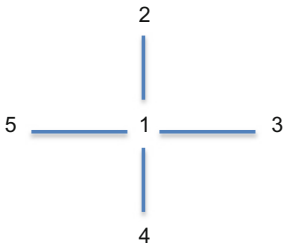


Fig. 3.28 A C-vine dependence tree

Table 3.6 Array to the dependence tree in Fig. 3.28

	12	13	14	15
		23|1	24|1	25|1
			34|12	35|12
				45|123

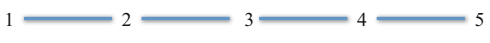


Fig. 3.29 A D-vine dependence tree

For a so-called D-vine tree (see Fig. 3.29), all primary entries of the array in Table 3.7 are in the first upper sub-diagonal and the corresponding (de)composition of the 5-variate density function is

Table 3.7 Array to the dependence tree in Fig. 3.29

	12	13|2	14|23	15|234
		23	24|3	25|34
			34	35|4
				45

$$
\begin{aligned}
\mathrm{f}_{12345}(p_1,p_2,p_3,p_4,p_5) = \quad & \mathrm{f}_1(p_1)\mathrm{f}_2(p_2)\mathrm{f}_3(p_3)\mathrm{f}_4(p_4)\mathrm{f}_5(p_5) \\
& \mathrm{c}_{12}(\mathrm{F}_1(p_1),\mathrm{F}_2(p_2)) \\
& \mathrm{c}_{13|2}\big(\mathrm{F}_{1|2}(p_1|p_2),\mathrm{F}_{3|2}(p_3|p_2)\big) \\
& \mathrm{c}_{14|23}\big(\big(\mathrm{F}_{1|23}(p_1|p_2,p_3),\mathrm{F}_{4|23}(p_4|p_2,p_3)\big)\big) \\
& \mathrm{c}_{15|234}\big(\big(\mathrm{F}_{1|234}(p_1|p_2,p_3,p_4),\mathrm{F}_{5|234}(p_5|p_2,p_3,p_4)\big)\big) \\
& \quad \mathrm{c}_{23}(\mathrm{F}_2(p_2),\mathrm{F}_3(p_3)) \\
& \quad \mathrm{c}_{24|3}\big(\mathrm{F}_{2|3}(p_2|p_3),\mathrm{F}_{4|3}(p_4|p_3)\big) \\
& \quad \mathrm{c}_{25|34}\big(\big(\mathrm{F}_{2|34}(p_2|p_3,p_4),\mathrm{F}_{5|34}(p_5|p_3,p_4)\big)\big) \\
& \quad\quad \mathrm{c}_{34}(\mathrm{F}_3(p_3),\mathrm{F}_4(p_4)) \\
& \quad\quad \mathrm{c}_{35|4}\big(\mathrm{F}_{3|4}(p_3|p_4),\mathrm{F}_{5|4}(p_5|p_4)\big) \\
& \quad\quad\quad \mathrm{c}_{45}(\mathrm{F}_4(p_4),\mathrm{F}_5(p_5)).
\end{aligned}
\tag{3.49}
$$

Given the primary entries, the construction of the set of pairs requiring specification of the copula densities can be readily performed by computer program following the instructions given above.

There are $K(K-1)/2$ copulas to be chosen together with the same number of correlation values for the probability integral transformed uncertain parameters involved (rank correlations for the primary array entries and conditional rank correlations for all others). Probability integral transformation does not change the rank correlations specified for the state of knowledge dependent uncertain parameters. The copulas are defined on the unit cube of dimension K with their pairs of probability integral transformed uncertain parameters having uniform marginal distributions. The values of the rank correlations are, therefore, equal to those of Pearson's correlation.

Discussion

Apart from the specification of the marginal subjective probability distributions for the uncertain parameters P_1, ..., P_K, the copula approach to random sampling requires the following steps:

1. The specification of rank correlation values for pairs of uncertain parameters with their indices given as the primary entries of the array. The rank correlation values simultaneously apply to the pairs of probability integral transformed uncertain parameters over the unit square where they are equal to the value of Pearson's correlation coefficient.
2. The specification of conditional rank correlation values for the secondary, tertiary, etc. entries of the tree array. However, the dimension K is limited by the fact that it will become increasingly more difficult for the expert to specify a conditional rank correlation value the larger the number of uncertain parameters behind the condition stroke.

   Conditional correlations may vary with the values of the conditioning variables. The latter values do vary in the course of the sampling process. As a simplifying step, it is assumed that the conditional correlations are constant over the range of values of the conditioning uncertain parameters.
3. Copula functions need to be chosen for the entries of the tree array with their parameters determined according to the rank correlation and conditional rank correlation values specified in steps 1 and 2.
4. Random sampling needs to be done in stages as determined by the decomposition of the K-variate subjective probability density function.

The advantage of the use of copulas lies in the fact that a multivariate or joint subjective probability density function is defined for the uncertain parameters according to which the sampling of sets of values is carried out. The resulting matrix of correlations is, therefore, guaranteed to be positive definite.

How to sample at random sets of values for the state of knowledge dependent uncertain parameters using the fully specified set of pair copulas?

The normal copula distribution is of great advantage for this task because:

- Given a K-variate normal distribution any conditional distribution is also a normal distribution (Kurowicka and Cooke 2006, page 54—proof for $K = 3$).
- The given conditional rank correlation coefficients can be transformed into conditional correlation coefficients via Eq. (3.39).
- Furthermore, the conditional correlation coefficients (they are assumed to be constant over the full value range of the conditioning uncertain parameters) equal partial correlation coefficients. Consequently, the well-known recursive formulae (Freund and Minton 1979), connecting partial correlation coefficients to correlation coefficients, can be used to replace the conditional correlation coefficients in the tree array by correlation coefficients. In this way, the representation of the K-variate density by $K(K-1)/2$ pairwise copulas can be replaced by one K-dimensional copula density with the correlation matrix $\boldsymbol{R}_{\boldsymbol{Z}}$. The matrix

$\boldsymbol{R_Z}$ is guaranteed to be positive definite since it is derived from a K-variate density function.

– Sampling according to the K-variate normal copula distribution is now done as follows:

From a random sample $(u_1, \ldots, u_K)$, each u_k independently sampled according to the uniform distribution over (0, 1), a random sample $(x_1, \ldots, x_K)$ is obtained through $x_k = \phi^{-1}(u_k)$ and the K-dimensional vector $\boldsymbol{x}$ is transformed into a K-dimensional vector $\boldsymbol{z}$ according to $\boldsymbol{z} = \boldsymbol{Dx}$ where $\boldsymbol{D}$ is the lower triangular matrix from the decomposition of $\boldsymbol{R_Z} = \boldsymbol{DD}'$ (Cholesky Decomposition). Sample values $(p_1, \ldots, p_K)$ are then obtained through the inversion $p_k = F^{-1}(\Phi(z_k))$, $k = 1, \ldots, K$.

3.5.2.8 State of Knowledge Dependence Modelling with Elementary Uncertainty Contributors

State of knowledge dependence of uncertain parameters P_i and P_j is due to uncertainty contributors that are shared by P_i and P_j. An approach to state of knowledge dependence quantification is suggested in this subsection that starts out with a countable set of state of knowledge independent elementary uncertainty contributors with subjective probability distributions of equal variance. Let them be denoted by $Q_1, Q_2, \ldots$.

Instead of asking experts for rank correlation coefficients and conditional rank correlation coefficients, as in the approach with copulas described in Sect. 3.5.2.7, they are asked to symbolically assign elementary uncertainty contributors to each uncertain parameter. The stronger the state of knowledge dependence is perceived to be, the larger the number of elementary uncertainty contributors that P_i and P_j share among those assigned to P_i and P_j together. Since the elementary uncertainty contributors shared are identified, the expert is enabled to traceably account for interrelationships of parameter uncertainties as he conceives them to exist. This is not possible with the information contained in correlation coefficients. For instance, the uncertainty contributors that are pairwise shared by P_i, P_j and P_k may all be different or P_k, P_j and P_i may share the same uncertainty contributor. Yet the latter case may not look any different from the first if the resulting state of knowledge dependence is quantified by pairwise correlation coefficients.

The assignment of elementary uncertainty contributors to K uncertain parameters is documented in a K by K symmetric assignment array $\boldsymbol{A}$. An example is given in Table 3.8 where only the entries in the diagonal and in the upper triangle are shown.

The indices of all the elementary uncertainty contributors, that have symbolically been assigned to the uncertain parameter P_1, are entered into the diagonal field $a_{1,1}$. The off-diagonal field $a_{1,j}$ with row number 1 and column number $j > 1$ carries the indices of those contributors that P_1 shares with the uncertain parameter P_j. The index of a shared contributor is entered with a minus sign in the off-diagonal field $a_{1,j}$ if its contribution to the uncertainty of P_1 and of P_j is in opposite directions.

Table 3.8 Example of an assignment of 16 state of knowledge independent uncertainty contributors to 4 pairwise state of knowledge dependent uncertain parameters

Assignment array **A=**

	1	2	3	4
1	Q_1, Q_2, Q_3, Q_4, Q_5, Q_6, Q_7, Q_8, Q_9, Q_{10}	Q_3, Q_4, Q_6, $-Q_7$	Q_1, Q_2, Q_5	$-Q_3$, Q_6, Q_5
2		Q_3, Q_4, Q_6, $-Q_7$, Q_{11}, Q_{12}, Q_{13}, Q_{14}	Q_{11}, Q_{12}	$-Q_3$, Q_6, $-Q_{11}$
3			Q_1, Q_2, Q_5, Q_{11}, Q_{12}	Q_5, $-Q_{11}$
4				$-Q_3$, Q_5, Q_6, $-Q_{11}$, Q_{15}, Q_{16}

The diagonal field $a_{1,1}$ must contain all elementary contributors that have been assigned to P_1. Therefore, it must also include all elements that are contained in the off-diagonal fields.

A standard normal subjective probability distribution is assigned as state of knowledge expression to each of the elementary uncertainty contributors. From this follows a normal distribution with mean value zero and standard deviation $t_{k,k}^{1/2}$, where $t_{k,k}$ is the number of elementary contributors contained in the diagonal field $a_{k,k}$, $k = 1, \ldots, K$, as symbolic uncertainty expression for each uncertain parameter. The K-variate subjective probability distribution is a joint normal distribution, and the resulting pairwise correlations of the uncertain parameters are given as

$$\rho_{ij} = h_{i,j} / \left(t_{i,i} t_{j,j}\right)^{1/2}, h_{i,j} = \sum\nolimits_{m \epsilon a_{i,j}} c_m \tag{3.50}$$

with $c_m = -1$ if the sign of the element index in the array field $a_{i,j}$ is negative and +1 otherwise. The upper triangle of the correlation matrix that results from the assignment array in Table 3.8 is shown below.

$$\rho_{12} = \frac{2}{\sqrt{80}} = 0.224 \quad \rho_{13} = \frac{3}{\sqrt{50}} = 0.424 \quad \rho_{14} = \frac{1}{\sqrt{60}} = 0.129$$
$$\rho_{23} = \frac{2}{\sqrt{40}} = 0.316 \quad \rho_{24} = -\frac{1}{\sqrt{48}} = -0.144$$
$$\rho_{34} = \frac{0}{\sqrt{30}} = 0$$

The expert may want to differentiate the uncertainty contribution of Q to the uncertain parameters P_i and P_j by its strength. For instance, the contribution of Q to the uncertainty of P_i may be judged to be twice as strong if compared to its contribution to the uncertainty of P_j. Table 3.9 shows elementary contributors with coefficients differing from unity. They account for multiples or fractions of their strength.

Table 3.9 Example of an assignment of 16 state of knowledge independent elementary uncertainty contributors to 4 pairwise state of knowledge dependent uncertain parameters using a strength factor for Q_6

Assignment array **A**=

	1	2	3	4
1	Q_1, Q_2, Q_3, Q_4, Q_5, 1.5 Q_6, Q_7, Q_8, Q_9, Q_{10}	Q_3, Q_4, Q_6, -Q_7	Q_1, Q_2, Q_5	-Q_3, 1.5 Q_6, Q_5
2		Q_3, Q_4, Q_6, -Q_7, Q_{11}, Q_{12}, Q_{13}, Q_{14}	Q_{11}, Q_{12}	-Q_3, 1.5 Q_6, -Q_{11}
3			Q_1, Q_2, Q_5, Q_{11}, Q_{12}	Q_5, -Q_{11}
4				-Q_3, Q_5, 1.5 Q_6, -Q_{11}, Q_{15}, Q_{16}

The corresponding correlation coefficients are obtained as

$\rho_{ij} = h_{i,j}/(t_{i,i}t_{j,j})^{1/2}$, $h_{i,j} = \sum_{m \epsilon a_{i,j}} c_m$ with $c_m = d_{i,m}d_{j,m}$ and $d_{i,m}$ is the factor of Q_m in the field $a_{i,i}$ and $d_{j,m}$ is the factor of Q_m in the field $a_{j,j}$ while $t_{i,i} = \sum_{m \epsilon a_{i,i}} d_{i,m}^2$ and $t_{j,j} = \sum_{m \epsilon a_{j,j}} d_{j,m}^2$.

The correlation coefficients resulting from the assignments in Table 3.9 are shown below.

$$\rho_{12} = \frac{2.5}{\sqrt{8*11.25}} = 0.264 \quad \rho_{13} = \frac{3}{\sqrt{5*11.25}} = 0.400 \quad \rho_{14} = \frac{2.25}{\sqrt{7.25*11.25}} = 0.249$$

$$\rho_{23} = \frac{2}{\sqrt{8*5}} = 0.316 \quad \rho_{24} = -\frac{0.5}{\sqrt{8*7.25}} = -0.066 \quad \rho_{34} = \frac{0}{\sqrt{5*7.25}} = 0$$

The K-variate normal distribution is a joint normal distribution with correlation matrix determined as explained above. Consequently, the correlation values can be converted into the corresponding Spearman (rank) correlation values ρ_{ij}^S as follows:

$$\rho_{ij}^S = (6/\pi)\arcsin\left(\frac{\rho_{ij}}{2}\right). \tag{3.51}$$

These rank correlation values are not changed by the probability integral transformation and the subsequent back transformation into the specified marginal subjective probability distributions of the uncertain parameters.

The experts who specified the entries in the assignment matrix may want to check the resulting (rank) correlation values at this stage and decide whether they are sufficiently compatible with values they might have in mind. If the difference is judged to be too large, a decision has to be made as to whether to prefer the existing value (given the considerations that have gone into building the assignment array $\boldsymbol{A}$) or to perform changes to $\boldsymbol{A}$ in order to reduce the difference. These changes will not only affect the correlation value in question but also others. The array $\boldsymbol{A}$ shows which correlation values are affected and by how much.

The next step is the actual sampling of the parameter values. Sampling at random a set of values $p_1, \ldots, p_K$ of the uncertain parameters is straightforward:

1. Independently, sample one value each according to the standard normal distribution for each of the elementary uncertainty contributors in the assignment array $\boldsymbol{A}$.
2. To obtain a value p_k for the uncertain parameter P_k, add the values sampled in step 1 for all elements in the diagonal field of row k of the assignment array thereby observing the sign (and, if given, the strength factor) of the element index.
3. Probability integral transform the values of the sums obtained in step 2 according to the marginal normal distributions assigned to the uncertain parameters through the assignment array $\boldsymbol{A}$.
4. Back transformation of the values obtained in the previous step for each uncertain parameter according to the inverse of the specified marginal subjective probability distribution.

Repeat steps 1–4 N times if N is the sample size required.

How can the special cases of state of knowledge dependence quantification, discussed in Sects. 3.5.2.1–3.5.2.3, be embedded into this approach?

- Conditional distributions for P_j given P_i:
  P_i goes into the assignment array while the sampling for P_j is done after the random sampling for all uncertain parameters in the assignment array.
- Constraint for P_j as function of P_i:
  P_i goes into the assignment array while the sampling for P_j is done after the random sampling for all uncertain parameters in the assignment array.
- Functional relationship like $P_k = g(P_i,P_j)$ between uncertain parameters:
  All uncertain parameters in the argument list of the functional relationship go into the assignment array while the sampling for P_k is done after the random sampling for all uncertain parameters in the assignment array. It is obvious how to proceed if conditional distributions or constraints are defined for some of the parameters in the argument list.
- Proportions like $P_k = 1 - P_i - P_j$, $0 \leq P_i, P_j, P_k \leq 1$:
  The pairwise state of knowledge independent uncertain parameters Q_i, Q_j go into the assignment array (as individual uncertain parameters and not as elementary uncertainty contributors) only if they are state of knowledge dependent on other uncertain parameters. The values for P_i, P_j, P_k are computed, using their relationships to Q_i, Q_j (see Sect. 3.5.2.3), after the random sampling for all uncertain parameters in the assignment array has been completed.

Comments

- This approach encourages the experts to investigate the reasons and nature of the state of knowledge dependence and thereby the interrelationships of parameter uncertainties.
- The experts do not need to specify (rank) correlation coefficients and conditional (rank) correlation coefficients.
- A multivariate normal distribution is obtained according to the assignment array $\boldsymbol{A}$. Using the inverses of the specified marginal subjective probability distribution functions, it translates into a joint subjective probability distribution for the uncertain parameters.
- The pairwise correlation values are known for the multivariate normal distribution. The corresponding matrix is positive definite. The pairwise correlations can be transformed into rank correlations using Eq. (3.51). Probability integral transformation provides a multivariate distribution over the unit hypercube. Its correlation matrix equals the matrix of the rank correlations as these are unchanged by the transformation. Their matrix is, therefore, positive definite. Transformation into the specified marginal subjective probability distributions does not change this matrix.
- The resulting (rank) correlations may be checked whether they are sufficiently compatible with values the experts might have in mind. If the difference is judged to be too large, a decision has to be made as to whether to prefer the existing value

(given the considerations that have gone into the building of the assignment array $\boldsymbol{A}$) or to perform changes to $\boldsymbol{A}$ in order to reduce the difference. These changes will not only affect the correlation value in question but also others. The array $\boldsymbol{A}$ shows which correlation values are affected and by how much.

– Random sampling of parameter values as explained in steps 1–4 above is straightforward.

3.5.3 Sample Expressions of State of Knowledge Dependence

The correlation coefficients discussed in the previous section are population measures of association. They apply to the whole population of all possibly true pairs of values for P_i and P_j. A random sample taken according to a joint distribution for P_i and P_j, that satisfies the specified marginal distributions and the population measure for state of knowledge dependence, will provide a sample correlation value that differs from the value specified for the population. Sometimes, the measure of association is given as a sample measure, i.e. it is specifically required that the sample correlation value is as specified. In this case, the sampling process is carried out such that the sample satisfies this requirement. The disadvantage is, however, that the sample is no longer purely random but has the specified sample measure of association as a deterministic component. Some useful uncertainty statements, derived in Chap. 5 for computer model results, require that the sample be random. They can, therefore, not be obtained if sampling is done according to sample measures of association.

In the next analysis step (Step 3), a joint subjective probability distribution satisfying the state of knowledge quantifications as well as the state of knowledge dependence expression for all uncertainties is propagated through the model. This is done by Monte Carlo simulation, i.e. a random sample of size N is drawn according to this joint distribution. The sample consists of N sets of M values each where M is the number of uncertain parameters. The model is then evaluated for each of the N sets. With this procedure in mind, it is quite natural to hope for a random sample that has (rank) correlation values sufficiently close to those specified.

Since the sample (rank) correlations are specified as a property of the random sample, the description of how to achieve the specified sample (rank) correlations is given in Sect. 4.4.1.14 of the subchapter dealing with sampling methods.

3.5.4 A Multivariate Sample

In some cases, probability distributions, summarizing stochastic variability, are used as state of knowledge expressions for the epistemic uncertainties $P_1, \ldots, P_K$. A sufficiently large multivariate sample obtained according to their joint probability distribution would be the best way to quantify state of knowledge dependence. For

Table 3.10 Eight sets of parameter values derived from experiments

Parameters of the model				
a	b	c	d	Source (year)
4.32	1.35	2.35	4.33	Expert(group) 1 (1991)
6.98	1.98	2.45	3.03	Expert(group) 2 (1993)
5.74	2.76	2.64	3.12	Expert(group) 3 (1994)
3.79	3.25	3.94	3.36	Expert(group) 4 (2000)
4.26	2.98	2.03	4.73	Expert(group) 5 (1999)
5.93	1.75	3.65	2.74	Expert(group) 6 (2001)
6.32	2.03	2.02	1.98	Expert(group) 7 (2003)
4.58	1.95	2.76	4.03	Expert(group) 8 (1995)
4.32	0.278	1.98	4.27	Expert(group) 9 (1997)

They may serve as joint state of knowledge expression for the four uncertain parameters $P_1 = a$, $P_2 = b$, $P_3 = c$ and $P_4 = d$ once a subjective probability sw(l) has been assigned to each of the $L = 8$ sets such that $\sum_{l=1}^{L} \mathrm{sw}(l) = 1$

instance, $p_{1,l}, \ldots, p_{K,l}$ could be the coefficients of a model fitted to experimental data provided by expert l. If such sets of data are provided by L experts, each fitting the model to his own experiments, the L sets of values $p_{1,l}, \ldots, p_{K,l}$, $l = 1,\ldots, L$, may be thought of as a random sample and may therefore be used directly for the purpose of uncertainty analysis. In this special case, the sample size will be considerably smaller than the size N that is to be drawn for the purpose of uncertainty propagation in Step 3 of the analysis. Sampling with replacement from the L sets of K values each may still be better than estimating values of population measures of association for the K parameters.

Table 3.10 shows eight sets of values for the uncertain parameters P_1, P_2, P_3 and P_4. Once a subjective probability has been assigned to each set, state of knowledge dependence of P_1, P_2, P_3, P_4 is inherently taken into account. The probabilities have to add up to unity.

A typical example would be the following: If the computer model receives several results from another model (feeder model) as input, the N sets of output values (scalar values, sequences, arrays and/or tables of values) from a Monte Carlo uncertainty analysis of the application of the feeder model are a multivariate sample for the Monte Carlo uncertainty analysis of the computer model application. Any state of knowledge dependence is automatically taken into account (within the accuracy bounds that may be achieved with the given sample size) by using this sample.

3.5.5 Summary of Sect. 3.5

– State of knowledge dependence is due to shared (not necessarily additive) contributors to uncertainty.

- State of knowledge dependence can have a significant effect on the results of the uncertainty analysis and must, therefore, be identified and quantified with care (see Sect. 3.5.1).
- State of knowledge dependence can be quantitatively expressed in many ways depending on what is known about its cause (see Sect. 3.5.2).
- The approximation of the joint subjective probability distribution of a pair of state of knowledge dependent parameters by a small set of conditional distributions for one parameter (given the value of the other parameter) has great potential for expressing state of knowledge dependence. It is the most natural approach both for expert judgment elicitation as well as for the Monte Carlo simulation. The practical use is, however, limited to very small sets of state of knowledge dependent uncertain parameters.
- Correlation coefficients have to satisfy a number of restrictions (see the summary of Sect. 3.5.2.4 and the comparison in Sect. 3.5.2.5) and are not always easy to elicit from experts. The larger the number K of state of knowledge dependent uncertain parameters, the less likely it is that the matrix of specified correlation coefficients will satisfy the condition of positive definiteness.
- The copula approach (see Sect. 3.5.2.7) defines a multivariate subjective probability distribution for the state of knowledge dependent uncertain parameters and thereby guarantees a positive definite (rank) correlation matrix. However, the number of copula functions as well as rank correlations and conditional rank correlations that need to be specified grows rapidly with the number K of state of knowledge dependent uncertain parameters and so does the number of uncertain parameters behind the condition stroke. The copula approach is, therefore, limited to small sets of state of knowledge dependent uncertain parameters.
- The approach using the symbolic assignment of elementary uncertainty contributors (see Sect. 3.5.2.8) also defines a multivariate subjective probability distribution for the state of knowledge dependent uncertain parameters and thereby guarantees a positive definite (rank) correlation matrix. Neither does it require the specification of copula functions nor are the values of (rank) correlation or conditional (rank) correlation coefficients to be provided. The expert is encouraged to identify cause, communalities and strength of the shared uncertainty contributions. The random sampling of sets of parameter values is straightforward for this approach.
- Scatterplots greatly aid the expert in deciding about the suitability of the specified state of knowledge dependence expression.

The next subchapter deals with the practical aspects of the elicitation of expert judgment. The judgment is required for the specification of the marginal subjective probability distributions that quantify the state of knowledge for each of the uncertain parameters. It also deals with the elicitation of expressions quantifying state of knowledge dependence.

3.6 State of Knowledge Elicitation and Probabilistic Modelling

The state of knowledge is elicited from experts familiar (or individuals quite familiar) with the respective uncertain parameter, model formulation or input datum of the computer model application. These so-called substantive experts are characterized in Council on Environmental Quality (1980) as follows:

"EXPERT–You should consider yourself an expert if you belong to that small community of people who currently study, work on and dedicate themselves to the subject matter. Typically, you know who else works in this area; you know the relevant domestic and probably the foreign literature; you attend conferences and seminars on the subject, sometimes reading a paper and sometimes chairing the sessions; you are most likely to have written up and/or published the results of your work. If a leading national or international institution were to convene a seminar on this subject, you would expect to be invited. Other experts in this field may disagree with your views but invariably respect your judgment; comments such as "this is an excellent person on this subject" would be typical when inquiring about you.

QUITE FAMILIAR—You are quite familiar with the subject matter either if you were an expert some time ago but feel somewhat rusty now because other assignments have intervened (even though, because of the previous interest, you have kept reasonably abreast of current developments in the field); or if you are in the process of becoming an expert but still have some way to go to achieve mastery of the subject; or if your concern is with integrating detailed developments in the area, thus trading breadth of understanding for depth of specialization".

An uncertainty analyst (also called "normative expert") organizes the elicitation sessions. He needs to know how to conduct an elicitation so as to be able to assist the substantive experts in providing the necessary information while keeping the influence of bias as small as possible. Furthermore, it is essential that he be familiar with the ways of quantitatively expressing state of knowledge by subjective probability. He also needs to be familiar with the concepts and tools from probability calculus and statistics used to propagate the quantified states of knowledge through the computer model. It is his task to also interpret the resulting state of knowledge quantification for the model result with respect to uncertainty ranges and uncertainty importance. Last but not least, the uncertainty analyst needs to be familiar with the question to be answered by the application of the computer model and with the intended use of the answer in any decision-making process. This background information may, however, also be contributed by a third person (it may be called "the client") involved in the elicitation sessions. The client writes the session minutes that would otherwise need to be produced by the analyst. The minutes should carry the names and signatures of the three individuals involved (expert, analyst and client) and should be filed together with the documentation sheet for the respective uncertain parameter, model formulation or input datum. A detailed description of the contents of the documentation sheet is given below in Sect. 3.6.1.5.

Obviously, not every uncertainty analysis of a computer model application permits or requires this separation of responsibilities. Quite often the other extreme will be the case, namely substantive expert, analyst and client united in one person.

3.6.1 State of Knowledge Elicitation and Probabilistic Modelling for Data

Model uncertainties are represented by uncertain parameters (see Sect. 3.4), and uncertain parameters are categorized as uncertain data in order to simplify matters. Therefore, without loss of generality, this section only uses the term "data".

The elicitation process is described for the following two extreme situations:

- No separation of responsibilities, i.e. only one substantive expert who also acts as analyst and client;
- Total separation of responsibilities, i.e. a team consisting of an analyst, a client and one substantive expert per uncertain parameter, model formulation or input datum.

  The state of knowledge elicitation involving several substantive experts for the same uncertain parameter, model formulation or input datum is discussed in Sect. 3.7.

3.6.1.1 No Separation of Responsibilities

It is the (substantive) expert's task to provide a subjective probability distribution that expresses his state of knowledge for the uncertain datum P. To perform this task, the expert executes a work plan. The steps of this plan are as follows:

1. Search for information on P or similar data by asking:
   - 1.1 Are there directly applicable and sufficient observations that would enable a statistical approach?
   - 1.2 Are there only observations that are not directly applicable or that are insufficient and therefore need to be supplemented by theoretical considerations?
   - 1.3 Are there no observations so that the state of knowledge is based solely on theoretical and/or plausibility considerations?
   - 1.4 Can P be represented as a function of two or more uncertain data that are more readily amenable to state of knowledge quantification (Hora et al. 1993)? If yes, go back to the top for each of these data.
   - 1.5 Has the state of knowledge already been expressed by a subjective probability distribution for a similar, related datum or can the state of knowledge be more readily quantified for a similar datum that does not quite correspond to

the description of *P*, and is it possible to perform the required adjustments (taking their uncertainty into account in turn)?

2. Quantify your state of knowledge by asking:
   - 2.1 Which is the range of all possibly true values?
   - 2.2 Which are the endpoints of a range such that the subjective probability (your degree of belief) is 0.9 for the true value to be within this range? Or in other words: You see only 5% chance for values below this range and 5% chance for values above this range to be true.
   - 2.3 Which is the value that is most likely true?
   - 2.4 Is there any information that would suggest a specific shape of the subjective probability density function towards the endpoints of the range given in 2.1 or that would suggest even one of the customary distribution types?
3. Use the information given in steps 1 and 2 to specify a subjective probability distribution for the uncertain datum *P*. Apply the maximum entropy principle (Buckley 1985) if the information does not uniquely determine the distribution.

Finally, the expert needs to check his quantifications by verifying whether all the information available to him is adequately reflected with respect to its strength and weight for the assessment task of the computer model application. While executing the work plan, bias of the state of knowledge quantifications may creep in at every step. The search for information in step 1 may subconsciously be limited to those observations, state of knowledge quantifications for similar data, theoretical and plausibility considerations that are most readily available or retrieved from memory (availability bias). The use of observations and/or information on data similar to *P* may be flawed by overemphasizing certain similarity aspects while neglecting features that would suggest to give this information less weight in the state of knowledge quantification for *P* (representativeness bias). When choosing the range of all possibly true values in step 2, one might subconsciously first think of a most likely value (that accounts for only part of the information) and then allow for too little deviation from the latter in order to account for all of the information (anchoring bias). The endpoints of a range of values that is supposed to contain the true value with subjective probability 0.9 may exhibit a tendency of being chosen too close to the most likely value (overconfidence bias). A thorough description and discussion of biases and their sources can be found in Vick (2002) and Kahneman (2011) together with illustrative and practical examples.

3.6.1.2 Total Separation of Responsibilities

The elicitation is carried out with a team consisting of an analyst, a client and one substantive expert per uncertain datum *P*. The analyst elicits the state of knowledge in individual sessions with each substantive expert after all substantive experts participated in a “background” session and in a training session.

– The background session

  The client supplies all substantive experts with sufficient information about the computer model and about the purpose of its application. The analyst provides all necessary background information about how an uncertainty analysis is conducted, what kind of information is required from the experts and in which format the expert's answers will be expected.

– The training session

  All substantive experts identified for the uncertain data are invited to a training session that is conducted by the analyst (Hora and Hora 1992) prior to the actual elicitation sessions. The main purpose of the training session is the "hands-on" preparation of the experts for the format in which the elicitation questions will be presented to them and in which the answers are expected. Additionally, the session is to make them aware of the sources of bias along the path to state of knowledge quantification. The training session should conclude with a set of 10 questions the answers of which are not likely to be known by the experts but are known to the analyst. The statistical yearbook or statistical annual is usually a good information source for the formulation of such questions. They should be chosen such that there is potential for all of the types of bias mentioned above. Since the questions are most likely from outside their expertise, the experts must allow for wide uncertainty ranges if they do not want to fall prey to overconfidence bias. The questions should ask for the range of all possibly true values and for the endpoints of an interval that contains the true value with subjective probability 0.9. An expert may be called well calibrated if his intervals contain the true value for 9 out of 10 questions (see also the remarks on "training" under point 5 of Sect. 3.7).

– The elicitation session

  Participants of the session are the expert, the analyst and the client. In what follows it is assumed that there is only one expert per uncertain datum. The case where several experts are questioned with respect to the same datum is discussed in Sect. 3.7.

Basically, the same work plan as in Sect. 3.6.1.1 has to be executed by the expert. Different to Sect. 3.6.1.1, the expert is now supported by the analyst. At the beginning of the session, the analyst shows the expert a summary of the outcome from the training session and in particular the answers given by the expert himself and how they relate to the true answers. Any influence of bias is pointed out and it is explained how it could have been avoided.

The analyst gives the uncertain data consecutive numbers and opens a documentation sheet for each. The documentation sheet contains the number of the uncertain datum and its name in the encoded computer model. Together with the expert, the analyst enters a short description of the meaning of the datum. This is followed by the symbol used for the datum in the mathematical model and by the physical unit (if any) in which the computer model expects the numbers for this datum to be provided. The best estimate value (see Sect. 3.6.1.4) used for a point result of the computer model application and a reference value (see Sect. 3.6.1.4) are also entered

but should only be chosen once the state of knowledge has been quantified in order to avoid anchoring bias.

The actual state of knowledge quantification starts with the analyst first asking the expert for the endpoints of the range of all possibly true values for the uncertain datum. The question goes as follows:

1. "Which is the smallest value at which you are 100% sure that the true value does not lie above?"
   "Which is the largest value at which you are 100% sure that the true value does not lie below?"

The analyst points out that these two endpoints must be chosen such that any values above (respectively below) can be excluded, i.e. these endpoints delimit the range of possibly true values and the expert must see no possibility for values beyond these endpoints to be true. The expert is asked to provide his arguments for the exclusion of these values. To support him in his argumentation, the analyst challenges the expert:

"Suppose, tomorrow you read in a report that a value beyond the endpoints of your 100% range was shown to be true. What could be the reason for this?"

Experts from engineering often associate a different meaning with the terms "minimum" and "maximum". They tend to think of "reasonably possible" ranges meaning that they still see a chance, albeit small, for values beyond the endpoints of the range to be true. If an expert can think of a reason why a value outside his 100% range could have turned out to be true, he is asked to adjust the two endpoints accordingly. Asking for a most likely value first is not recommended as it is seen as a source of anchoring bias.

Next, the expert is asked to provide a value for the uncertain datum such that his subjective probability is 0.1 (0.05) for values below to be true and another value such that his subjective probability is 0.1 (0.05) for values above to be true. The question goes as follows:

2. "At which value do you see 10% (5%) chance for the true value to lie above and at which value 10% (5%) chance for the true value to lie below?"

Again, these values need to be accompanied by supporting argumentation. Later, the analyst may ask the same question using the complementary probability values, i.e. by replacing the 10% (5%) by 90% (95%) to see whether there is any contradiction.

Next, the analyst asks for a value such that the expert has no preference with respect to the true value lying either above or below. The question goes as follows:

3. "At which value is it impossible for you to say whether the true value is more likely to lie above or below?"

This value constitutes the 50% quantile (or median) of the subjective probability distribution. Supporting argumentation would be appreciated.

It might be even possible to go one step further and to ask the following questions:

4. "Let us consider only values from the range below the value given under (3). At which value from this sub-range is it impossible for you to say whether the true value is more likely to lie above or below?"

and analogously:

5. "Let us consider only values from the range above the value given under (3). At which value from this sub-range is it impossible for you to say whether the true value is more likely to lie above or below?"

These latter two values would now constitute 25% and 75% quantiles of the subjective probability distribution, i.e. a quartile range.

As a last question, the analyst may ask:

6. It may be that not every value below your 10% (5%) quantile (above your 90% (95%) quantile) can be considered as equally likely true. Values closer to those quantiles may be more likely true than those further away. If that is the case, can you give an impression (graphically) of how your degree of belief changes between the 10% (5%) quantile and the lower endpoint and again between the 90% (95%) quantile and the upper endpoint of the range of possibly true values?

Again, argumentation for the depicted changes in degree of belief would be appreciated.

While the expert provided answers to the above questions the analyst made a graphical representation (possibly using a suitable software package) showing how the degree of belief is distributed according to the answers given under (1)–(6). This graphical representation is now presented to the expert and the expert is asked to check whether it is an adequate expression of his state of knowledge. This inspection may give rise to some adjustments that need to be accompanied by supporting argumentation. A subjective probability distribution needs to be chosen that contains a minimum of information in addition to the information contained in the expert's answers to question (1)–(6). This task is supported by the maximum entropy principle (Buckley 1985). According to this principle, the following distributions may, for instance, be used in the uncertainty analysis:

Expert's quantifications	Distribution type according to the maximum entropy principle
– Only minimum and maximum possibly true values	Uniform
– Minimum, maximum and quantiles	Piecewise uniform
– Mean value and variance	Normal.

The following is a presentation of the distribution types that are most commonly used in uncertainty analysis. The presentation of each type starts with an explanation why an expert might opt to use the distribution type as the probabilistic expression of

his state of knowledge. This is followed by the formulae for the probability density function (pdf) and the cumulative distribution function (cdf) as well as by the mean value, variance and median value. Last but not least, it is indicated how random samples might be generated according to each of the distributions specified as state of knowledge expression.

3.6.1.3 Subjective Probability Distributions Frequently Used as State of Knowledge Expressions

Most elicited state of knowledge quantifications may be satisfactorily represented by one of the following types of subjective probability distributions:

- Uniform
- Triangular
- Trapezoidal
- Piecewise uniform
- Polygonal line
- Truncated normal
- t- or Student
- Beta
- Logarithmic uniform (loguniform)
- Logarithmic normal (lognormal)
- Discrete.

Each of these distributions is presented below together with an illustrative example of the density graph. The mathematical expressions for mean value, variance and median value are also given, followed by instructions for the generation of a random sample of values drawn according to the distribution in a Monte Carlo simulation.

Subjective probability density functions fitted to elicited state of knowledge quantifications need to be inspected by the expert in order to decide whether any modifications to his state of knowledge quantifications are suggested by the density graph.

- Uniform distribution

The uniform subjective probability distribution (see Fig. 3.30) assigns equal degree of belief to any value between the elicited minimum (p_{min}) and maximum (p_{max}) possibly true value of the uncertain datum P. It says that the expert does not see any reason nor is he aware of any information that would enable him to prefer any value between the minimum and the maximum possibly true value.

Pdf:
$f(p) = 1/(p_{max} - p_{min})$

Cdf:
$F(p) = (p - p_{min})/(p_{max} - p_{min})$

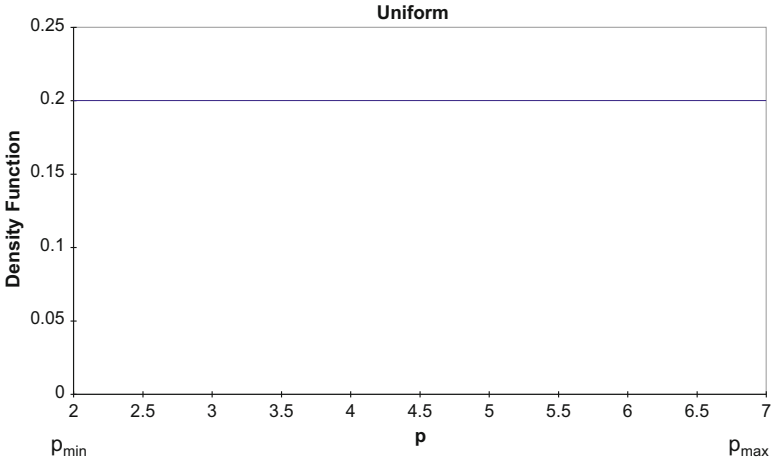


Fig. 3.30 Density function of a uniform distribution

for $p_{min} \leq p \leq p_{max}$

Mean value:	$E\{P\} = (p_{min} + p_{max})/2$
Variance:	$Var\{P\} = (p_{max} - p_{min})^2/12$
Median value:	$p_{50\%} = E\{P\}$.
Monte Carlo simulation:	Sample a value u from the (0, 1) uniform distribution (standard subroutine) and transform it into a value p using the transformation $p = F^{-1}(u)$ where F^{-1} denotes the inverse of the distribution function. In other words, p is that value for P that satisfies $F(p) = u$, i.e. that solves the equation $u = (p - p_{min})/(p_{max} - p_{min})$ and therefore $p = u(p_{max} - p_{min}) + p_{min}$.

– Triangular distribution

This subjective probability distribution is the simplest state of knowledge expression in cases where the expert restricts the range of possibly true values for P to a finite interval delimited by p_{min} and p_{max} and where there is information that suggests one value p_{mod} as most likely true while values below and above are to receive degrees of belief that decrease linearly to zero towards both endpoints (see Fig. 3.31).

Pdf:
$f(p) = h(p - p_{min})/(p_{mod} - p_{min})$
for $p_{min} \leq p < p_{mod}$
and

Cdf:
$F(p) = h(p - p_{min})^2/2(p_{mod} - p_{min})$

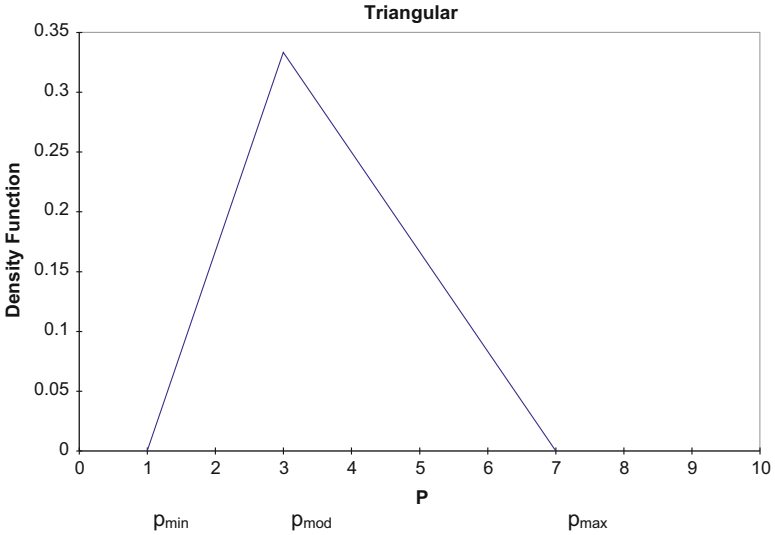


Fig. 3.31 Density function of a triangular distribution; h is the maximum density value

$f(p) = h(p_{max} - p)/(p_{max} - p_{mod})$ $\quad F(p) = 1 - h(p_{max} - p)^2/2(p_{max} - p_{mod})$
for $p_{mod} < p \leq p_{max}$
$h = 2/(p_{max} - p_{min}) = f(p_{mod})$ $\quad F(p_{mod}) = (p_{mod} - p_{min})/(p_{max} - p_{min})$.

Note, that p_{mod} may be equal to p_{min} or p_{max}.

Mean value: $E\{P\} = (p_{min} + p_{mod} + p_{max})/3$ if $p_{min} < p_{mod} < p_{max}$
$E\{P\} = (2p_{min} + p_{max})/3$ if $p_{min} = p_{mod} < p_{max}$
$E\{P\} = (p_{min} + 2p_{max})/3$ if $p_{min} < p_{mod} = p_{max}$

Variance: $\mathrm{Var}\{P\} = \left(p_{min}^2 + p_{mod}^2 + p_{max}^2 - p_{min}p_{mod} - p_{min}p_{max} - p_{mod}p_{max}\right)/18$

Median value: See "Monte Carlo simulation" for $u = 0.5$.

Monte Carlo simulation: Sample a value u from the (0, 1) uniform distribution (standard subroutine) and transform it into a value p using the transformation $p = F^{-1}(u)$ where F^{-1} denotes the inverse of the distribution function. In other words, p is that value for P that satisfies $F(p) = u$.
For $u \leq F(p_{mod})$:
Solve $u = (p - p_{min})^2/[(p_{mod} - p_{min})(p_{max} - p_{min})]$ for p;
For $u \geq F(p_{mod})$:
Solve $u = 1 - (p_{max} - p)^2/[(p_{max} - p_{mod})(p_{max} - p_{min})]$ for p.

– Trapezoidal distribution

This subjective probability distribution (see Fig. 3.32) is the simplest state of knowledge expression in cases where the expert restricts the range of possibly true

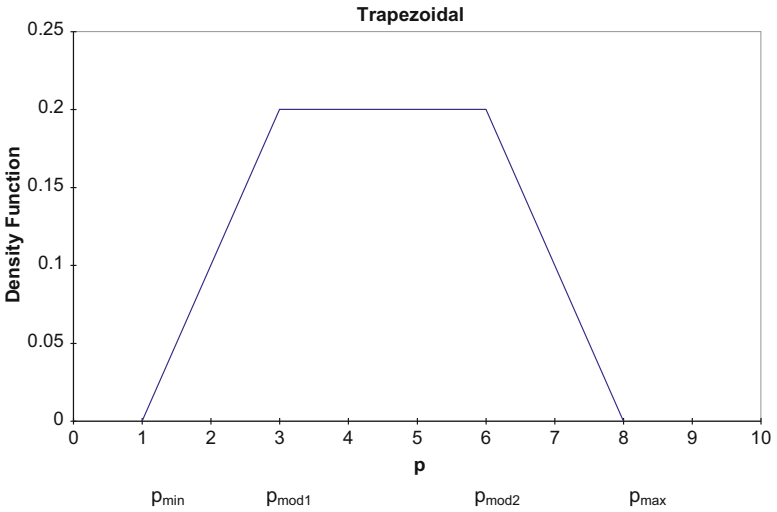


Fig. 3.32 Density function of a trapezoidal distribution; h is the maximum density value

values for P to a finite interval delimited by p_{min} and p_{max} and where there is not enough information to suggest only one value as most likely true. Each value between p_{mod1} and p_{mod2}, with $p_{min} \leq p_{mod1} < p_{mod2} \leq p_{max}$, receives the highest and equal degrees of belief while values below p_{mod1} and values above p_{mod2} receive degrees of belief that decrease linearly to zero towards both endpoints.

Pdf:

$f(p) = h(p - p_{min})/(p_{mod1} - p_{min})$ for $p_{min} \leq p < p_{mod1}$

$f(p) = h$ for $p_{mod1} \leq p \leq p_{mod2}$

$f(p) = h(p_{max} - p)/(p_{max} - p_{mod2})$ for $p_{mod2} < p \leq p_{max}$

$h = 2/(p_{max} - p_{min} + p_{mod2} - p_{mod1})$

Cdf:

$F(p) = h(p - p_{min})^2/2(p_{mod1} - p_{min})$

$F(p) = h(p - p_{mod1}) + F(p_{mod1})$

$F(p) = 1 - h(p_{max} - p)^2/2(p_{max} - p_{mod2})$

$F(p_{mod1}) = (p_{mod1} - p_{min})/(p_{max} - p_{min} + p_{mod2} - p_{mod1})$

$F(p_{mod2}) = F(p_{mod1}) + (p_{mod2} - p_{mod1})h$

To compute mean value and variance, it is assumed that P is a mixture of three state of knowledge independent uncertain data, namely

$P_1 \sim$ triangular over (p_{min}, p_{mod1})
$P_2 \sim$ uniform over (p_{mod1}, p_{mod2})
$P_3 \sim$ triangular over (p_{mod2}, p_{max})

with the nonnegative mixture coefficients c_1, c_2, c_3 where $c_1 + c_2 + c_3 = 1$ and

$$c_1 = \int_{p_{min}}^{p_{mod1}} \mathrm{f}(p)\mathrm{d}p = \mathrm{F}(p_{mod1}), \quad c_2 = \int_{p_{mod1}}^{p_{mod2}} \mathrm{f}(p)\mathrm{d}p = \mathrm{F}(p_{mod2}) - \mathrm{F}(p_{mod1}),$$
$$c_3 = \int_{p_{mod2}}^{p_{max}} \mathrm{f}(p)\mathrm{d}p = 1 - \mathrm{F}(p_{mod2})$$

Mean value: $\mathrm{E}\{P\} = \sum_{i=1}^{3} c_i \mathrm{E}\{P_i\}$

Variance: $\mathrm{Var}\{P\} = \sum_{i=1}^{3} c_i \mathrm{Var}\{P_i\} + \sum_{i=1}^{3} c_i (\mathrm{E}\{P_i\} - \mathrm{E}\{P\})^2$

This follows from the relationship $\mathrm{Var}\{P\} = \mathrm{E}\{\mathrm{Var}\{P|P{=}P_i\}\} + \mathrm{Var}\{\mathrm{E}\{P|P{=}P_i\}\}$ with $\mathrm{E}\{P_i\}$ and $\mathrm{Var}\{P_i\}$ as given above for the uniform and the triangular distribution.

Median value: See "Monte Carlo simulation" for $u = 0.5$.

Monte Carlo simulation: Sample a value u from the (0, 1) uniform distribution (standard subroutine) and transform it into a value p using the transformation
$p = \mathrm{F}^{-1}$ (u) where F^{-1} denotes the inverse of the distribution function. In other words, p is that value for P that satisfies $\mathrm{F}(p) = u$.
For $u \le \mathrm{F}(p_{mod1})$:
Solve $u = h(p - p_{min})^2/2(p_{mod1} - p_{min})$ for p;
For $\mathrm{F}(p_{mod1}) \le u \le \mathrm{F}(p_{mod2})$:
Solve $u - \mathrm{F}(p_{mod1}) = h(p - p_{mod1})$ for p;
For $u \ge \mathrm{F}(p_{mod2})$:
Solve $u = 1 - h(p_{max} - p)^2/2(p_{max} - p_{mod2})$ for p.

– Piecewise uniform distribution

The piecewise uniform or stepwise uniform subjective probability distribution (see Fig. 3.33) assigns equal degrees of belief to the values between two successive elicited quantiles. The value p_{min} is the 0%-quantile denoted by p_1 and p_{max} is the 100%-quantile denoted by p_K if the elicitation provided $K-2$ quantiles in addition to p_{min} and p_{max}. The distribution says that the expert does not see any reason nor is he aware of any information that would enable him to prefer any value from those between two successive elicited quantiles. The values q_k, k = 1, …, K are the corresponding quantile percentages divided by 100, with $q_1 = 0$ and $q_K = 1$.

Pdf: $\mathrm{f}(p) = (q_k - q_{k-1})/(p_k - p_{k-1})$

Cdf: $\mathrm{F}(p) = q_{k-1} + (q_k - q_{k-1})(p - p_{k-1})/(p_k - p_{k-1})$

for $p_{k-1} \le p \le p_k$, $k = 2, \ldots, K$

where q_k is the quantile percentage divided by 100 of the quantile p_k

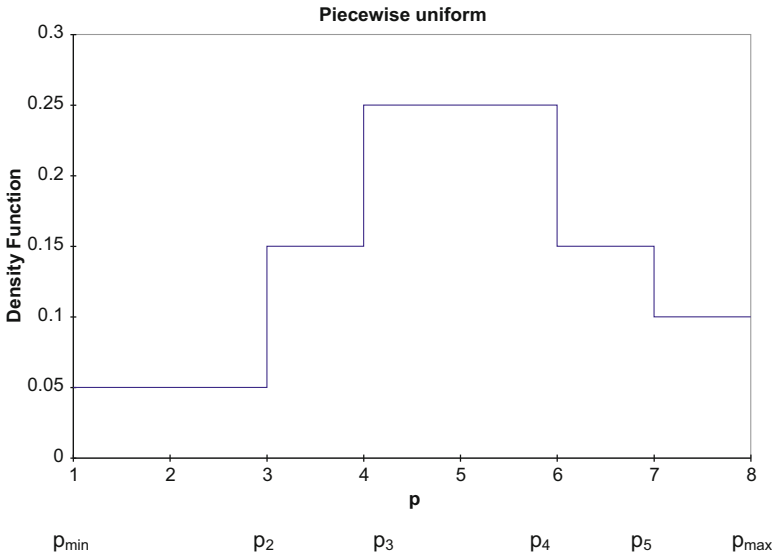


Fig. 3.33 Density function of a piecewise uniform distribution with the heights $h_1 = q_2/(p_2 - p_{min})$, $h_2 = (q_3 - q_2)/(p_3 - p_2)$ and so forth

Just as for the trapezoidal distribution, it is assumed that P is a mixture of $K-1$ state of knowledge independent uncertain data $P_1,\ldots,P_{K-1}$, each with a uniform distribution over the respective interval. The mixture coefficient c_i is the integral of $f(p)$ over the i-th interval, $i = 1, \ldots, K-1$.

Mean value: $$\mathrm{E}\{P\} = \sum_{i=1}^{K-1} c_i \mathrm{E}\{P_i\}$$

Variance: $$\mathrm{Var}\{P\} = \sum_{i=1}^{K-1} c_i \mathrm{Var}\{P_i\} + \sum_{i=1}^{K-1} c_i (\mathrm{E}\{P_i\} - \mathrm{E}\{P\})^2$$

This follows from the relationship $\mathrm{Var}\{P\} = \mathrm{E}\{\mathrm{Var}\{P|P{=}P_i\}\} + \mathrm{Var}\{\mathrm{E}\{P|P{=}P_i\}\}$ where $\mathrm{E}\{P_i\}$ and $\mathrm{Var}\{P_i\}$ are the mean value and variance of the datum P_i over the i-th interval as given above for the uniform distribution.

Median value: See "Monte Carlo simulation" for $u = 0.5$

Monte Carlo simulation: Sample a value u from the (0, 1) uniform distribution (standard subroutine) and transform it into a value p using the transformation $p = \mathrm{F}^{-1}(u)$ where F^{-1} denotes the inverse of the distribution function. In other words, p is that value for P that satisfies $\mathrm{F}(p) = u$.
Find k such that $q_{k-1} \le u \le q_k$,
$p = p_{k-1} + (u - q_{k-1})(p_k - p_{k-1})/(q_k - q_{k-1})$.

It is essential to inspect a graphical representation of the density function in order to judge its suitability as state of knowledge expression. For instance, Fig. 3.34 may suggest that the expert's state of knowledge quantification needs to be modified if

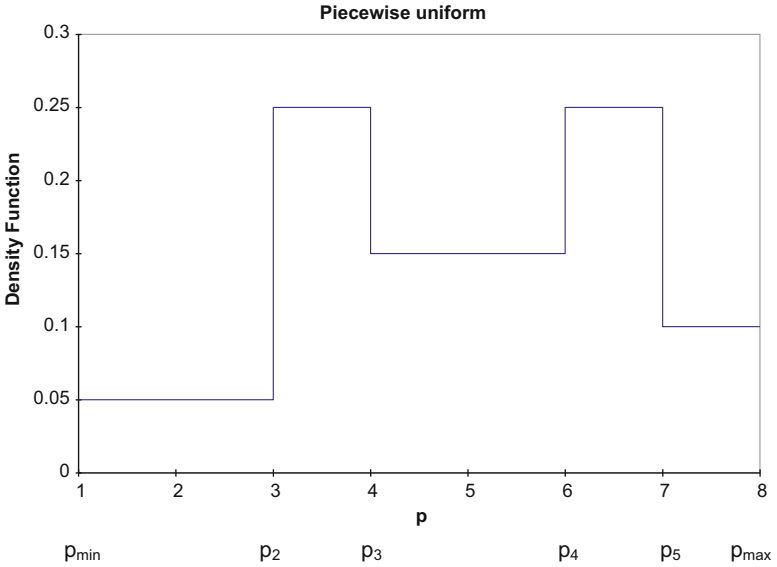


Fig. 3.34 Density function of a piecewise uniform distribution possibly needing modification

there is no reason to assume that the values of the third interval should receive a lower degree of belief than those of the two neighbouring intervals.

– Polygonal line

Sometimes the expert may find it a more adequate expression of his state of knowledge if the density function were sloping linearly between two successive quantiles. This and similar state of knowledge expressions can be modelled by using the polygonal line distribution (see Fig. 3.35).

If the elicitation provided $K-2$ quantiles in addition to p_{min} and p_{max} and with p_{min} the 0% quantile denoted by p_1 and p_{max} the 100% quantile denoted by p_K, the heights $h_2, \ldots, h_{K-1}$ at $p_2, \ldots, p_{K-1}$ are determined from q_k, $\mathrm{k} = 1, \ldots, K$, with q_k the quantile percentages divided by 100.

$$h_1 = 0$$

$$h_2 = 2q_2/(p_2 - p_{min})$$

$$h_k = 2(q_k - q_{k-1})/(p_k - p_{k-1}) - h_{k-1} \quad \text{for k} = 3, \ldots, \mathrm{K} - 1$$

$$h_K = 0.$$

Pdf:

– If $k = 2$ then

$$\mathrm{f}(p) = h_2(p - p_{min})/(p_2 - p_{min})$$

Cdf:

$$\mathrm{F}(p) = h_2(p - p_{min})^2/2(p_2 - p_{min})$$

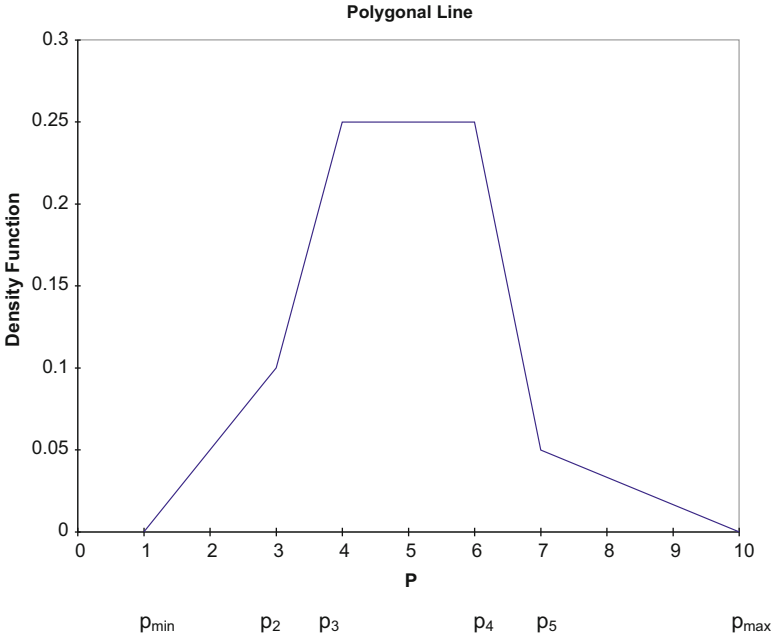


Fig. 3.35 Density function of a polygonal line distribution

for $p_{min} \leq p \leq p_2$;
– If $K > k > 2$ and $h_{k-1} < h_k$ then

$$f(p) = (h_k - h_{k-1})(p - p_{k-1})/(p_k - p_{k-1}) + h_{k-1}$$

$$F(p) = (h_k - h_{k-1})(p - p_{k-1})^2/2(p_k - p_{k-1}) + h_{k-1}(p - p_{k-1}) + F(p_{k-1})$$

for $p_{k-1} \leq p \leq p_k$;
– If $K > k > 2$ and $h_{k-1} = h_k$ then

$$f(p) = h_k$$

$$F(p) = h_k(p - p_{k-1}) + F(p_{k-1})$$

for $p_{k-1} \leq p \leq p_k$;
– If $K > k > 2$ and $h_{k-1} > h_k$ then

$$f(p) = (h_{k-1} - h_k)(p_k - p)/(p_k - p_{k-1}) + h_k$$

$$F(p) = (h_{k-1} - h_k)(p_k - p_{k-1})/2 - (h_{k-1} - h_k)(p_k - p)^2/2(p_k - p_{k-1}) + h_k(p - p_{k-1}) + F(p_{k-1})$$

for $p_{k-1} \leq p \leq p_k$;
– If $k = K$ then

$$f(p) = h_{K-1}(p_{max} - p)/(p_{max} - p_{K-1})$$

$$F(p) = 1 - h_{K-1}(p_{max} - p)^2/2(p_{max} - p_{K-1})$$

for $p_{K-1} \leq p \leq p_{max}$.

Just as for the trapezoidal distribution, it is assumed that P is a mixture of $K-1$ state of knowledge independent uncertain data $P_1, \ldots, P_{K-1}$. The mean value and

variance for the datum P with the state of knowledge expressed by a polygonal line distribution can then be obtained as

Mean value: $$\mathrm{E}\{P\} = \sum_{i=1}^{K-1} c_i \mathrm{E}\{P_i\}$$

Variance: $$\mathrm{Var}\{P\} = \sum_{i=1}^{K-1} c_i \mathrm{Var}\{P_i\} + \sum_{i=1}^{K-1} c_i (\mathrm{E}\{P_i\} - \mathrm{E}\{P\})^2$$

This follows from the relationship $\mathrm{Var}\{P\} = \mathrm{E}\{\mathrm{Var}\{P|P{=}P_i\}\} + \mathrm{Var}\{\mathrm{E}\{P|P{=}P_i\}\}$.

The coefficient c_i is the integral of $f(p)$ over the interval (p_i, p_{i+1}), $i = 1,\ldots, K–1$. In the illustrative example of Fig. 3.35, P_2 and P_4 are themselves considered to be mixtures of two state of knowledge independent uncertain data. For example, P_2 is a mixture of $P_{2,1}$ with triangular distribution ($p_{2,1mod} = p_3$) and of $P_{2,2}$ with a uniform distribution, both over the interval (p_2, p_3) and therefore $\mathrm{E}\{P_2\} = d_{2,\,1}E\{P_{2,\,1}\} + d_{2,\,2}\mathrm{E}\{P_{2,\,2}\}$ with $d_{2,\,1}$ and $d_{2,\,2}$ the mixture coefficients for P_2 alone such that $d_{2,\,1} + d_{2,\,2} = 1$. The coefficient $d_{2,\,1}$ is the area of the upper triangular part divided by d while $d_{2,\,2}$ is the area of the lower rectangular part divided by d and d is the sum of both areas so that

$$\mathrm{Var}\{P_2\} = \sum_{i=1}^{2} d_{2,i} \mathrm{Var}\{P_{2,i}\} + \sum_{i=1}^{2} d_{2,i} (\mathrm{E}\{P_{2,i}\} - \mathrm{E}\{P_2\})^2.$$

P_4 is a mixture of $P_{4,1}$ with triangular distribution ($p_{4,1mod} = p_4$) and of $P_{4,2}$ with a uniform distribution, both over the interval (p_4, p_5) and therefore

$$\mathrm{Var}\{P_4\} = \sum_{i=1}^{2} d_{4,i} \mathrm{Var}\{P_{4,i}\} + \sum_{i=1}^{2} d_{4,i} (\mathrm{E}\{P_{4,i}\} - \mathrm{E}\{P_4\})^2.$$

For the remaining mixture components, $\mathrm{E}\{P_i\}$ and $\mathrm{Var}\{P_i\}$ are as given above for the uniform and triangular distribution;

Mean value: See "Monte Carlo simulation" for $u = 0.5$.

Monte Carlo simulation: Sample a value u from the (0, 1) uniform distribution (standard subroutine) and transform it into a value p using the transformation $p = \mathrm{F}^{-1}(u)$ where F^{-1} denotes the inverse of the distribution function. In other words, p is that value for P that satisfies $\mathrm{F}(p) = u$.

1) Find the two quantile values $q_{k\,-\,1}$ and q_k such that $q_{k\,-\,1} < u \leq q_k$;

2) Solve for p from the interval $[p_{k\,-\,1}, p_k]$:

– If $h_{k\,-\,1} = 0$

$u = h_k(p - p_{k\,-\,1})^2/2(p_k - p_{k\,-\,1}) + F(p_{k\,-\,1})$;

– If $h_{k\,-\,1} < h_k$

$u = (h_k - h_{k\,-\,1})(p - p_{k\,-\,1})^2/2(p_k - p_{k\,-\,1}) +$
$h_{k\,-\,1}(p - p_{k\,-\,1}) + F(p_{k\,-\,1})$;

– If $h_{k\,-\,1} = h_k$

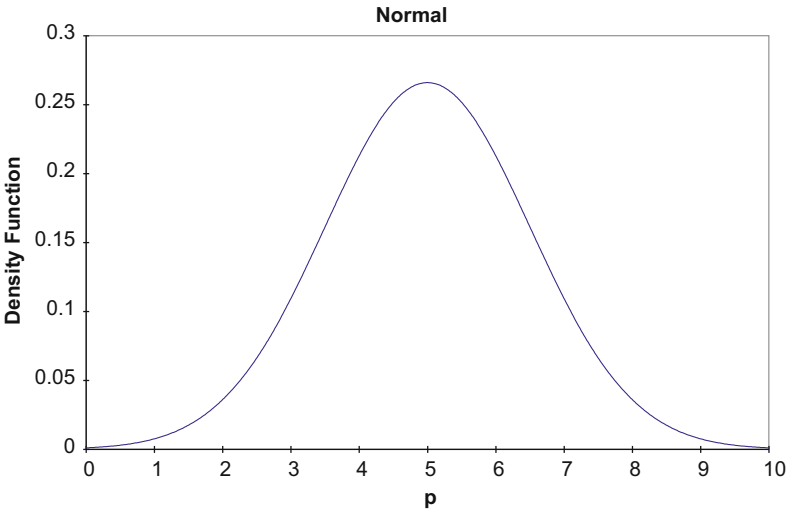


Fig. 3.36 Density function of a normal distribution. The density is shown only over a portion of its carrier

$u = h_k(p - p_{k-1}) + F(p_{k-1})$;
– If $h_{k-1} > h_k$
$u = (h_{k-1} - h_k)(p_k - p_{k-1})/2 -$
$(h_{k-1} - h_k)(p_k - p)^2/2(p_k - p_{k-1}) +$
$h_k(p - p_{k-1}) + F(p_{k-1})$;
– If $h_k = 0$
$u = (h_{k-1}(p_k - p_{k-1})/2 - h_{k-1}(p_k - p)^2/$
$2(p_k - p_{k-1}) + F(p_{k-1})$.

– Normal or Gaussian distribution

This state of knowledge expression says that the expert considers the uncertainty of P to be due to a number of independent additive uncertainty contributors. The central limit theorem suggests the well-known bell shape (see Fig. 3.36) for state of knowledge quantification. The shape is symmetrical about the specified mean value μ and is spread out according to the specified variance σ^2. If two quantiles are specified in the elicitation session, instead of mean value and variance, the corresponding values of mean and variance are computed from these quantiles as shown below.

Pdf:

$$f(p) = \left(1/\sqrt{2\pi}\sigma\right)\exp\left[-(p-\mu)^2/2\sigma^2\right]$$

for $-\infty < p < \infty$ and with μ and σ^2 the specified mean value and variance.

Cdf:

$F(p) = \phi(z)$ with $z = (p - \mu)/\sigma$ and $\phi(z)$ the value of the tabulated standard normal distribution (mean value is

0 and standard deviation is 1) function at z.

Mean value:	$\mathrm{E}\{P\} = \mu$
Variance:	$\mathrm{Var}\{P\} = \sigma^2$
Median value:	$p_{50\%} = \mathrm{E}\{P\}$.

Instead of mean value and variance, the expert may have provided his state of knowledge quantification in the form of two quantiles $p_1 < p_2$, with q_1 and q_2 the quantile percentages and with "normal" as the specified distribution type. Mean value and variance may then be obtained from the quantiles as

$$\mu = p_1 - z_1(p_2 - p_1)/(z_2 - z_1) \text{ and } \sigma = (p_2 - p_1)/(z_2 - z_1)$$

where z_1 and z_2 are the $q_1\%$ and the $q_2\%$ quantiles of the standard normal distribution. The latter are available from standard subroutines or tables.

Note, that the sum of K state of knowledge independent uncertain data, with normal distribution as their state of knowledge expression, has a normal distribution with mean value $\mu_s = \Sigma_{k=1}^{K}\mu_k$ and variance $\sigma_s^2 = \Sigma_{k=1}^{K}\sigma_k^2$ as its state of knowledge expression.

Monte Carlo simulation:	Sample a value z from the standard normal distribution (standard subroutine) and transform it into a value p using the transformation $p = \mu + z\sigma$.

The set of possibly true values for P will in most practical situations not extend to infinity but will rather be limited from below at a value p_{min} and from above at a value p_{max}, both specified in the elicitation session. The normal distribution will, therefore, need to be truncated at these values.

Figure 3.37 shows the density function of the normal distribution in Fig. 3.36 after truncation at $p_{min} = 2.0$ and $p_{max} = 8.0$:

Monte Carlo simulation:	Sample a value z from the standard normal distribution (standard subroutine) and transform it into a value p using the transformation $p = \mu + z\sigma$ and discard any values p outside the specified limits. The variance and, if the truncation is not symmetric, the mean value of the truncated distribution will differ from σ and μ. The quantiles after truncation will not agree with those specified and used in the calculation of μ and σ. The software for uncertainty analysis (Kloos 2015) searches for a normal distribution that is truncated at p_{min} and at p_{max} and complies, after truncation, with the pair of specified quantiles or, as closely as requested by the user, with the specified mean value and variance.

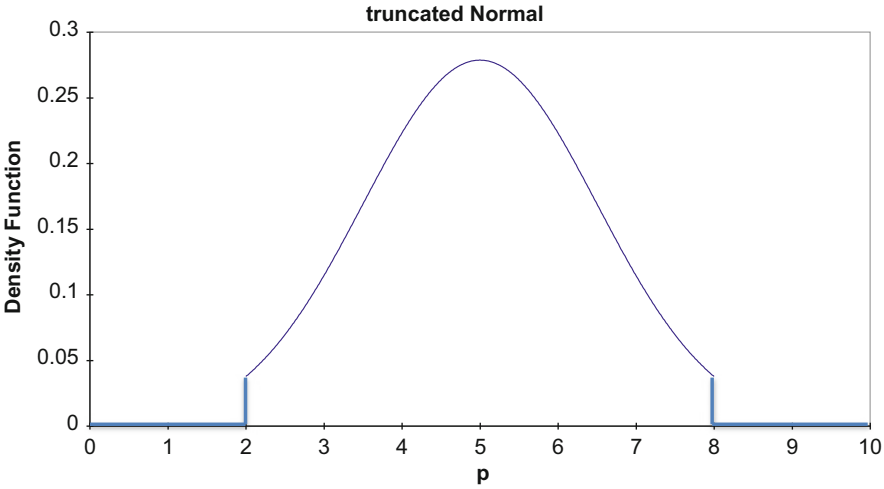


Fig. 3.37 Density function of the normal distribution in Fig. 3.36 after truncation

– t- (or Student) distribution

If the uncertain datum P is the mean value of a population of values and a set of values $p_1, \ldots, p_K$, sampled at random from this population, is available, then

$$\bar{p} = \left(\frac{1}{K}\right) \sum_{k=1}^{K} p_k$$

is an unbiased estimate of P and the state of knowledge of P can be quantified with the help of the t-distribution with $K-1$ degrees of freedom (see Fig. 3.38a) as was shown in Sect. 3.3.1.1.

Pdf:
(for K degrees of freedom):
$f(p) = \Gamma\left(\frac{K+1}{2}\right)$
$/\left\{(1+p^2/K)^{\frac{K+1}{2}}(K\pi)^{\frac{1}{2}}\Gamma\left(\frac{K}{2}\right)\right\}$
for $-\infty < p < \infty$.

Cdf:
$F(p)$ from Tables
$\Gamma(x)$ is the value of the Gamma function at x.

Mean value:	$E\{P\} = 0$ (does not exist for $K = 1$)
Variance:	$\mathrm{Var}\{P\} = K/(K-2)$ (exists only for $K > 2$)
Median value:	$p_{50\%} = E\{P\}$.

Monte Carlo simulation:	A straightforward but possibly not the most efficient way would be to sample independently $K+1$ values z according to

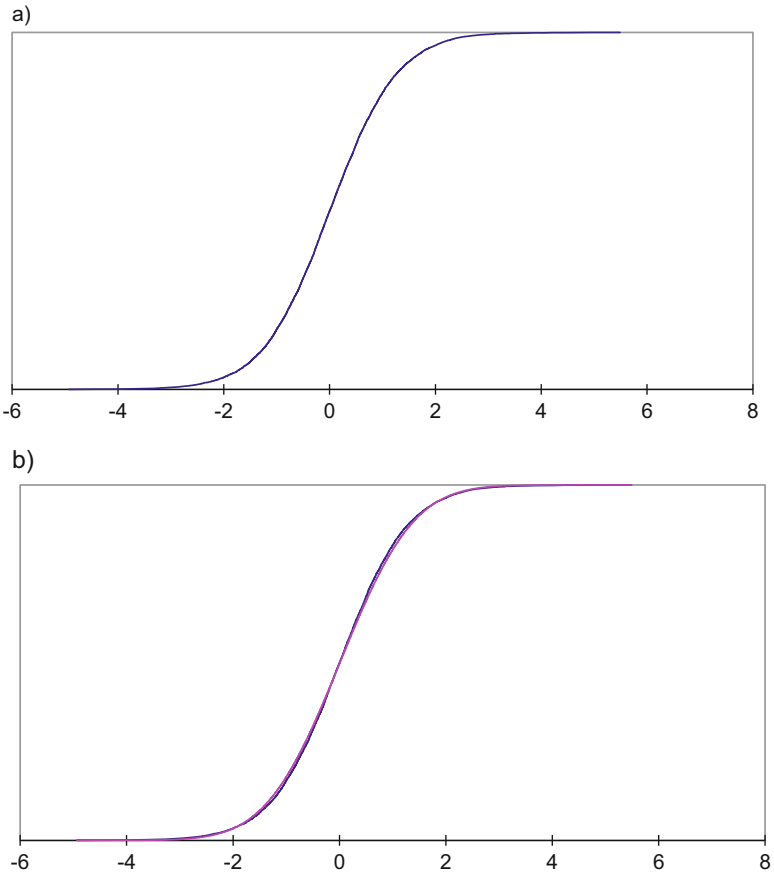


Fig. 3.38 (**a**) Cumulative distribution function of the t-distribution with 13 degrees of freedom. (**b**) Normal distribution fitted to the t-distribution in (**a**)

the standard normal distribution (standard subroutine). Multiply the first by the square root of K and divide the result by the square root of the sum of the squares of the K others to obtain a sample value from the t-distribution of K degrees of freedom. For other ways of generating random samples according to a Student distribution, see Fishman (2000).

Figure 3.38b shows a Normal distribution fitted to the t-distribution in Fig. 3.38a.

– Beta distribution

The family of beta distributions offers a wide variety of probabilistic expressions for the state of knowledge of an uncertain datum P with the range of possibly true values restricted to a finite interval.

The density function of the Beta distribution is defined over the interval (0, 1) as

$$g(y) = \left(\frac{1}{B(\alpha,\beta)}\right) y^{\alpha-1}(1-y)^{\beta-1} \text{ with } \alpha,\beta > 0 \text{ and } 0 < y < 1.$$

$B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$ and $\Gamma(x)$ is the value of the Gamma function at x.

To use the probabilistic modelling flexibility of the Beta distribution for the state of knowledge of P over (p_{min}, p_{max}), the following linear transformation is performed:

$P = p_{min} + (p_{max} - p_{min})Y$ with $0 \le p_{min} < p < p_{max}$ which leads to the density and distribution function given below.

Pdf:
$f(p) = [(p - p_{min})^{\alpha-1}(p_{max} - p)^{\beta-1}]/[B(\alpha,\beta)(p_{max} - p_{min})^{\alpha+\beta-1}]$

Cdf:
$F(p) = G[(p - p_{min})/(p_{max} - p_{min})] = G(y)$
where G(y) is the Beta distribution with the same values of α and β but over the interval (0, 1). G(y) is taken from tables possibly using the relationship $Y = mV_{m,n}/(n + mV_{m,n})$ with $V_{m,n}$ following the F-distribution with (m, n) degrees of freedom and $\alpha = m/2$, $\beta = n/2$.
$p_{min} < p < p_{max}$

Mean value: $E\{P\} = p_{min} + (p_{max} - p_{min})\alpha/(\alpha+\beta)$
Variance: $\text{Var}\{P\} = (p_{max} - p_{min})^2\alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$
Median value: $p_{50\%} = p_{min} + (p_{max} - p_{min})y_{50\%}$ with $y_{50\%}$ from ables for G(y).

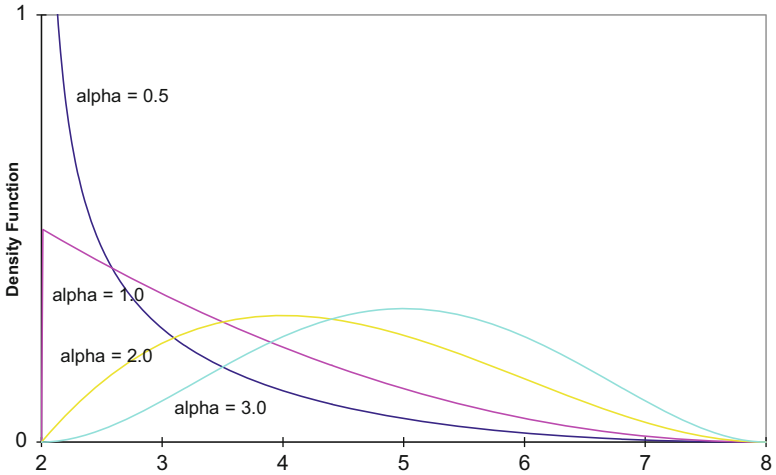


Fig. 3.39 Density functions for P with $\beta = 3.0$ and varying values of α

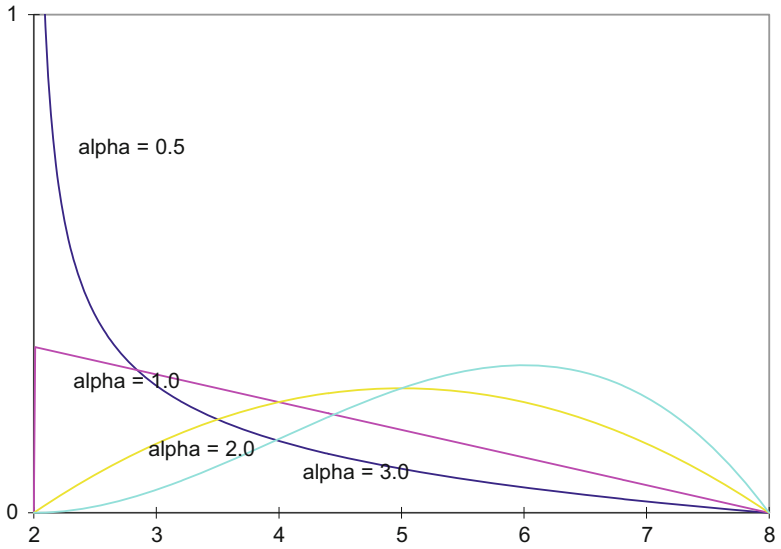


Fig. 3.40 Density functions for P with $\beta = 2.0$ and varying values of α

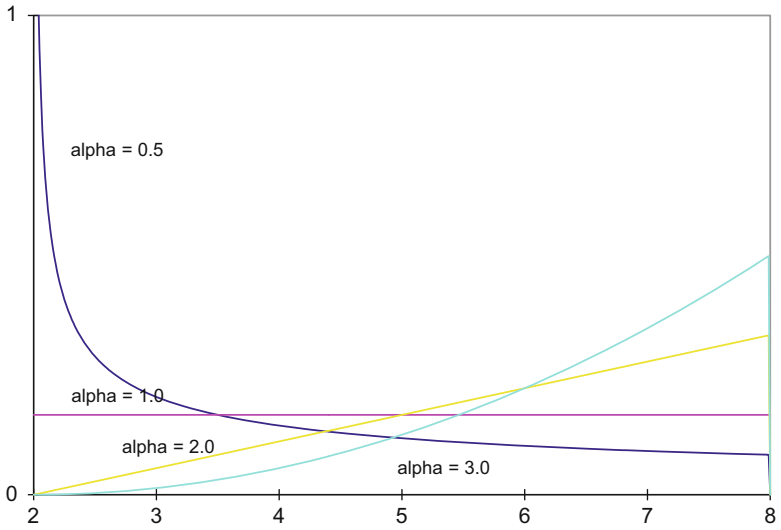


Fig. 3.41 Density functions for P with $\beta = 1.0$ and varying values of α

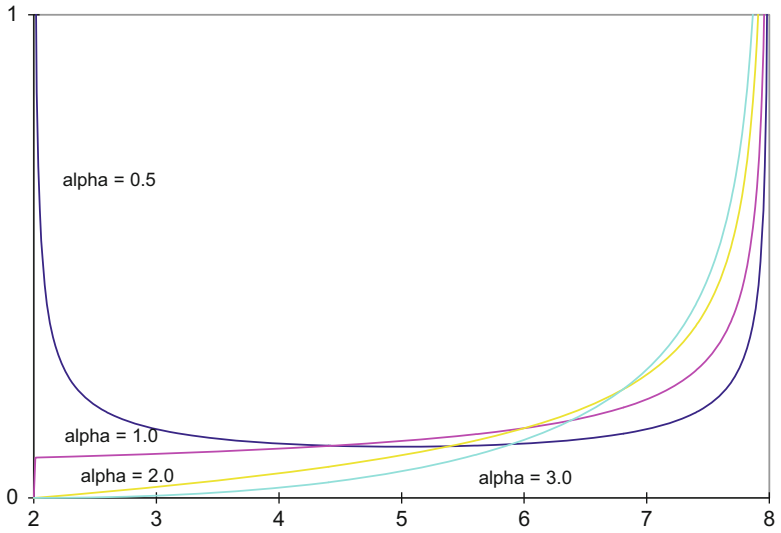


Fig. 3.42 Density functions for P with $\beta = 0.5$ and varying values of α

Monte Carlo simulation: See Fishman (2000)

The Figs. 3.39, 3.40, 3.41 and 3.42 show some of the infinitely many shapes offered by the family of beta density functions after the linear transformation $P = p_{min} + (p_{max} - p_{min})Y$ of the Beta-distributed Y and with $p_{min} = 2$, $p_{max} = 8$.

– Loguniform distribution

This type of distribution (see Figs. 3.43 and 3.44) is chosen if the expert considers only positive values as possibly true for P and the uncertainty to be so large that he is only able to quantify his state of knowledge for orders of magnitude between a lowest order of magnitude $\log_{10}(p_{min})$ and a largest order of magnitude $\log_{10}(p_{max})$ while the degrees of belief are equally spread between these limits.

Pdf:
$f(\log_{10}(p)) = 1 / [\log_{10}(p_{max}) - \log_{10}(p_{min})]$

Cdf:
$F(\log_{10}(p)) = [\log_{10}(p) - \log_{10}(p_{min})] / [\log_{10}(p_{max}) - \log_{10}(p_{min})]$

Using the relationship $\log_{10}(x)\ln(10) = \ln(x)$:

Pdf:
$f(p) = 1 / \log_{10}(p_{max} / p_{min}) p \ln(10)$
$0 < p_{min} \leq p \leq p_{max}$

Cdf:
$F(p) = \log_{10}(p / p_{min}) / \log_{10}(p_{max} / p_{min})$

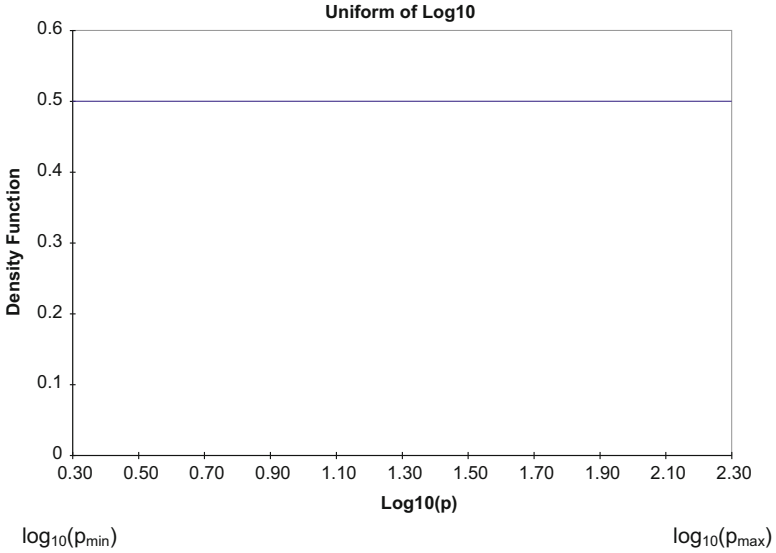


Fig. 3.43 Density function over $\log_{10}(P)$ of a loguniform distribution for P

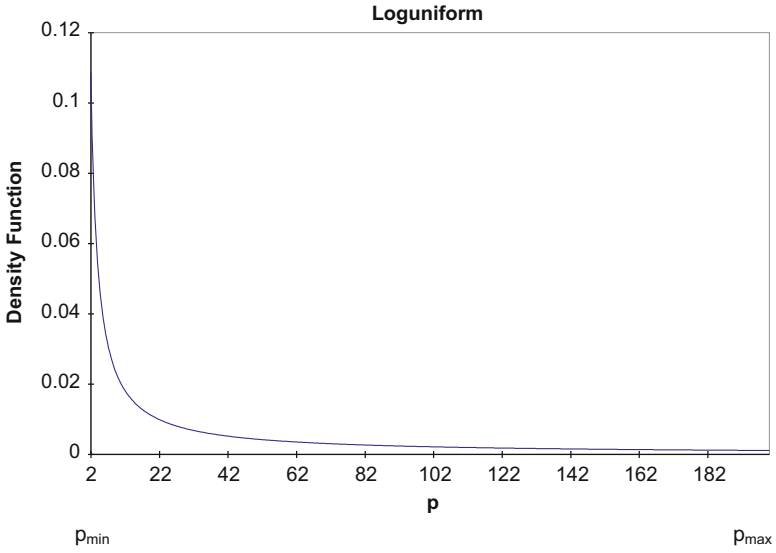


Fig. 3.44 Density function of a loguniform distribution for P shown over part of its carrier

Mean value: $\mathrm{E}\{P\} = (p_{max} - p_{min}) / \log_{10}(p_{max} / p_{min})\ln(10)$

Variance: $\mathrm{Var}\{P\} = (p_{max}{}^2 - p_{min}{}^2) / 2\log_{10}(p_{max} / p_{min})\ln(10) - (p_{max} - p_{min})^2 / [\log_{10}(p_{max} / p_{min})\ln(10)]^2$

Median value: $p_{50\%} = (p_{min}p_{max})^{1/2}$.

Monte Carlo simulation: Sample a value u from the (0, 1) uniform distribution (standard subroutine) and transform it into a value p using the transformation $p = \exp\{[\log_{10}(p_{min}) + u\log_{10}(p_{max}/p_{min})]\ln(10)\}$.

Note that a uniform distribution for P over the range of values (0, 1000] would assign subjective probability 0.01 to the range of values (0, 10] and subjective probability 0.9 to the range of values (100, 1000]. If there is no information that would justify this preference for the upper order of magnitude, i.e. if the state of knowledge is only such that each of the three ranges (0, 10], (10, 100] and (100, 1000] are considered to equally likely contain the true value of P, the uniform distribution would favour the upper magnitude out of proportion.

– Lognormal distribution

Choosing this distribution as state of knowledge expression says that the expert considers only positive values to be possibly true for the uncertain datum P and that the uncertainty is due to a number of independent multiplicative uncertainty contributors all of which have only positive values. Following the central limit theorem, the state of knowledge of $\ln(P)$ is then approximately expressed by the normal distribution (see Fig. 3.45) with specified mean value μ and according to the specified variance σ^2. The state of knowledge expression for P is a lognormal distribution (see Fig. 3.46). If instead of μ and σ of the normal distribution of $\ln(P)$ two quantiles of the lognormal distribution for P are specified in the elicitation

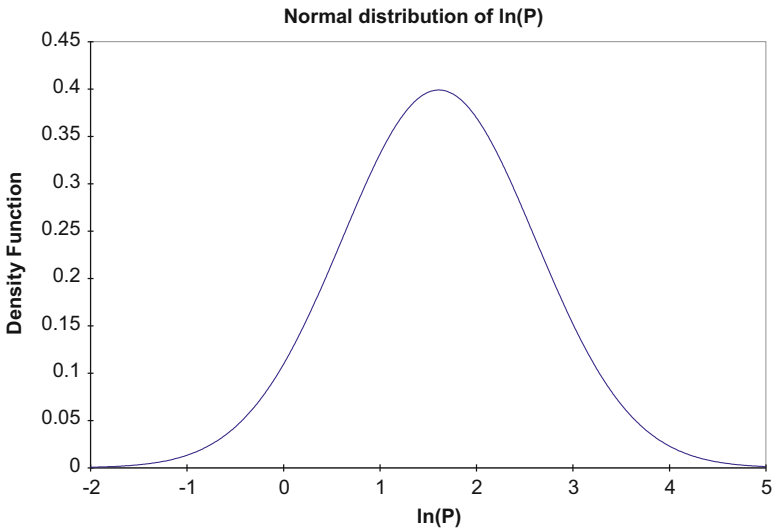


Fig. 3.45 Normal density function of $\ln(P)$ shown over part of its carrier

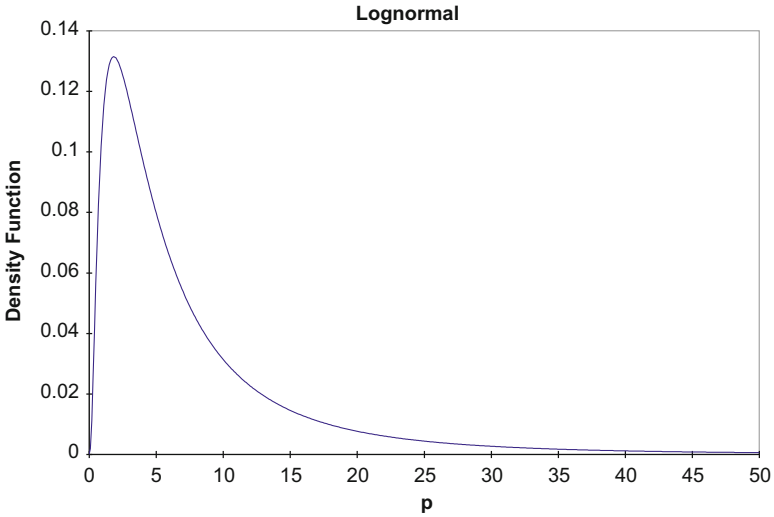


Fig. 3.46 Density function of a lognormal distribution for P shown over part of its carrier

session, then the corresponding values of μ and σ are computed from these quantiles as shown below.

Pdf:
$$f(\ln(p)) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)\exp\left[-(\ln(p)-\mu)^2/2\sigma^2\right]$$

Cdf:
$F(\ln(p)) = \phi(z)$ with $z = (\ln(p)-\mu)/\sigma$ and $\phi(z)$ is the value of the tabulated standard normal distribution function at z.

for $0 < p < \infty$.

With μ the mean value and σ the standard deviation of the normal distribution for $\ln(P)$, the density and distribution function for P are

Pdf:
$$f(p) = \left(1/\sqrt{2\pi}\sigma p\right)\exp\left[-(\ln(p)-\mu)^2/2\sigma^2\right]$$

Cdf:
$F(p) = \phi(z)$ with $z = (\ln(p)-\mu)/\sigma$ and $\phi(z)$ is the value of the tabulated standard normal distribution function at z.

for $0 < p < \infty$
$f(p) = 0$ for $p \leq 0$.

Mean value: $E\{P\} = \exp(\mu + \sigma^2/2)$
Variance: $\mathrm{Var}\{P\} = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$
Median value: $p_{50\%} = \exp(\mu)$
Modal value: $\exp(\mu - \sigma^2)$
The density function assumes its maximum at the modal value.

The term $\exp(\mu)$ is called the geometric mean and $\exp(\sigma)$ the geometric standard deviation of the state of knowledge expression for P.

If the elicitation session provided p_q, the q% quantile, and $p_{(100-q)}$, the $(100-q)$% quantile, then parameter values μ and σ are obtained from the two symmetrical (in percentage) quantiles as:

$$\mu = \left[\ln\left(p_q\right) + \ln\left(p_{100-q}\right)\right]/2 \text{ and } \sigma = \left[\ln\left(p_{100-q} - \mu\right]/z_{100-q}\right.$$

where z_{100-q} is the $(100-\text{q})$% quantile of the standard normal distribution. The value of z_{100-q} is available from standard subroutines and tables.

The product of K uncertain state of knowledge independent uncertain data with their states of knowledge expressed by lognormal distributions has again a lognormal distribution as its state of knowledge expression. The parameters μ and σ of the lognormal distribution for the product are obtained as $\mu = \Sigma_{k=1}^{K}\mu_k$ and $\sigma = \left(\Sigma_{k=1}^{K}\sigma_k^2\right)^{\frac{1}{2}}$.

Monte Carlo simulation: Sample a value z from the standard normal distribution (standard subroutine) and transform it into a value p using the transformation $p = \exp(\mu + z\sigma)$.

– Discrete distribution

If there are only K discrete values (see Fig. 3.47) that could be possibly true for the uncertain datum P, then the expert's state of knowledge quantification needs to

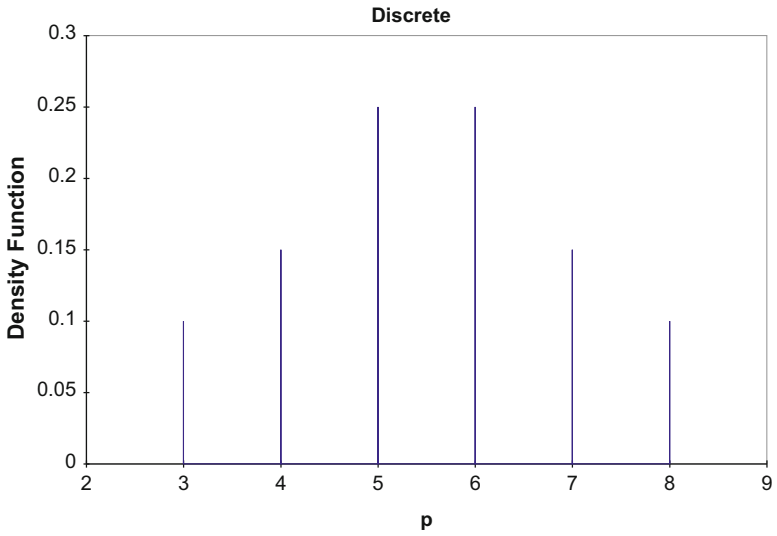


Fig. 3.47 Density function of a discrete distribution for P with $K = 6$ possibly true values

specify the set of K values $p_1, \ldots, p_K$ considered as possibly true together with the associated subjective probabilities $q_1, \ldots, q_K$.

Pdf:	Cdf:	
$f(p_1) = q_1$	$F(p) = \sum_{i=1}^{k} q_i$	for $p_k \le p < p_{k+1}$
.		$k = 1, \ldots, K-1$
.	$F(p) = 0$	for $p < p_1$
$f(p_K) = q_K$	$F(p) = 1$	for $p \ge p_K$.
$p_1 < p_2 < \ldots < p_K$.		

Note, that the density function could be a discrete approximation to any continuous distribution with the number and positions of the discrete values chosen such that the approximation is satisfactory.

Mean value:	$E\{P\} = \sum_{k=1}^{K} p_k q_k$
Variance:	$Var\{P\} = \sum_{k=1}^{K} (p_k - E\{P\})^2 q_k$
Median value:	See "Monte Carlo simulation" for $u = 0.5$. However, the value $F(p)$ at the obtained value p will be larger than 0.5 in most situations.
Monte Carlo simulation:	Sample a value u from the (0, 1) uniform distribution (standard subroutine) and find k such that $F(p_{k-1}) < u \le F(p_k)$, $p = p_k$.

Guidance for the probabilistic modelling of the state of knowledge of uncertain data may also be found in Stephens et al. (1993) and Joint Committee for Guides in Metrology (2008).

3.6.1.4 Specific Values for Uncertain Data

Any single value p of P with subjective probability density $f(p) > 0$ is called a "point value" for P. This is to emphasize the fact that it is only one value out of a population of possibly true values. Specific point values are:

– Best estimate value

The mean value of the subjective probability distribution is usually considered to be the best estimate for P. The integral of the product "p times the density value $f(p)$" taken over the values of P below the mean value is equal to the integral of this product taken over the values of P above the mean value. In this sense, the mean value is balanced with respect to the subjective probability weighted contents of possibly true values on either side.

– Reference value

The median value of the subjective probability distribution is usually considered to be the reference value for P. The integrals of the density function f(p) taken over the values of P on either side of the median value are equal. The median is balanced with respect to the subjective probability content on either side.

– Most likely or preferred value

The modal value of the subjective probability density function is the value for P that is believed to be most likely true. The subjective probability density function has its maximum at the modal value. Quite frequently, the specified subjective probability density function does not have a single modal value but is flat on top with all values from a finite interval receiving the highest subjective probability density.

– q% value

This is the q%-quantile of the subjective probability distribution. The true value of P is believed to be below or equal to this value with subjective probability q/100.

– Extreme values

The specified smallest possibly true value and the specified largest possibly true value delimit the range of all values of P that are thought to be possibly true. The subjective probability for values outside this range is zero.

3.6.1.5 The Documentation Sheet

The state of knowledge document is discussed for data uncertainties only since model uncertainties are represented by uncertain parameters (see Sect. 3.4), and uncertain parameters are categorized as uncertain data in order to simplify matters.

There is one document each per uncertain datum. Each document consists of three parts:

– Cover sheet
– State of knowledge quantification and probabilistic modelling sheet
– State of knowledge dependence information sheet

The **cover sheet** contains:

- Number of the uncertain datum.
  The uncertain data are successively numbered from 1 to M where M is the total number of uncertain data considered in the uncertainty analysis of the application of the computer model.
- Full name of the uncertain datum.
  This is a verbal explanatory description of the datum.
- Is the uncertain datum already part of the encoded computer model (yes/no)?

- Is the uncertain datum already part of the input file of the computer model (yes/no)?
- Short name of the uncertain datum.
  This is the name (to be) used in the encoded computer model.
- Symbol for the uncertain datum.
  This is the symbol used in the equations of the mathematical model.
- Unit of the uncertain datum (if any).
  This is the physical (or other) unit of the uncertain datum used in the encoded computer model.
- Probabilistic state of knowledge (sok) model for the uncertain datum.
  - The type of the subjective probability distribution.
  - The values of the distribution parameters.
  - Values at which the distribution is truncated (if any).
- State of knowledge dependence with
  Numbers of the uncertain data that are found to be state of knowledge dependent with the uncertain datum of this sheet.
- Notes.
  - Caveats to be remembered when interpreting the analysis results.
  - Any additional explanatory and supplementary remarks provided by the substantive expert.
  - Any additional explanatory and supplementary remarks provided by the analyst.
  - Status of the state of knowledge quantification (preliminary/exploratory/state of the art)?

The **state of knowledge quantification and probabilistic modelling sheet** contains:

- Number of the uncertain datum
  (as on the cover sheet)
- Short name of the uncertain datum
  (as on the cover sheet)
- Unit of the uncertain datum
  (as on the cover sheet)
- Name and affiliation of the substantive expert
- Detailed description of the uncertain datum and of its role in the application of the computer model (provided by the client)
- Full text of the elicitation questions (provided by the analyst)
- The state of knowledge quantifications elicited from the substantive expert
- The substantive expert's arguments in support of his quantifications including any computations and caveats
- The subjective probability distribution fitted to the substantive expert's state of knowledge quantifications

  - Type, parameters and truncations (if any)
  - Plot of the density and of the distribution function

- References (supporting documents)
  The references should contain a list of all the information sources used by the expert in the course of the state of knowledge quantification. These may range from observations permitting a statistical estimate of the uncertain datum to the fitting of a model to experimental results, or to common sense plausibility considerations. Within this range are: Practical experience, literature studies, extrapolations from related fields, laboratory experiments, field tests, theoretical models and interview data from other experts or persons intimately familiar with the uncertain datum
- Notes
  - Caveats
  - Status of the state of knowledge quantification (preliminary/exploratory/state of the art)?
- Date and signatures of the substantive expert, the analyst and the client.

The **state of knowledge dependence information sheet** contains:

- Number of the uncertain datum
  (as on the cover sheet)
- Short name of the uncertain datum
  (as on the cover sheet)
- Unit of the uncertain datum
  (as on the cover sheet)
- Name and affiliation of the substantive expert
- Full text of the explanation of state of knowledge dependence provided by the analyst;
- Numbers of the uncertain data that are found to be state of knowledge dependent to the uncertain datum of this sheet
- For each datum listed above:
  Either
  - the number where state of knowledge dependence is already specified

  or
  - full text of the elicitation question posed by the analyst in order to identify and quantify state of knowledge dependence between the pair of uncertain data
  - Source of the state of knowledge dependence as seen by the substantive expert
  - Type and details of the dependence modelling chosen
  - Scatterplot of the dependence modelling chosen
- References (supporting documents)
- Notes
  - Caveats

Table 3.11 Example of a cover sheet

COVER SHEET	
Number	4
Full name	Loss from cohort 1 due to predation by non-bird-enemies[a], given as fraction of the total loss L (see no. 10)
Datum is part of the model	Yes
Datum is part of the input file	Yes
Short name	LOSS1
Symbol	l_1
Unit	L
Probabilistic state of knowledge model	(0.7, 0.9) uniform
State of knowledge dependent with number(s)	5, 6, 7, 8, 9
Notes	Preliminary state of knowledge quantification

[a]Excluding fishing activities

  - Status of the state of knowledge dependence quantifications (preliminary/exploratory/state of the art)?

- Date and signatures of the involved substantive expert, the analyst and the client (Tables 3.11, 3.12 and 3.13).

3.6.2 State of Knowledge Elicitation and Probabilistic Modelling for Model Uncertainties

It is the substantive expert's task to provide a subjective probability distribution that expresses his state of knowledge for the computational representation of a process or state of affairs. To perform this task, the expert first needs to determine the category of the model uncertainty. Four categories may be distinguished:

– Category 1

It is uncertain how to represent the process or state of affairs because it is unknown which representation out of a finite and exhaustive set of mutually exclusive possibly true alternatives is actually true. Specifically, there are only two alternatives if it is uncertain whether to include or exclude a known phenomenon.

– Category 2

There are infinitely many different ways of simplifying the representation of a continuum or of a very large set of data that cannot practically be represented in full.

– Category 3

Table 3.12 Example of a state of knowledge quantification and probabilistic modelling sheet

STATE OF KNOWLEDGE QUANTIFICATION AND PROBABILISTIC MODELLING SHEET	
Number	4
Short name	LOSS1
Unit	L
Name and affiliation of the expert	A. Fisher, Fisheries Board
Detailed description of the datum	Quantifies the loss (per period of 4 months), of the youngest anchovies as fraction of the total loss (see no. 10), due to enemies other than guano-birds and fishing activities.
Elicitation questions	a) Which is the smallest value at which you are 100% sure that the true value of this fraction does not lie above? Which is the largest value at which you are 100% sure that the true value of this fraction does not lie below? b) Do you see any reason, or are you aware of any information, that would enable you to prefer any value between the minimum and the maximum possibly true value (given under a)) of this fraction?
State of knowledge quantifications by the expert	The fraction is not less than 0.7 and not more than 0.9 of the total loss (measured in MTs of biomass)—no preference for any values between these limits.
Expert's argumentation	The Youngest are highly vulnerable to predation by non-bird enemies—limits are preliminary.
Fitted subjective probability distribution	(0.7, 0.9) uniform
References	None
Notes	Preliminary state of knowledge quantification
Date and signatures	Oct. 21, 2011

There are an infinite number of different ways of fitting a function to observations or there are an infinite number of different mathematical expressions based on theoretical and/or plausibility considerations that could serve as an approximate representation.

– Category 4

The process or state of affairs is declared an issue with no model existing.

3.6.2.1 Elicitation for Category 1

The model uncertainty reduces to a parameter uncertainty. The parameter P can assume only a finite number of integer values. These values are the index numbers assigned to the alternatives. In the case of a phenomenon, the relevance of which is

Table 3.13 Example of a state of knowledge dependence information sheet

STATE OF KNOWLEDGE DEPENDENCE INFORMATION SHEET	
Number	4
Short name	LOSS1
Unit	L
Name and affiliation of expert	A. Fisher, Fisheries Board
Meaning of state of knowledge dependence	The state of knowledge for no. 4 is dependent if it is different for different values thought to be true for any of the other uncertain data.
Numbers of state of knowledge dependent uncertain data	5, 6, 7, 8, 9
Details of dependence:	
a) Reason or source of dependence	Fractions of the total loss have to add up to 1.
b) Type of dependence modelling	State of knowledge for conditional fractions
References	None
Notes	"State of the art" dependence modelling
Date and signatures	Oct. 25, 2011

uncertain, *P* can assume only two index numbers, namely "one" for inclusion of the phenomenon and "zero" for exclusion from the computer model application.

The first step of the work plan for state of knowledge quantification is analogous to the description given in Sect. 3.6.1.1 for an uncertain datum *P*. In the second step, the expert is asked to give his subjective probabilities for the truth of each of the alternatives. For each of the alternatives, the expert is asked as follows:

> "Considering the information identified in the first step, which is your subjective probability for this alternative to be true?"

The expert is asked to supplement his answer by his arguments for the choice of probability values. His arguments are to be documented together with the result of the elicitation. The probability values given by the expert have to add up to 1 given the alternatives are mutually exclusive and the set of alternatives is exhaustive. If they do not add up to 1, normalization may help unless the expert comes up with a correction supported by arguments that are to be documented. The third step of the work plan, as described for an uncertain datum, is obsolete as the subjective probability distribution for *P* is discrete. The subjective probabilities for the elements of the finite set of index values (assigned to the alternatives) are already specifying the subjective probability distribution.

3.6.2.2 Elicitation for Category 2

The model uncertainty reduces to a parameter uncertainty. The parameter *P* can assume only a finite number of integer values. These values are the index numbers assigned to the finite number of alternatives considered.

The expert has to make a decision as to how many and which representations to include in the set. After this decision is made, the search for information on the alternative simplified representations proceeds very much as described in Sect. 3.6.1.1 for an uncertain datum *P*. In the first step, the expert has to specify the set of simplified representations. All the information available about the elements of this set needs to be collected. This includes the experience gained with simplified representations, in particular their success in representing the continuum or large set sufficiently well for problems to be solved by the computer model. In the second step, the expert is asked to give his subjective probability for each of the chosen representations to come sufficiently close to the full representation of the continuum or large set. For each of the representations, the expert is asked:

> "Considering the information identified in the first step, which is your subjective probability for this simplified representation to come sufficiently close to the full representation?"

The expert is asked to supplement his answer by his arguments for the choice of probability values. His arguments are to be documented together with the result of the elicitation. The subjective probabilities will most likely not add up to 1. Normalization will in this case be required. Again, the third step is obsolete as the subjective probability distribution for *P* is discrete. The probability values for the elements in the finite set of index values (assigned to the representations) are already specifying the distribution. This model uncertainty needs to be earmarked for a caveat that is in place when interpreting the analysis results. Reason is that the question could not ask for possibly true alternatives but for alternatives that come sufficiently close to the truth.

3.6.2.3 Elicitation for Category 3

The model uncertainty reduces to a parameter uncertainty in either of three ways:

- There is either only one model formulation available or the alternative that is judged to come closest to the true representation is chosen. An uncertain parameter (or a set of parameters) is defined as the value(s) of a correction term that is possibly needed so that the output of the model formulation may be used in the computer model application. The elicitation of the state of knowledge for the correction parameter(s) follows then the procedure described in Sect. 3.6.1.1 for an uncertain datum *P*.
- There are two extreme models, and the uncertainty is expressed by forming a weighted average of the output from both models with the weight now being a new uncertain parameter *P*. Elicitation proceeds as described in Sect. 3.6.1.1 for an uncertain datum *P*.
- There is a finite number of alternative functions fitted to observations or a finite number of alternative mathematical expressions based on theoretical and/or plausibility considerations that could be thought of as coming sufficiently close

to the true representation. P assumes then only a finite number of integer values, namely the index numbers assigned to the alternatives.

The second step is as for category 2 and the third step is again obsolete as the subjective probability distribution for P is discrete with the subjective probabilities for the elements of the finite set of index values already specifying the subjective probability distribution.

3.6.2.4 Elicitation for Category 4

In this case, the state of knowledge is directly elicited for the value of the quantity required in the computer model application. To this end, it may be appropriate to consider an exhaustive set of cases, the so-called case structure, and to elicit the state of knowledge from the substantive expert in each of the cases. State of knowledge dependence across cases may need to be considered. For each case, the elicitation proceeds as described in Sect. 3.6.1.1 for an uncertain datum P. The expert quantifies his state of knowledge by a subjective probability distribution for the value in question. Often the subjective probability distributions are obtained from a decomposition of the issue into a sequence of sub-issues that may be more amenable to state of knowledge quantification. The state of knowledge for the actual issue result is then obtained by propagation of the states of knowledge for the sub-issues through this decomposition. The argumentation for the chosen case structure as well as for any decomposition into sub-issues and for the state of knowledge quantifications needs to be documented. Different experts may choose different issue decompositions and may arrive at different state of knowledge quantifications (see Sect. 3.7).

3.6.3 Elicitation for State of Knowledge Dependence

The elicitation session closes with the question whether state of knowledge dependence needs to be taken into account. If the expert has provided his state of knowledge quantifications for two or more uncertain data (model uncertainties are represented by uncertain parameters, as was shown in Sect. 3.4, and uncertain parameters are categorized as data uncertainties for the sake of simplicity), the analyst needs to explain the concept of state of knowledge dependence. He will do so with the help of an example of two uncertain data P_i and P_j. The explanation will show that in the case of dependence, the state of knowledge of P_j will vary with the value that is thought to be true for P_i. He will then go through all pairs of uncertain data for which the expert provided his state of knowledge and will ask whether the state of knowledge quantification given for one member of the pair remains valid irrespective of the value that is thought to be true for the other member. If the expert indicates for one such pair that the assumption of independence is not justified, the analyst will explain the ways available for state of knowledge dependence quantification. This explanation will include the many possibilities presented in Sect. 3.5

except for Sect. 3.5.2.3 since any known functional relationship will have most likely been specified already.

For any unknown monotone functional relationship (Sect. 3.5.2.6), that the expert might think exists between the two uncertain data, he needs to indicate the direction namely whether it is increasing (large values are true for P_j together with large values for P_i) or decreasing (large values are true for P_j together with small values for P_i).

If the range of values, that are possibly true for P_j, is restricted by a function of the value that is thought to be true for P_i (Sect. 3.5.2.2), then the expert needs to specify this function.

If none of these clear-cut situations apply, then the expert is shown how to represent state of knowledge dependence by conditional subjective probability distributions (Sect. 3.5.2.1). Should he choose to use this means of quantification, then the expert has to decide about an exhaustive set of disjoint intervals covering the range of possibly true values of the so-called free datum P_i. The elicitation continues with the conditional state of knowledge quantification (as described in Sect. 3.6.1) for P_j under the condition that the true value of P_i lies in the respective interval. Instead of conditional distributions for P_j to each of the finite number of intervals for P_i, the expert may be able to specify a distribution type for the conditional state of knowledge of P_j and to express the values of its parameters as continuous functions of the value that is thought to be true for P_i (see Fig. 3.6).

If none of the above options apply or are feasible, then the expert may specify a measure of association (Sects. 3.5.2.4 and 3.5.2.5) as quantitative expression of state of knowledge dependence, rather than conditional distributions. In this case, the analyst will put the following question to the expert:

Given the true value p_i of P_i lies somewhere below $p_{i,\ 50\%}$ (i.e. $\mathrm{F_i}(p_i) < 0.5$), what would be your subjective probability for the true value p_j of P_j to also lie below $p_{j,50\%}$ (i.e. $\mathrm{F_j}(p_j) < 0.5$)?[5]

The provided subjective probability is

$$\mathrm{sw}\left(P_j < p_{j,50\%} | P_i < p_{i,50\%}\right) = \left(b_{ij} + 1\right)/2$$

with b_{ij} known as Blomquist's quadrant measure (Kruskal 1958).

If P_i and P_j are transformed as follows:

$$Z_i = \phi^{-1}(\mathrm{F_i}(P_i)) \text{ and } Z_j = \phi^{-1}\left(\mathrm{F_j}\left(P_j\right)\right)$$

where ϕ^{-1} is the inverse of the standard normal distribution function, then b_{ij} also applies to Z_i and Z_j and may be transformed into the Pearson correlation coefficient ρ_{ij} of Z_i and Z_j as follows (Kruskal 1958)

[5] If his judgment is such that there is no state of knowledge dependence, then he should give 0.5 as his subjective probability. If he judges the data to be completely state of knowledge dependent, then his subjective probability should be 1 (positive complete dependence) or 0 (negative complete dependence).

$$\rho_{ij} = \sin\left(b_{ij}\pi/2\right).$$

The corresponding Spearman rank correlation coefficient is obtained through the relationship (Kruskal 1958)

$$\rho_{ij}^{S} = (6/\pi)\arcsin\left(\rho_{ij}/2\right).$$

Back transformation of Z_i and Z_j into P_i and P_j leaves ρ_{ij}^{S} unchanged so that the expert has indirectly specified a Spearman rank correlation coefficient as association measure quantifying his judgment of the state of knowledge dependence between P_i and P_j.

If Spearman rank correlation coefficients are specified for more than two pairwise state of knowledge dependent data, then the corresponding matrix needs to be positive definite (see Sect. 3.5.2.5). The difficulty of specifying a positive definite matrix of pairwise correlations grows rapidly with the dimension of this matrix. Besides, the subjective probability distributions specified for the uncertain data individually (the so-called marginal distributions) together with the correlation coefficients do not define a unique multivariate distribution. However, a multivariate normal distribution is uniquely defined by the marginal distributions and the correlation matrix.

The copula approach (Sect. 3.5.2.7) defines a multivariate subjective probability density function and thereby yields a positive definite correlation matrix. The expert needs to provide the dependence tree and to fill in the corresponding fields of the dependence array, i.e. he has to provide the values of rank correlation coefficients and of conditional rank correlation coefficients. The elicitation of these values runs as explained above for measures of association. However, the larger the number of uncertain data in the condition (of the conditional correlation coefficients) the more difficult it will be for the expert to arrive at a judgment for the conditional subjective probabilities needed for the Blomquist measure of association and therefore to arrive at a value for the conditional rank correlation coefficient.

The approach described in Sect. 3.5.2.8 gives the expert the opportunity to build the dependence structure using elementary uncertainty contributors that are additive at the multivariate normal level. This way the expert is led to investigate the reasons and nature of the state of knowledge dependence and thereby the interrelationship of state of knowledge dependent uncertain data. All the expert has to provide is the array assigning elementary contributors to the various state of knowledge dependent pairs of uncertain data. The elicitation is only concerned with the rather simple yet intuitive step of consistently assigning elementary uncertainty contributors to the fields in the diagonal and the upper triangle of the assignment array. No correlation coefficients or conditional correlation coefficients need to be specified. The correlation matrix that results from this approach is guaranteed to be positive definite and there is no limitation to its dimension. The following is an example how the analyst would guide the expert through the elicitation for this approach:

Given an arbitrarily large set $\{Q_1, Q_2, \ldots\}$ of state of knowledge independent elementary uncertainty contributors, let the state of knowledge of each be expressed

by a standard normal distribution. Consider now the first uncertain datum P_1 among the set of four state of knowledge dependent uncertain data in the example of Sect. 3.5.2.8 and assign the contributor Q_1 to it, i.e. enter Q_1 in the diagonal field $a_{1,1}$ of the assignment array A. Think of an explanation for this contributor (e.g. the uncertainty about the ambient temperature) and consider now the other uncertain data P_2, P_3, P_4 whether Q_1 could also contribute to their uncertainty. If Q_1 is judged to be also a contributor to the uncertainty of P_3, then enter Q_1 in the field $a_{1,3}$ of the array A. If there is no other uncertain datum that could be seen as a recipient of the contributor Q_1 and if there is more uncertainty of P_1 to be explained, then take Q_2 and enter it in the field $a_{1,1}$. Does Q_2 also contribute to any of the other uncertain data? If so, enter Q_2 in the respective off-diagonal field of row 1 of the array A. Q_3 may be able to account for some of the remaining uncertainty of P_1 but may act on P_4 in the negative direction (i.e. a positive value of Q_3 acting on the value of P_1 in the positive direction would act on the value of P_4 in the negative direction and vice versa); therefore, Q_3 is to be entered into the field $a_{1,1}$ with the plus sign and into the field $a_{1,4}$ with the minus sign. Once all conceivable causes of uncertainty of P_1 have been taken care of by elementary uncertainty contributors (there are 10 contributors in the example of Table 3.8), the elicitation continues with the uncertain datum P_2. All entries in the field $a_{1,2}$ need to be copied into the field $a_{2,2}$. Are there any further causes of the uncertainty of P_2 that need to be taken care off by an elementary uncertainty contributor? If so, enter Q_{11} and possibly others into the field $a_{2,2}$ and see whether any of these are shared with P_3 and/or P_4. Proceed as for P_1, and so forth.

Table 3.9 shows an assignment array with multiples or fractions of uncertainty contributors assigned to uncertain data. This is to account for any knowledge the expert might have of different strengths of the same contributor for different uncertain data (for instance, the contribution of the uncertainty about the ambient temperature to the uncertainty of P_1 may be half or double as strong as to the uncertainty of P_3). The elicitation of these multiples or fractions is self-evident.

After the individual elicitation sessions with each substantive expert have been completed, there still remains the question whether there is state of knowledge dependence among uncertainties with state of knowledge quantifications provided by different experts. To answer this question, all substantive experts are invited to join into a final session where these additional dependences are identified and jointly quantified by the involved experts.

In every instance of state of knowledge dependence quantification, the expert's argumentation is to be documented. He is presented with a graphical representation of the consequences of his quantification for the sample of pairs of values that will be used in the Monte Carlo simulation for uncertainty propagation. This is done with the help of scatterplots. His dependence quantification is finalized and documented for use in the uncertainty analysis only once his agreement has been obtained.

3.7 Survey of Expert Judgment

It is the aim of the uncertainty analysis to compute the combined influence of all possibly important epistemic uncertainties on the result of the computer model application. Thereby a subjective probability distribution is obtained for the model result. It is understood that this distribution expresses the state of knowledge of the team of experts involved in the analysis. Other teams of experts may arrive at a different subjective probability distribution. It may, therefore, be necessary to base the state of knowledge quantifications for data and model uncertainties on the judgment of several experts, depending on the importance of the model result. This chapter, therefore, closes with a discussion of structured approaches to expert judgment elicitation and with recommendations of how to conduct a survey of expert judgment. As was already emphasized at the beginning, "state of knowledge quantification by subjective probability" is the most laborious step of the uncertainty analysis and the quality of the analysis result depends primarily on its outcome.

A properly conducted survey of expert judgment is time consuming and expensive. For this reason it will be an option only in those situations where the results from the application of the computer model are intended to serve as input to sufficiently important decisions. The following description is based on Hora and Iman (1989) and Ortiz et al. (1991).

While expert judgment is part and parcel of most computer models, a structured formal survey of expert judgment is not justified in every instance. Instances that require a formal procedure need to be identified by the project team that develops and/or applies the computer model. They are characterized by:

- Lack of observations.
- Observations that allow for a wide range of interpretations.
- Models that are far from being validated or are severely simplified or may even be called fragmentary.
- A state of knowledge riddled with gaps.
- Modelling attempts that are subject to controversy.
- High complexity.
- Reasoning that is still in the state of work hypotheses.

In these situations (also called "Issues"), use of the judgment of a single expert may be insufficient. The spectrum of possible judgments will most likely lead to significantly different results of the computer model application. Consequently, a quantitative expression of the state of knowledge among experts is required. It is the aim of a survey of expert judgment to obtain this quantitative expression. Elicitation and application of so-called point values without quantification of the state of knowledge is the most common misuse of expert judgment. It gives the impression of precision that is obviously never attainable.

A structured formal survey of expert judgment needs to proceed according to a protocol that is composed of consistently defined, well planned, and carefully documented steps. It must be ensured that the experts do not disassociate themselves

from the quantitative statements that are derived from their judgments or from the use of their judgments in the computer model application and its uncertainty analysis. Quite frequently, expert judgment is elicited in an informal way and/or is implicitly used. The disadvantage of this approach is that the results from the model application and from its uncertainty analysis are not repeatable by interested parties and are therefore not defendable.

3.7.1 The Structured Formal Survey of Expert Judgment

A survey of expert judgment is laborious if executed according to the standard that is required for model results that are intended to support far-reaching decisions. Therefore, it has to follow a procedure that is tightly organized, carefully documented (like a scientific experiment) and conclusively structured.

The essential tasks are shown in Table 3.14 and are described in detail below. Practical examples are documented in Merkhofer (1987), Hora and Iman (1989), Keeney and von Winterfeldt (1991), Ortiz et al. (1991), Hora (1992), Hora and Hora (1992), Otway and von Winterfeldt (1992), Reliability Engineering and System Safety (1997), Goossens and Harper (1998) and Cojazzi et al. (2001). Guidelines and a justification for the use of expert judgment in technical problems are given, for instance, in Keeney and von Winterfeldt (1989), Cooke and Goossens (2000) and Vick (2002).

Task 1: Selection of the Issues
The criteria may be summarized as follows:

- There is disagreement among experts with respect to a quantitative assessment of the issue and possibly also with respect to the importance of the issue for the results of the computer model application.
- Either there are no observations from experiments, tests, operation experience and results from validated computer models or their suitability is subject to controversy.
- Preliminary sensitivity analyses show that the issue may contribute significantly to important results from the computer model application and/or to their uncertainty.

The project team needs to compile all issues that are found to necessitate a survey of expert judgment. The experts, that are to be included in the survey, must be given the opportunity to modify this list if they see the need to do so.

Task 2: Selection of the Experts
The literature (Council on Environmental Quality 1980) provides a rather comprehensive and useful definition of the term “expert” with respect to an issue and a corresponding ranking of expertise into the categories “expert”, “quite familiar”, “familiar”, “casually acquainted” and “unfamiliar”. The description of the first two ranks is quoted in Sect. 3.6. The relevant scientific literature, lists of participants of

conferences, membership lists of scientific societies active in the field, etc. will be valuable sources of information that may be used in an attempt to identify experts for a specific issue. The experts should come from consultancies, research centres, government organizations, universities and other organizations and institutions that are known to be active in the field. The aim is to cover as much as possible of the associated scientific background and of the available expertise. It will be recommendable to employ a more or less formal nomination procedure depending on the degree to which the model results and their uncertainty analysis will have to strive for acceptance among peers. The nomination procedure must be designed in a way that excludes partiality. The selection criteria need to be specific and thoroughly documented. They must include:

- Proof of expertise.
- Reputation in the trade.
- Indication of availability and willingness to participate in the survey.
- Assurance of impartiality, i.e. no personal or economical interests in the model results and especially in the decision they are intended to support.
- Achievement of a balanced representation of the spectrum of existing viewpoints.

Similar criteria must also be satisfied by the selection procedure for the members of the elicitation team. Experts that, for reasons of their professional involvement, cannot satisfy the second last criterion above may present their "insider" information to the expert panel in task 4 below.

Task 3: Grouping of the Experts

Usually, there is an expert panel for each issue. Every expert in the panel deals with the complete issue. Such panels should comprise not less than five and not more than nine experts.

Issues are frequently decomposed and the judgment is elicited at the level of the components of the decomposition. This is, for instance, necessary in the case of issues that extend over several specialized fields. Each expert of the panel contributes his expertise in the field and to the component he is most familiar with. For each component of the decomposition, there are possibly several experts that are providing their judgment. Another approach would be to form several teams with one expert each per component. This approach is, however, only feasible if the number of available experts is sufficiently large for most of the components so that several teams can be put together.

Task 4: Description of the Issue

The project team (represented by the client—see Task 9) has to provide a correct, logically conclusive, comprehensive and unequivocal formulation of the issue that is free of any unmentioned assumptions. This formulation or issue description is to be supplemented by:

- A compilation of the background information including references to the pertinent literature, experimental data, models, computer model results, etc.

– A description of where the issue is going to play a role in the application of the computer model as well as how and where the expert judgment will be made use of (the actual position within the computer model) and also how the issue is to be seen in relation to other issues identified in Task 1.
– A case structure that accounts for the different boundary conditions under which the issue plays a role in the application of the computer model.

When presenting this information to the expert panel in task 6, it will be of advantage to give those experts that could not be included in the panel for various reasons, like possible partiality (see Task 2), an opportunity to also present their information to the panel.

Task 5: Training
The actual elicitation will profit qualitatively from a preparatory, conceptually well-designed training session. As far as this has not been done up to now, this session is to familiarize the experts with:

– The differentiation between aleatoric and epistemic uncertainty.
– The sources of epistemic uncertainty.
– The possible ways of quantifying model uncertainty.
– The difference between "subjective probability" and the classical interpretation of "probability" as the limit of relative frequencies.
– The quantification of the state of knowledge by subjective probability distributions.
– The causes for state of knowledge dependence and the possible ways of its quantification.
– The causes for bias and their effect on expert judgment as well as possible countermeasures.
– The possibilities for and advantages of issue decomposition.

Task 6: Detailed Discussion of the Issue
The project team has to discuss the issue with the experts in sufficient detail in order to:

– Eliminate any latitude of the issue interpretation.
– Achieve broad communality of the background knowledge about the issue and of the appreciation of the issue importance.
– Avoid misunderstandings about the expected format of the answers. Ultimately, subjective probability distributions are to be specified and misunderstandings about what these distributions should quantify are to be avoided.

These discussions should also serve the identification of any necessary extensions of the issue formulation and should help to supplement the issue description accordingly. In addition they will be useful in the identification of further documents that might be required by the experts and in making sure that they are distributed among all experts of the panel in good time.

Task 7: Experts Prepare Their Analyses
The experts prepare their analyses individually in their usual working environment where they have access to all necessary facilities. The exchange of information among experts is not only permitted but is even desirable. The time allotted to this step must be sufficient in order to allow for the issue to be treated in the necessary depth.

In the case of issues that do not allow the preparation of analyses due to the limited state of knowledge, the expert is expected to provide at least a protocol of the scientific or plausibility considerations that led to his judgment. These protocols are then presented and discussed in Task 8.

Task 8: Discussion of the Individual Analyses or Protocols by the Expert Panel
Each expert presents his analyses or considerations that led to his judgment to the panel. In doing so, they use the given case structure (see Task 4) or modifications thereof if found necessary. The presentations do not include the experts' detailed numerical conclusions but only the way and means by which he has arrived at the answers. This includes the information sources (literature references, experiments, model calculations, etc.) they are based upon. The presentations also include any issue decompositions the expert might have used. The decomposition could be a computer model. The purpose of these discussions is to show how the expert arrived at his answer and not the actual judgment itself.

Task 9: Elicitation of the Expert Judgment
The expert judgment is elicited in individual sessions. The only participants in these sessions are the expert, the analyst and the client. The client is a member of the project team, the project being the intended development and/or application of the computer model to derive the input needed for decision-making. He knows how the expert judgment is used in the computer model, and he is familiar with the corresponding boundary conditions and how the model result is used in the decision-making process. Furthermore, the client documents the scientific considerations, issue decompositions and arguments of the experts together with their conclusions and he checks for consistency.

Again, it is recommended to conduct these sessions in the working environment of the expert. Standardized formulations of the questions ensure that the same wording is used with all experts. Consequently, it is of advantage if the elicitation team (client and analyst) is the same for all experts. The analyst is familiar with the concept of subjective probability, with probability calculus and with the requirements of the elicitation of probability statements. He is expected to support the experts in their state of knowledge quantifications by subjective probability distributions, to assist them in avoiding bias, to check the consistency of their probability statements and to document them suitably for later use in the propagation of the state of knowledge through the computer model application.

Task 10: Recombination
The expert may have decomposed the issue in order to quantify his state of knowledge at the level of the components of the decomposition. The time has now

come to recombine these quantifications to the issue assessment, i.e. his judgment for the issue. If the decomposition was in the form of a computer model the recombination is straightforward. In all other situations, the recombination will most likely need to be performed by the elicitation team possibly supported by suitable software. The result of the recombination will then be presented to the expert for his final approval. He needs to judge whether it is an adequate expression of his state of knowledge. The reason for any discrepancies must be found and eliminated in cooperation with the expert. A review of the issue decomposition and of the expert's state of knowledge quantifications will enable the elicitation team together with the expert to discover misunderstandings and any documentation or processing errors or any flaws in the expert's reasoning. Any changes to the expert's answers that might have been carried out upon the expert's request, after he was presented with the result of the recombination, need to be backed by sufficient reasoning. Reasoning and the performed changes need to be documented.

Task 11: Aggregation
The assessments of the individual experts or expert teams must now be aggregated to the issue assessment. This is achieved by mixing the corresponding subjective probability distributions provided by the experts (or expert teams). Usually, the individual distributions receive equal mixing weights $1/K$ with K being the number of experts or teams providing distributions. At any value of the quantity in question, the value of the mixture distribution is the arithmetic mean of the cumulative probabilities read from the individual distributions at this value. In Council on Environmental Quality (1980), unequal weights, derived from self and colleague ratings of the experts, were used. The sum of the weights has to add up to 1.

Within the framework of uncertainty analysis by Monte Carlo simulation, N replications of the computer model evaluation (for short " N model runs") are performed, each using a new set of values for the epistemic uncertainties chosen at random according to their state of knowledge quantifications. The K subjective probability distributions provided by the experts (or expert teams) may be seen as those obtained using alternative "model formulations". It seems, therefore, preferable not to use the mixture distributions mentioned above for sampling but to assign an index value to each of the K experts or expert teams and for the n-th model run, $n = 1, \ldots, N$, to first sample from the set of K index values using the mixing weights as subjective probabilities. Then a value for the uncertain datum in question is sampled according to the subjective probability distribution provided by the expert (or expert team) that corresponds to the sampled index value. This value of the uncertain datum is then used in the n^{th} replication of the computer model evaluation.

Task 12: Documentation
The documentation of the issue assessment ideally contains:

- The reasons why the issue was selected for a survey of expert judgment.
- The procedure used to select the experts.
- The issue description given to the experts.
- The contents of the training session.

Table 3.14 Structured approach to the survey of expert judgment

No.	Task
1	Selection of the issue
2	Selection of the experts
3	Grouping of the experts
4	Description of the issue
5	Training
	Tasks 5 and 6 are the subject of the first meeting of all panel experts
6	Discussion of the issue
7	Experts prepare their analyses
8	Discussion of the analyses
	Second meeting of all experts
9	Elicitation
10	Recombination
11	Aggregation
12	Documentation

- The session minutes to Task 6.
- The analyses (or protocols) produced by the experts in Task 7 together with a compilation of the information sources used.
- The session minutes to Task 8.
- The session minutes to Task 9, in particular the formulation of the questions presented to the expert, the expert's answers, the expert's argumentation to support his answers and the way and result of the processing of his answers by the elicitation team.
- The recombination procedure, the results of the recombination and the corresponding documentation as well as the documentation of any corrections of the recombined results that might have been initiated by the experts.
- The individual issue assessments and the results of their aggregation over the experts, including the procedure used, the choice of the aggregation weights, reasons for the use of unequal weights (if any) and the information source used for their specification.

3.7.2 The Structured Formal Survey of Expert Judgment by Questionnaire

The approach described in the previous section involves two meetings of the experts. Apart from the time and financial resources needed for travel, these meetings require the availability of all experts over the same period of time and on two occasions. This requirement can often not be met. A structured formal survey of expert judgment by questionnaire would be a way out of this dilemma. In this approach, Tasks 1–3 of Table 3.14 would be performed as in Sect. 3.7.1 but Tasks 4, 5 and 9 would need to

be covered by the questionnaire. The client, in cooperation with the analyst, will design the questionnaire. The client then travels to each of the experts and presents the questionnaire. He performs Task 6 together with each expert individually. Task 7 is performed as in Sect. 3.7.1 and Task 8 requires some interaction between the experts via the usual communication channels. This interaction may need to be encouraged by the client. There will be ample time for communication between the experts, as the calendar time allotted to the completion of the questionnaire needs to be chosen such that all experts can fit the involved workload into their individual schedules. The experts send their completed questionnaires to the elicitation team, and Tasks 10–12 are then performed the same way as in the previous section.

References

Armstrong, B. G. (1998). Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occupational and Environmental Medicine, 55*, 651–656.

Bayes, T. (1958, December). Thomas Bayes' Essay towards solving a problem in the doctrine of chances. In *Studies in the history of probability and statistics* (Vol. 45). *Biometrika*, Parts 3 and 4.

Berkson, J. (1950). Are there two regressions? *Journal of the American Statistical Association, 45*, 164–180.

Box, G. E. P., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Reading, MA: Addison-Wesley.

Buckley, J. J. (1985). Entropy principles in decision making under risk. *Risk Analysis, 5*(4), 303–313.

Cojazzi, G., et al. (2001). Benchmark exercise on expert judgement techniques in PSA Level 2. *Nuclear Engineering and Design, 209*, 211–221.

Conover, W. J. (1980). *Practical nonparametric statistics* (2nd ed.). New York: Wiley.

Cooke, R. M., & Goossens, L. J. H. (2000). *Procedures guide for structured expert judgment*. European Commission, Nuclear Science and Engineering, EUR 18820EN.

Council on Environmental Quality. (1980). *The Global 2000 Report to the President*. Washington, DC: U.S. Government Printing Office.

Fishman, G. S. (2000). *Monte Carlo – Concepts, algorithms and applications*, *Springer series in operations research*. New York: Springer.

Freund, R. J., & Minton, P. D. (1979). *Regression methods – A tool for data analysis*. New York: Marcel Dekker.

Goossens, L. H. J., & Harper, F. T. (1998). Joint EC/USNRC expert judgment driven radiological protection uncertainty analysis. *Journal of Radiological Protection, 18*(4), 249–264.

Heinhold, J., & Gaede, K.-W. (1968). *Ingenieur-Statistik*. München: R. Oldenbourg.

Hofer, E. (2008). How to account for uncertainty due to measurement errors in an uncertainty analysis using Monte Carlo simulation. *Health Physics, 95*(3), 277–290.

Hora, S. C. (1992). Acquisition of expert judgment: Examples from risk assessment. *Journal of Energy Engineering, 118*(2), 136–148.

Hora, S. C., & Hora, J. A. (1992). Assessment of probability distributions for continuous random variables: A comparison of the bisection and fixed value methods. *Organizational Behavior and Human Decision Processes, 51*, 133–155.

Hora, S. C., & Iman, R. L. (1989). Expert opinion in risk analysis: The NUREG-1150 methodology. *Nuclear Science and Engineering, 102*, 323–331.

Hora, S. C., et al. (1993). The use of decomposition in probability assessments of continuous variables. *Journal of Behavioral Decision Making, 6*, 133–147.

Iman, R. L., & Conover, W. J. (1982). A distribution-free approach to inducing rank correlations among input variables. *Communications in Statistics: Simulation and Computation, 11*(3), 311–334.

Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (1st ed.). JCGM 100 (GUM 1995 with minor corrections).

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Keeney, R. L., & von Winterfeldt, D. (1989). On the uses of expert judgment on complex technical problems. *IEEE Transactions on Engineering Management, 36*(2), 83–86.

Keeney, R. L., & von Winterfeldt, D. (1991). Eliciting probabilities from experts in complex technical problems. *IEEE Transactions on Engineering Management, 38*(3), 191–201.

Kloos, M. (2015). *Main features of the tool SUSA 4.0 for uncertainty and sensitivity Analyses*. ESREL 2015, European Safety and Reliability Conference, Zuerich, Switzerland.

Kruskal, W. H. (1958). Ordinal measures of association. *Journal of the American Statistical Association, 53*, 814–861.

Krzykacz, B. (1993). *Guidelines for representation and probabilistic quantification of parameter uncertainty and dependence*. GRS-A-2124, Gesellschaft fuer Anlagen- und Reaktorsicherheit, Garching.

Kurowicka, D., & Cooke, R. M. (2006). *Uncertainty analysis with high dimensional dependence modelling*. Chichester: Wiley.

Lipschutz, S. (1964). *Set theory and related topics*. New York: McGraw-Hill.

Merkhofer, M. W. (1987). Quantifying judgmental uncertainty: Methodology, experiences and insights. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-17*(5), 741–752.

Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed.). New York: Springer.

Ortiz, N. R., et al. (1991). Use of expert judgment in NUREG-1150. *Nuclear Engineering and Design, 126*, 313–331.

Otway, H., & von Winterfeldt, D. (1992). Expert judgment in risk analysis and management: Process, context, and pitfalls. *Risk Analysis, 12*(1), 83–93.

Park, I., et al. (2010). A Bayesian approach for quantification of model uncertainty. *Reliability Engineering & System Safety, 95*, 777–785.

Reliability Engineering and System Safety. (1997). Special issue on Expert Judgement.

Rosner, B. (1995). *Fundamentals of biostatistics* (4th ed.). New York: Duxbury Press.

Ross, T. J. (1995). *Fuzzy logic with engineering applications*. New York: McGraw-Hill.

Schirmacher, D., & Schirmacher, E. (2008). *Multivariate dependence modeling using pair-copulas*. Society of Actuaries.

Siu, N. O., et al. (1992). Bayesian assessment of modeling uncertainties: Application to fire risk assessment. In B. M. Ayyub, M. M. Gupta, & L. N. Kanal (Eds.), *Analysis and management of uncertainty: Theory and applications*. New York: Elsevier Science Publishers.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institute de Statistique de l'Université de Paris, 8*, 229–231.

Smith, A. E., et al. (1992). The effect of neglecting correlations when propagating uncertainty and estimating the population distribution of risk. *Risk Analysis, 12*(4), 467–474.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15*, 72–101.

Stephens, M. E., et al. (1993). Deriving parameter probability density functions. *Reliability Engineering and System Safety, 42*, 271–291.

Vasquez, V. R., & Whiting, W. B. (2006). Accounting for both random errors and systematic errors in uncertainty propagation analysis of computer models involving experimental measurements with Monte Carlo methods. *Risk Analysis, 25*(6), 1669–1681.

Vick, S. G. (2002). *Degrees of belief, subjective probability and engineering judgment*. Reston, VA: ASCE Press.

Chapter 4
STEP 3: Propagate

4.1 Introduction

The computer model maps the possibly true values of the uncertain data[1] into values of the model results. This is illustrated in Fig. 4.1 for the case of a model h(P_1, P_2) that maps each pair of possibly true values of (P_1, P_2) into one model result value.

For M uncertain data and J model results, the mapping is given as

$$\mathrm{h}: \boldsymbol{P}^{\mathrm{M}} \rightarrow \boldsymbol{Y}^{\mathrm{J}}. \tag{4.1}$$

$\boldsymbol{P}^{\mathrm{M}}$ is the space of possibly true M-tuples of the uncertain data, and $\boldsymbol{Y}^{\mathrm{J}}$ is the space of the corresponding J-tuples of model result values. According to Eq. (4.1), there is one and only one J-tuple for each M-tuple while in the case of

$$\mathrm{h}: \boldsymbol{P}^{\mathrm{M}} \leftrightarrow \boldsymbol{Y}^{\mathrm{J}} \tag{4.2}$$

there is also one and only one M-tuple for each J-tuple, i.e. the mapping performed by h is reversible

$$\mathrm{h}^{-1}: \boldsymbol{Y}^{\mathrm{J}} \rightarrow \boldsymbol{P}^{\mathrm{M}}. \tag{4.3}$$

Many computer models do not represent a reversible mapping, i.e. there is more than one M-tuple for some or all J-tuples.

In Fig. 4.2, the states of knowledge of the two uncertain data of Fig. 4.1 are quantified by the joint subjective probability density function $\mathrm{f}_{1,2}$ of P_1 and P_2. This state of knowledge expression at the data level is now included in the mapping of Fig. 4.1.

[1]To simplify matters, this and the following chapters only use the term "uncertain data" since model uncertainties are represented by uncertain parameters and the latter are, together with uncertain input data, categorized as uncertain data.

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_4

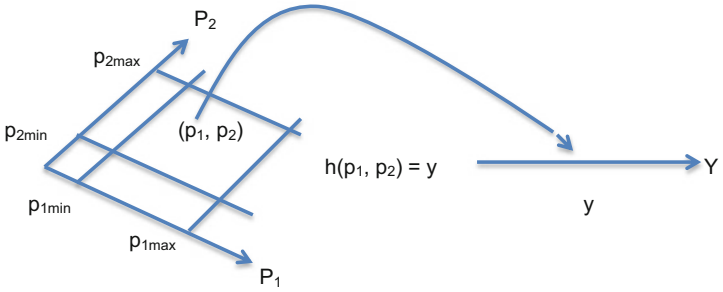
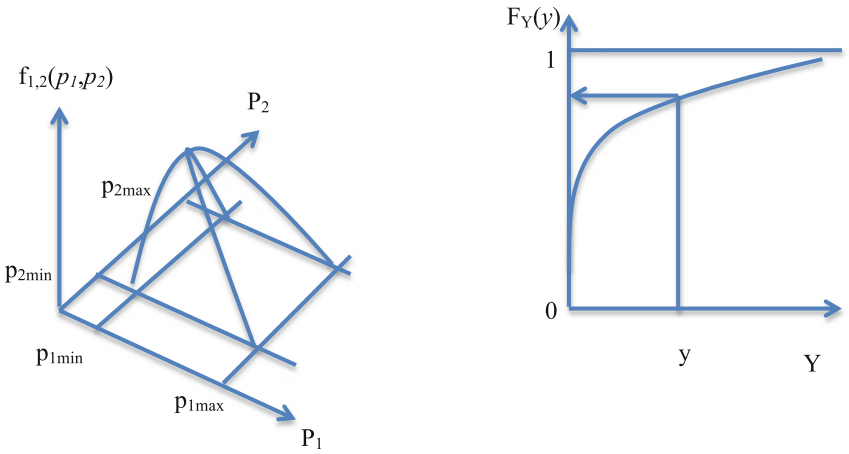


Fig. 4.1 Mapping of pairs of parameter values into one model result value



$$\int\int \mathbf{I}_{(h(p_1,p_2)\leqq y)}\, f_{1,2}(p_1,p_2)\mathrm{d}p_1\mathrm{d}p_2 = F_Y(y)$$

Fig. 4.2 The degrees of belief of all possibly true pairs of parameter values that are mapped, using the model h, into a value smaller or equal to y are accumulated and presented as the cumulative subjective probability $F_Y(y)$. The cumulative distribution function F_Y quantifies the state of knowledge that follows for the model result

Figure 4.2 shows how the computer model translates the joint state of knowledge expression for the uncertain data into the state of knowledge expressed by the cumulative subjective probability distribution function F_Y for the model result. The use of capital letters for data and model results indicates that their values are associated with state of knowledge expressed by subjective probability.

An analytical approach to propagation is usually out of the question due to the complexity of most computer models of practical relevance. Consequently, the subjective probability distributions that follow for the model results remain unknown. A statistical approach needs to be taken in this situation. It generates a random sample of size N that is drawn according to these unknown distributions. This is achieved by drawing a random sample of size N according to the subjective probability distributions and state of knowledge dependence expressions specified

for the M uncertain data and by evaluating the model for each of the N sets of M sample values. The corresponding N sets of J model result values are a random sample drawn according to their unknown joint subjective probability distribution. This sample can then be processed using standard statistical techniques to quantify the uncertainty of the model results in Step 4 and to rank the uncertain data with respect to their contribution to the uncertainty of each of the model results in Step 5.

4.2 Random Sampling

Random sampling makes use of the inverse transform method. For an uncertain datum P, it starts out with drawing a value at random according to the uniform distribution over the unit interval. With the help of the inverse of the subjective probability distribution function specified for P, this value is then transformed into a value sampled according to the specified distribution. If a value for the uncertain datum P is to be drawn according to the cumulative subjective probability distribution function F_P specified as state of knowledge expression for P, then

$$F_P(p) = \mathrm{sw}(P \leq p) = u$$
and the inverse ${F_P}^{-1}$ of the distribution function provides ${F_P}^{-1}(u) = p.$

"sw" stands for "subjective probability", and u is drawn at random according to the uniform distribution over the interval (0, 1). From the relationship ${F_P}^{-1}(u) = p$, the sample value p of P is found. Random sampling according to truncated distributions as well as discrete distributions can also be achieved by this method. Section 3.6.1.3 explains how to derive p from u for each of the subjective probability distributions commonly used in uncertainty analysis.

The statistical approach to state of knowledge propagation involves taking a sample of N M-tuples and mapping each M-tuple into the corresponding J-tuple by evaluating the model for the M-tuple.

For example, if the uncertain data P_1 and P_2 are not state of knowledge dependent, a pair of values (u_1, u_2) is chosen independently at random according to the uniform distribution over the $\boldsymbol{U}^2$ space (i.e. the unit square). The corresponding pair of values (p_1, p_2) for the uncertain data is obtained via the inverses of the distribution functions specified for P_1 and P_2, namely $p_1 = {F_1}^{-1}(u_1)$ and $p_2 = {F_2}^{-1}(u_2)$. Evaluation of the model h for (p_1, p_2) provides the corresponding model result value y'. The chance of drawing a pair of values (u_1, u_2) such that $\mathrm{h}({F_1}^{-1}(u_1), {F_2}^{-1}(u_2)) \leq y$ is equal to the chance of drawing a value u according to the uniform distribution over the unit interval such that ${F_Y}^{-1}(u) \leq y$. Due to this equivalence, one can obtain sample values of Y according to F_Y although F_Y is unknown. This opens the door to the estimation of distribution characteristics of F_Y, like mean value, variance, the value y to any given value $u = F_Y(y)$, the subjective probability u for Y to exceed any given value y (i.e. $u = 1 - F_Y(y)$) and so forth, from a sufficiently large sample of values drawn at random according to the joint subjective probability distribution specified for the uncertain data.

The inverse distribution function always exists, but it is not always readily available in analytical form. Fishman (2000) presents a table containing inverse functions for some commonly used distributions. Approximations are given for those not in the table. The reference also describes other approaches, in addition to the inverse transform method, particularly with a view to efficiency and accuracy. Algorithms are provided for a number of distribution functions that can be turned into computer programs for sampling. Commercially available subroutine libraries provide ready-to-use subroutines for a number of commonly used distribution functions. They generally start off with drawing a value at random according to the uniform distribution over the unit interval. "At random" means that each number on the unit interval has the same chance of being drawn. How can one draw numbers from the unit interval truly at random? Radioactive decay is, among other natural phenomena, considered to be a source of true randomness. Numbers from the unit interval, generated this way and stored on the computer, could be used as arguments of the inverse distribution functions. For most applications of random numbers, like in the propagation of states of knowledge in uncertainty analysis, the use of computerized pseudorandom number generators is preferred for obvious reasons. The properties of such generators play a crucial role. Fishman (2000) discusses and compares several pseudo-random number generators. The sequences of sample values generated are called "pseudo-random" as there is no real randomness involved in their generation. A sequence of N pseudo-random numbers has to satisfy certain criteria in order to qualify as statistically random, i.e. as a satisfactory realization of a sequence of N numbers drawn at random according to the uniform distribution over the unit interval. For instance, any sequence of N digits from the string of digits of π is a statistically random sample of numbers drawn according to the discrete uniform distribution of the digits 0, 1, …, 9. To have a (statistically) random sample of size N from the unit interval, one might take a sequence of kN digits with $k = 4$ (or any other suitable integer), separate them into N sets of k digits each and divide each by 10^k.

The suitability of the pseudo-random number generator for the intended use is determined by the degree to which the criteria for statistical randomness are met. Probability theory defines the characteristics of real randomness. Based on this definition, it is possible to formulate criteria (Fishman 2000) that have to be met by pseudorandom number generators in order to produce sufficiently long sequences of numbers that are statistically random.

The suitability of the pseudorandom number generator for uncertainty analysis depends on how the length of statistically random sequences compares to the number M of uncertain data and to the number N of computer model evaluations to be performed in the course of the Monte Carlo simulation. A sequence of at least MN pseudorandom numbers is to be generated. Chapter 9 introduces the nested Monte Carlo simulation concept for the separation of M epistemic and K aleatoric uncertainties. This concept may require N^2 evaluations of the model and therefore the generation of sequences of at least $(M + KN)N$ pseudorandom numbers. The suitability of the pseudo-random number generator for uncertainty analyses with large values of M, N and K should therefore be assessed. Fishman (2000) provides some assistance in this effort.

4.3 Monte Carlo Simulation

The cumulative subjective probability distribution function F_Y summarizes the population of values of the model result Y that follows from the propagation of the state of knowledge for the uncertain data through the model. The statistical approach does not sample directly model result values y according to F_Y since this distribution function is usually unknown. Rather, it simulates possibly true values of Y via the mapping of sets of M values chosen at random from the $\boldsymbol{P}^M$ space of the uncertain data into the $\boldsymbol{Y}^J$ space of the model results through evaluation of the model for each set. This process is known as Monte Carlo simulation as it combines the chance element of random sampling with simulation according to the instructions of the encoded computer model.

Monte Carlo simulation generates a sample of size N (the input sample of the analysis), drawn according to a joint subjective probability distribution for the M uncertain data. This joint distribution complies with the individually specified marginal subjective probability distributions and with the state of knowledge dependence quantifications. Correlation coefficients specified in Step 2 as quantitative expressions of state of knowledge dependence will not generally determine a unique joint subjective probability distribution. There are usually other joint subjective probability distributions that comply with the same marginal distributions and the same correlation coefficient. This ambiguity is, however, not considered to be a handicap for uncertainty analysis as long as the generated joint distribution has (nearly) minimal information contents in addition to the information provided by the state of knowledge quantifications (see also Sect. 3.5.2.7).

Random (input) samples of small-to-moderate size N for M uncertain data, with M close to or even larger than N, will exhibit non-negligible spurious correlations. These correlations are a product of the chance element of random sampling. They fake state of knowledge dependence among the uncertain data thereby 'contaminating' the sample. Spurious correlations contribute to the relatively large variability of estimates of mean value, variance and percentiles of the model results derived from small (output) sample sizes N. Section 4.4.1.14 presents a procedure (Iman and Conover 1982) to induce specified sample rank correlations. It may also be used (in an iterative fashion) to eliminate (or reduce) spurious correlations if $M < N$. In the case of $M > N$, it may still be applied to subsets of $M' < N$ uncertain data. This matrix procedure does not change the actual (input) sample values of individual data but their combination into N sets of M values each.

An element of the input sample is a data set that carries one value for each of the M uncertain data at the prescribed position. The Monte Carlo simulation generates an input sample of size N consisting of N such data sets and evaluates the computer model for each. To this end, it will be necessary to supplement the input data file of the computer model by those data that are not yet represented in the file. This applies particularly to model uncertainties expressed by a set of alternative model formulations or by uncertain correction factors.

For each of the N model evaluations (called "runs" for short), the values of the M uncertain data in the input file need to be changed into those of the sample element that is to be used by the run. The changes are, for reliability reasons, best done by a computer program that generates the N needed copies of the input data file, i.e. it adjusts each copy according to the set of values in the corresponding sample element. Data that have been augmented to the file will require changes to the computer model so that their values can be read and used as intended.

At the end of the Monte Carlo simulation, there are N sets of values. Each set contains one value each for the J results $Y_1, \ldots, Y_J$. These N sets of J values each are a random sample of size N (the output sample of the analysis) drawn according to the unknown joint subjective probability distribution that logically results for $Y_1, \ldots, Y_J$ from the state of knowledge quantifications at the level of the uncertain data and from the sequence of logic and arithmetic instructions of the encoded model. Generally, the state of knowledge for any pair of model results is dependent since they most likely share many of the uncertain data. Consequently, the output sample is rich in information on functional relationships among the model results. The N sets are ready for use in the subsequent analysis steps to derive measures of uncertainty and uncertainty importance for any of the model results Y_j. Suitable uncertainty measures are either a value range of Y_j with a specified subjective probability content or the subjective probability content of a specified value range of Y_j.

The value ranges with specified subjective probability content are either defined by two limiting values or by only one value, the other end of the range being the minimum or the maximum value in the population of values that follows for the model result from the state of knowledge propagation. The one-sided limits are quantile values to the specified subjective probability (upper limit) or to its complementary value (lower limit). A frequently specified subjective probability value is 0.95. The limiting values are estimated from a random sample of model result values. Since sample sizes are usually small due to processor time and cost considerations the estimates need to be supplemented by confidence intervals to specified confidence levels. A frequently specified level is 95%.

Well-established methods from statistics may be applied, using the sample of model results in combination with the corresponding sample values of the uncertain data, in order to obtain the uncertainty measures and to rank the data with respect to their uncertainty importance for each of the model results. This does not require a sample size N in the order of several thousands or even millions, as is often claimed. A moderate sample size in the order of N = 100 (i.e. 100 runs of the computer model) or several hundreds is frequently sufficient to obtain the necessary uncertainty measures and uncertainty importance measures for any of the model results. For instance, the minimum sample size N needed for the estimation of two limiting values such that the subjective probability content is at least 0.95 at a confidence level of at least 95% (a so-called (95%, 95%) statistical tolerance interval) is N = 93, independent of the number M of uncertain data involved in the analysis (see Chap. 5). Uncertainty statements like these require the sample to be a simple random sample, i.e. to be purely (pseudo) random as described above. For these relatively small sample sizes, it will be affordable to evaluate the full computer model

including all of its feeder models. The use of a simplified model or of a so-called meta-model (Storlie et al. 2009) for the purpose of the uncertainty analysis, as is sometimes suggested, is not recommendable and is unnecessary. Better 100 runs of the full model than a million runs of a so-called meta-model. The influence of such simplifications is itself uncertain and generally difficult to judge especially if a set of model results of extremely small subjective probability is of particular interest. Sampling methods that address ranges of model result values of small to very small subjective probability are discussed in Sects. 4.4.3 and 4.4.4.

4.4 Sampling Techniques

The sampling techniques most frequently used in uncertainty analysis are simple (sometimes also called "pure" or somewhat derogatory "crude") random sampling (SRS) and Latin Hypercube sampling (LHS). SRS is discussed in Sect. 4.4.1, and the discussion of LHS follows in Sect. 4.4.2. The state of knowledge for some uncertain data may depend on the values obtained for interim model results. The discussion of the sampling techniques will have to pay specific attention to how these situations are handled and particularly how the techniques fare in the case of the state of knowledge dependence models presented in Sect. 3.5.2.

The subjective probability content of specified critical value ranges is usually small. For moderately small probability content (10^{-1} to 10^{-3}), the sampling techniques of Sects. 4.4.1 and 4.4.2 will be affordable. For very small probability content, the necessary SRS or LHS sizes will most likely be unaffordable although upper confidence limits are always available from an affordable SRS but may often be judged as too coarse. Therefore, a number of techniques have been developed to arrive at an estimate for very small (10^{-4} and lower) subjective probability content of specified value ranges of the model result. Among these techniques are importance sampling (discussed in Sect. 4.4.3) and subset sampling (discussed in Sect. 4.4.4).

4.4.1 Simple Random Sampling (SRS)

The adjective "simple" stands for "not tampered with". The pseudorandom numbers are generated without constraints imposed and are used as they come (no reordering or unequal weighting, for instance). Constraints would be, among others, instructions like "Only one number per interval of probability content $1/N$" or "preference for specific subspaces of the value space of the uncertain datum" or "measures taken to achieve even coverage of the value space". None of these interferences that introduce a deterministic element into the sample are made. This is one reason why Monte Carlo simulation using SRS is sometimes in a somewhat derogatory manner called "crude Monte Carlo". The other reason is that no attempts are being

made to reduce the computational simulation effort by replacing the original model by a simplification (meta-model) thereof or by restricting the model evaluation to sequences of sub-models that are thought to be responsible for the model result values from the range of particular interest. All of these steps are thought to enhance efficiency but simultaneously introduce additional uncertainty that is often not quantifiable in a satisfactory manner.

SRS is in essence the sampling procedure described in Sects. 4.2 and 4.3. The following subsections are an extension of the description with a view on some desirable features of a sampling technique and on the case of state of knowledge dependence.

4.4.1.1 Statistics for the Estimation of Distribution Characteristics

If the estimates of uncertainty measures are derived from an SRS, confidence statements (limits and intervals with specified confidence level percentage) for the true values (see Chap. 5) can be obtained. This feature cannot be appreciated enough. It is due to the fact that the model result values obtained with simple random sampling are independently sampled according to their subjective probability distribution. Since computer models of practical relevance are usually processor time intensive, the number of affordable model runs is only moderate. This makes confidence statements a necessary supplement for the estimates of uncertainty measures.

4.4.1.2 Easy Extension of the Sample Size

An uncertainty analysis using an SRS of sample size N may suggest that tighter confidence limits would be desirable for the estimates derived for some of the uncertainty measures. A sample size $N' > N$ should therefore have been used. This does not require another Monte Carlo simulation of sample size N' since the existing sample can simply be extended by an additional SRS of size $N' - N$, and the estimates and confidence statements can be derived from the output sample that is extended by the results of the additional $N' - N$ model runs.

4.4.1.3 State of Knowledge Dependence on Interim Results

Frequently, all N sets of M values (the input sample) can be obtained before the N runs of the model are started. There are, however, model applications that require the sampling of some of the M values to be done dynamically, i.e. when needed. This might, by the way, not be the case in each of the N model runs. This 'just on time' sampling may be necessitated by information that is required for the state of knowledge quantification but becomes available only in the course of the model run. It is simply the only economical way since sampling in advance would mean to

keep megabytes of sampled values ready for a model run in order to cover every eventuality that might occur during the run, while at the end of it only a small fraction of this data deluge turns out to have been actually made use of. The following situations are considered:

– The state of knowledge depends on the value of an interim result

If the state of knowledge of the datum P_j depends on the value of the model interim result Q then the N sets of M values would initially contain sample values $u_{j,n}$, drawn according to the uniform distribution over the unit interval, in place of the actual values $p_{j,n}$, $n = 1, \ldots, N$. Once the value q_n of the interim result has been determined, the subjective probability distribution $\mathrm{F_{jlq}}(p_j)$ is either read from a table giving $\mathrm{F_{jlq}}$ for discrete values of Q or obtained from a function relating the parameter values of $\mathrm{F_{jlq}}$ to the value of Q (the table or the function would need to be part of the state of knowledge specifications provided as input to the uncertainty analysis). The sample value $p_{j,n}$ is found from the inverse relationship $p_{j,n} = \mathrm{F_{jlq}}^{-1}(u_{j,n})$. The value $p_{j,n}$ replaces $u_{j,n}$ in the n-th set of M sample values.

In some computer models, the solution path, followed during the n-th model run, may branch off before the interim result Q has been determined, i.e. q_n is not required and therefore the value $u_{j,n}$ is not replaced by a value $p_{j,n}$. In this case, $u_{j,n}$ needs to be replaced by a default value that is recognized as such by Step 5 (determines uncertainty importance measures). P_j could, for instance, be the correction factor to an uncertain model formulation. It would be needed only if the respective model formulation is applied during the model run. Q could be an interim result that is input to the model formulation such that the state of knowledge of P_j depends on its value.

– The state of knowledge depends on the value of an interim result that is function of an independent model variable like time or space.

With time t being the independent variable, the N sets of M values would initially contain a sample value $u_{j,n}$, drawn according to the uniform distribution over the unit interval, in place of the actual datum value $p_{j,n}(t)$, $n = 1, \ldots, N$ for all times t. Once the value $\mathrm{q_n}(t)$ of the interim result has been determined, the value $\mathrm{F_{jlq}}(p_j(t))$ of the conditional subjective probability distribution is either read from a table giving $\mathrm{F_{jlq}}$ for discrete values of Q or obtained from a function relating the parameters of $\mathrm{F_{jlq}}$ to the value of Q (the table or the function would need to be part of the state of knowledge specifications provided as input to the uncertainty analysis). The sample value $p_{j,n}(t)$ is found from the inverse relationship $p_{j,n}(t) = \mathrm{F_{jlq}}^{-1}(u_{j,n})$. Here, the datum value $p_{j,n}(t)$ is used in the model evaluation only while $u_{j,n}$ is retained in the n-th set of sample values for subsequent analysis steps. Obviously, this is a case of complete state of knowledge dependence between the uncertain data $P_j(t)$ at all times t. If this assumption of complete dependence is deemed unjustified, then the uncertain datum P_j would have to be treated as a model uncertainty. The correct sequence of values over time would be the relationship to be modeled. It is not always possible to sample a set of N alternative sequences beforehand. Rather, each of the N sequences is created during the respective model run. Not everyone of the N model runs may need a sequence $p_{j,n}(t)$ as some of the model runs may follow a

solution path that does not require the values of $P_j(t)$ for any of the times t used in the sequence. If this is the case for the n-th model run, a default value, recognizable as such by subsequent analysis steps, will need to replace $u_{j,n}$ in the n-th set of sample values.

How does SRS fare in situations where uncertain data are state of knowledge dependent on the values of interim model results? Since these results are only available in the course of the model run, sampling of these data values has to be done during the run, i.e. whenever the interim result is actually computed. This may be more or less frequently during a run and not necessarily in each of the N model runs. There is no problem with this requirement since the sample elements of an SRS do not need to be fully set up before the start of the N model runs.

4.4.1.4 State of Knowledge Dependence on Previously Sampled Values

The computer model may make use of a time (or space) advancement procedure that uses equal or unequal step lengths, the sequence of which becomes known only during the model run. The corresponding advancement of the values of some uncertain data may need to be determined. The state of knowledge of the value of their advancement (i.e. their value for the new step) depends on the value(s) they had in the previous step(s) and on the length of the new step. Clearly, this is a case of the 'just on time' sampling situation mentioned above. It does not cause any problems for the simple random sampling method.

How does SRS fare in the cases of state of knowledge dependence modeling discussed in Sect. 3.5.2?

4.4.1.5 Conditional Distributions

If the state of knowledge dependence between uncertain data P_i and P_j was expressed by conditional distributions, the value $p_{i,n}$ is sampled for the so-called free uncertain datum P_i according to its marginal distribution. For P_j, a value $u_{j,n}$ is sampled at random according to the uniform distribution over (0, 1). The value $p_{j,n}$ is obtained from the inverse of the conditional distribution for P_j that was specified given $P_i = p_{i,n}$ (the parameters of the conditional distribution may be a function of the value $p_{i,n}$) or for the interval of P_i that contains $p_{i,n}$ (see Sect. 3.5.2.1).

4.4.1.6 Constraints

How to account for state of knowledge dependence that is due to constraints on the range of possibly true values for the uncertain datum P_j, given a possibly true value for P_i? In practice, there is a choice between two procedures:

a) Sampling from the marginal distributions of P_i and P_j is done independently but any pair of values ($p_{i,n}$, $p_{j,n}$) that does not satisfy the inequality is discarded. The sampling proceeds as long as necessary to have the desired sample size. This can, in some instances, take a long time depending on how stringent the constraint is.
b) The relationships presented in Sect. 3.5.2.2 are used. This is, in essence, an approach using conditional distributions for P_j given the value for P_i. It is the equivalent of the procedure under a) but without the sampling of pairs that do not satisfy the inequality.

For the constraint $p_j \leq \mathrm{g_j}(p_i)$, it proceeds as follows:

1) Sample $p_{i,n}$ according to $\mathrm{F_i}$;
2) Sample $p_{j,n}$ according to $\mathrm{f_{j|i}}(p_j|p_i) = \mathrm{f_j}(p_j) \,/\, A_{j|i}$ as follows:
2a) Sample $u_{j,n}$ according to the uniform distribution over (0, 1)
2b) Compute $A_{j|\mathrm{i}} = \mathrm{F_j}(\mathrm{g_j}(p_{i,n}))$ and set

$$u_{j,n} = \int_{p_{jmin}}^{p_{j,n}} \left[\mathrm{f_j}\left(p'_j\right)/A_{j|i}\right] \mathrm{d}p'_j = \left(\frac{1}{A_{j|i}}\right) \mathrm{F_j}\left(p_{j,n}\right) \tag{4.4}$$

$$\text{and } p_{j,n} = \mathrm{F_j}^{-1}\left(u_{j,n} A_{j|i}\right) = \mathrm{F_j}^{-1}\left(\mathrm{F_j}\left(\mathrm{g_j}\left(p_{i,n}\right)\right) u_{j,n}\right).$$

For the constraint $p_j > \mathrm{g_j}(p_i)$, it proceeds as follows:

1) Sample $p_{i,n}$ according to $\mathrm{F_i}$;
2) Sample $p_{j,n} - \mathrm{g_j}(p_{i,n})$ according to $\mathrm{f_{j|i}}(p_j|p_i) = \mathrm{f_j}(p_j) \,/\, A_{j|i}$ as follows:
2a) Sample $u_{j,n}$ according to the uniform distribution over (0, 1)
2b) Compute $A_{j|\mathrm{i}} = 1 - \mathrm{F_j}(\mathrm{g_j}(p_{i,n}))$ and set

$$u_{j,n} = \int_{\mathrm{g_j}\left(p_{i,n}\right)}^{p_{j,n}} \left[\mathrm{f_j}\left(p'_j\right)/A_{j|i}\right] \mathrm{d}p'_j = \left(\frac{1}{A_{j|i}}\right) \left[\mathrm{F_j}\left(p_{j,n}\right) - \mathrm{F_j}\left(\mathrm{g_j}\left(p_{i,n}\right)\right)\right]$$

$$\text{and } p_{j,n} = \mathrm{F_j}^{-1}\left(u_{j,n} A_{j|i} + \mathrm{F_j}\left(\mathrm{g_j}\left(p_{i,n}\right)\right)\right).$$

4.4.1.7 Unknown Monotone Increasing or Decreasing Functional Relationship

If the state of knowledge dependence is due to an unknown functional relationship (see Sect. 3.5.2.6) between the uncertain data P_j and P_i, simple random sampling according to the marginal distribution of the so-called free datum P_i involves the random sampling of a value $u_{i,n}$ according to the uniform distribution over the unit interval and the subsequent transformation of the sampled value by the inverse of the cumulative distribution function.

In the case of an unknown monotone increasing relationship

$$p_{j,n} = \mathrm{F_j}^{-1}(u_{i,n})$$

and in the case of an unknown monotone decreasing relationship

$$p_{j,n} = \mathrm{F_j}^{-1}(1 - u_{i,n}).$$

4.4.1.8 Known Functional Relationship

In the case of state of knowledge dependence due to a known functional relationship $P_j = \mathrm{g_j}(P_i)$, a value $p_{i,n}$ is sampled according to the marginal distribution for P_i, and the corresponding value $p_{j,n}$ is found through evaluation of the relationship using $p_{i,n}$.

For instance, in the case of uncertain data that are fractions P_l, $l = 1, \ldots, L$ with the functional relationship

$$\sum\nolimits_{l=1}^{L} P_l = 1, \;\; 0 \le P_l \le 1, P_j = 1 - \sum\nolimits_{l=1}^{L-1} P_l = P_L$$

a simple random sample of values $u_{l,n}$, $l = 1, \ldots, (L-1)$ is taken independently according to the uniform distribution over (0, 1) and is turned into sample values $q_{j,n}$ by using the inverse functions of the marginal distributions specified for the uncertain conditional fractions Q_l, $l = 1, \ldots, (L-1)$ that are state of knowledge independent (see Sect. 3.5.2.3). The actual sample values $p_{l,n}$, $l = 1, \ldots, (L-1)$, are derived for the uncertain fractions P_l by post-processing the sampled conditional fractions according to

$$\begin{aligned} p_{1,n} &= q_{1,n} \\ p_{2,n} &= \left(1 - q_{1,n}\right) q_{2,n} \\ &\cdot \\ &\cdot \\ &\cdot \\ p_{L-1,n} &= \left(1 - q_{1,n}\right) \ldots \left(1 - q_{L-2,n}\right) q_{L-1,n} \end{aligned} \tag{4.5}$$

and the known functional relationship between the L uncertain data is used in

$$p_{j,n} = p_{L,n} = 1 - \sum\nolimits_{l=1}^{L-1} p_{l,n}.$$

4.4.1.9 Correlation Coefficients (Pearson)

Necessary condition is–as explained in Sect. 3.5.2.4–that the symmetric matrix $\boldsymbol{R}$ of specified correlations is positive definite.

To account for state of knowledge dependence expressed by the matrix $\boldsymbol{R}$ of Pearson's correlations, a procedure may be adopted that is based on its property to completely characterize the dependence structure of the multivariate normal distribution. In the case of K state of knowledge dependent uncertain data, with L the set of their indices, the procedure finds a K-variate distribution such that its marginal distributions are the given marginal subjective probability distributions $\mathrm{F_l}$, $l \in L$, and the correlation coefficients are the given values $r_{i,j}$, $j > i$ with $j, i \in L$. This problem is solved as follows. First the P_l, $l \in L$, are transformed into standard normal distributed variates

$$Z_l = \Phi^{-1}(\mathrm{F_l}(\mathrm{P_l})), l \in L$$

using the inverse Φ^{-1} of the standard normal distribution function. Secondly, it searches for a value $\mathrm{r}(\mathrm{Z_i},\mathrm{Z_j})$ such that

$$P_i = \mathrm{F_i}^{-1}(\Phi(Z_i)) \text{ and } P_j = \mathrm{F_j}^{-1}\left(\Phi\left(Z_j\right)\right)$$

satisfy $\mathrm{r}(P_i,P_j) = r_{i,j}$ as specified. The symmetric $(K \times K)$ correlation matrix $\boldsymbol{R}_z$ with elements $\mathrm{r}(Z_i,Z_j)$, $i,j \in L$ is required to be positive definite. $\mathrm{r}(Z_i,Z_j)$ can be obtained analytically for some types of marginal distributions $\mathrm{F_i}$, $\mathrm{F_j}$. For others, an iterative procedure will be required. As is explained in Sect. 3.5.2.4, the procedure will not have a solution for any combination of ($\mathrm{F_i}$, $\mathrm{F_j}$, $r_{i,j}$). Failure to find a solution indicates that an improper triplet ($\mathrm{F_i}$, $\mathrm{F_j}$, $r_{i,j}$) has been specified in Step 2 of the analysis.

After the above procedure was performed for all pairs of the K, uncertain data sampling is easily done as follows:

1. Take a simple random sample of values $u_{k,n}$ ($k = 1, \ldots, K$; $n = 1, \ldots, N$) independently drawn according to the uniform distribution over (0, 1) and transform each value into a sample value $x_{k,n}$, from the standard normal distribution, i.e. $x_{k,n} = \Phi^{-1}(u_{k,n})$. Decompose $\boldsymbol{R}_z$ into $\boldsymbol{DD}'$ where $\boldsymbol{D}$ is a lower triangular matrix (standard software packages are available) and obtain the vector $\boldsymbol{z_n} = \boldsymbol{Dx_n}$, $n = 1, \ldots, N$. The components $z_{k,n}$ are a random sample from the K-dimensional multivariate normal distribution with the K by K (K rows and K columns) correlation matrix $\boldsymbol{R}_z$. The multivariate normal distribution serves as a so-called normal copula $\mathrm{C}(\boldsymbol{u}, \boldsymbol{R}_z) = \Phi(\boldsymbol{\Phi^{-1}(u)};\boldsymbol{R}_z)$ (see Sect. 3.5.2.7) for the joint distribution of the uncertain data P_l, $l \in L$.
2. Obtain values $p_{l,n}$, $l \in L$, of the uncertain data P_l through element-wise transformation $p_{l,n} = \mathrm{F_l}^{-1}(\Phi(z_{k,n}))$ where l is the k-th index in the set L. The values $p_{l,n}$, $l \in L$, may be considered as randomly sampled according to a multivariate distribution with the marginal distributions and pairwise Pearson's correlations as specified. Their joint density function is $\mathrm{f}(\boldsymbol{p}) = c(\boldsymbol{u}, \boldsymbol{R_Z})\prod_{l \in L}\mathrm{f}_l(p_l)$, and $\mathrm{c}(\boldsymbol{u}, \boldsymbol{R}_z)$ is the derivative of the normal copula.

4.4.1.10 Rank Correlation Coefficients (Spearman)

The procedure for the specified Spearman's correlation coefficients $r^s(P_i,P_j)$ is very much the same as for Pearson's correlation coefficients. Once $r(Z_i,Z_j)$ is found from Eq. (4.6) such that $r^s(Z_i,Z_j)$ is as specified for P_i, P_j, the K-variate normal distribution to be used in the sampling process described above is fully determined. As $r(Z_i,Z_j)$ may have any value from $(-1, +1)$ and the relationship (Kruskal 1958)

$$r^S\left(Z_i, Z_j\right) = (6/\pi)\ \arcsin\left(r\left(Z_i, Z_j\right)/2\right) \tag{4.6}$$

holds for all values $r(Z_i,Z_j)$, there are no restrictions for the specified values $r^s(P_i,P_j)$ other than $-1 < r^s(P_i,P_j) < 1$. The symmetric $K \times K$ matrix $\boldsymbol{R}^s$ of rank correlations needs to be positive definite and so does the symmetric matrix $\boldsymbol{R}_z$ of Pearson's correlation coefficients for the reasons given in Sect. 3.5.2.4. The decomposition of $\boldsymbol{R}_z$ into the product $\boldsymbol{D}\boldsymbol{D}'$ (Cholesky decomposition—standard software packages available) will fail if the matrix is not positive definite.

Since r^s is invariant under monotone transformations, the operations $F_l^{-1}(\Phi(Z_k))$ (l is the k-th index in the set L) do not change its value.

4.4.1.11 Copulas

The copula approach for K state of knowledge dependent uncertain data involves the independent random sampling of K values $u_{k,n}$, $k = 1, \ldots, K$ according to a uniform distribution over the unit interval. How these values are processed into sample values of the K state of knowledge dependent uncertain data is explained in Sect. 4.4.1.9 above for the normal copula and in the discussion of Sect. 3.5.2.7 for the use of pair copulas.

4.4.1.12 Dependence Modeling with Elementary Uncertainty Contributors

Random sampling is performed at the level of the I state of knowledge independent elementary uncertainty contributors, of the set L of state of knowledge dependent uncertain data, by obtaining a random sample of values $u_{i,n}$, $i = 1, \ldots, I$ independently drawn according to the uniform distribution over the unit interval and by transforming it into a sample of values $z_{i,n}$, $i = 1, \ldots, I$ independently drawn according to the standard normal distribution. See Sect. 3.5.2.8 for a description of the subsequent steps to sample values for the state of knowledge dependent uncertain data.

4.4.1.13 A Multivariate Simple Random Sample

If J sets of values $p_{l,j}, j = 1, \ldots, J; l \in L$, are available as K-variate sample of the state of knowledge dependent uncertain data P_l, $l \in L$, the K-tuples are numbered arbitrarily as 1 to J, and the K data are represented by a quantity Q uniformly (or otherwise if there is reason to give preference to some of the sets) distributed over the interval $(0, J)$. Let F_Q be the distribution function of Q over the interval $(0, J)$. The simple random sampling method then deals with Q instead of P_l, $l \in L$. A value q_n is drawn at random according to the distribution F_Q, and the value q_n is rounded to q_n', the nearest integer that is larger than q_n, and q_n' is replaced in the sample by the K-tuple of sample values $p_{l,j}$, $l \in L$, with $j = q_n'$.

Note, that some of these K uncertain data could be sets of table entries, matrices, vectors or function values to a sequence of values of an independent variable like time and/or space.

4.4.1.14 Pearson's and Spearman's Correlation Coefficients Defined as Sample Measures of Association

The correlation coefficients discussed in the previous subsections were defined as population measures of association. They apply to the whole population of all possibly true pairs of values for P_i and P_j. A random sample taken according to a joint distribution for P_i and P_j, that satisfies the specified marginal distributions and the population measure for state of knowledge dependence, will exhibit a sample correlation value that differs from the correlation value specified for the population. Sometimes, the measure of association is, therefore, specified as a sample measure, i.e. it is specifically required that the sample correlation value is as specified. In this case, the sampling process is carried out such that the sample satisfies this requirement. The disadvantage is, however, that the sample is no longer a simple random sample but has a deterministic component—the specified sample measure of association. Some useful uncertainty statements derived in Chap. 5 for model results require the sample to be random. They can, therefore, not be obtained if sampling is done according to sample measures of association.

Pearson's sample correlation:
Instead of steps 1 and 2 of Sect. 4.4.1.9, the following steps are executed:

1. Take a simple random sample of values $u_{k,n}$, $k = 1, \ldots, K$; $n = 1, \ldots, N$ independently drawn according to the uniform distribution over (0, 1) and transform each value into a sample value $x_{k,n}$, from the standard normal distribution. The resulting $(K \times N)$ matrix is denoted by $\boldsymbol{X}^{(0)}$.
2. Compute the sample correlation matrix $\boldsymbol{R}^{(0)}$ of the rows in the $K \times N$ matrix $\boldsymbol{X}^{(0)}$.
3. Decompose $\boldsymbol{R}^{(0)}$ and $\boldsymbol{R}_z$ (see Sect. 4.4.1.9) into the products $\boldsymbol{R}^{(0)} = \boldsymbol{Q}\boldsymbol{Q}'$ and $\boldsymbol{R}z = \boldsymbol{D}\boldsymbol{D}'$ of the lower triangular $(K \times K)$ matrix $\boldsymbol{Q}$ (resp. $\boldsymbol{D}$) and its transpose $\boldsymbol{Q}'$ (resp. $\boldsymbol{D}'$) (Cholesky decomposition).

4. Determine the $(K \times N)$ matrix $\boldsymbol{Z} = \boldsymbol{D}\boldsymbol{Q}^{-1}\boldsymbol{X}^{(0)}$. $\boldsymbol{Z}$ will have the desired correlation structure.
5. Obtain values $p_{l,n}$, $l \in L$, of the uncertain data P_l through element-wise transformation $p_{l,n} = \mathrm{F}_1^{-1}(\Phi(z_{k,n}))$. The index l is the k-th index in the set L and Φ is the standard normal distribution function.

The values $p_{l,n}$, $l \in L$, may be considered as sampled according to a multivariate distribution with the marginal distributions and pairwise Pearson's sample correlations as specified.

Spearman's sample correlation:

If Spearman's sample rank correlation coefficient $r_{i,j}^S$ is specified for the uncertain data P_i and P_j, then pairs of values $(p_{i,1},p_{j,1})$,$(p_{i,2},p_{j,2})$, …, $(p_{i,N},p_{j,N})$ should be drawn such that the resulting empirical or sample rank correlation coefficient is as specified. The desired sample rank correlation structure, for any set L of $K \geq 2$ state of knowledge dependent uncertain data, is given by the symmetric rank correlation matrix $\boldsymbol{R}^{\boldsymbol{s}}$ with $r_{i,i}^S = 1$ and the specified elements $r_{i,j}^S$, $j > i$; $j, i \in L$. It is achieved through the following steps (Iman and Conover 1982):

1. Draw a random sample $(u_{k,n}, k = 1, ..., K; n = 1, \ldots, N)$, i.e. N sets of K values each, independently according to the uniform distribution over the unit interval. Transform the NK values by using the inverse functions of the marginal distributions that express the state of knowledge for each of the K data independently. The resulting $(K \times N)$ matrix is denoted by $\boldsymbol{X}^{(0)}$.
2. Express the state of knowledge dependence between any pair of the K uncertain data by the respective sample rank correlation coefficient $r_{i,j}^S$, $j > i$ with $j, i \in L$ and $r_{j,i}^S = r_{i,j}^S$;
3. Set up the $(K \times K)$ matrix $\boldsymbol{R}^{\boldsymbol{s}}$ of rank correlation values; $\boldsymbol{R}^{\boldsymbol{s}}$ is required to be positive definite in order to be a correlation matrix.
4. Set up an $(K \times N)$ random rank matrix $\boldsymbol{W}^{(0)}$, where every one of the K rows is an independent random permutation of the integer numbers 1 to N, and determine the corresponding $(K \times K)$ sample rank correlation matrix $\boldsymbol{R}^{(0)}$.
5. Decompose $\boldsymbol{R}^{(0)}$ into the product $\boldsymbol{R}^{(0)} = \boldsymbol{Q}\boldsymbol{Q}'$ of the lower triangular $(K \times K)$ matrix $\boldsymbol{Q}$ and its transpose $\boldsymbol{Q}'$ (Cholesky decomposition).
6. Decompose $\boldsymbol{R}^{\boldsymbol{s}}$ into the product $\boldsymbol{R}^{\boldsymbol{s}} = \boldsymbol{D}\boldsymbol{D}'$ with $\boldsymbol{D}$ a lower triangular $(K \times K)$ matrix (Cholesky decomposition).
7. Determine the $(L \times N)$ matrix $\boldsymbol{W}^{(1)} = \boldsymbol{D}\boldsymbol{Q}^{-1}\boldsymbol{W}^{(0)}$. $\boldsymbol{W}^{(1)}$ will have the desired rank correlation structure. Its elements will not be integer.
8. Determine the $(K \times N)$ rank matrix $\boldsymbol{W}^{(2)}$ of $\boldsymbol{W}^{(1)}$.
9. Order the N sample values of P_l, $l \in L$ in $\boldsymbol{X}^{(0)}$ according to the ranking given by the kth row (the index l is the k-th index in the set L) of the rank matrix $\boldsymbol{W}^{(2)}$. The resulting matrix $\boldsymbol{X}^{(1)}$ is obtained from $\boldsymbol{X}^{(0)}$ through this permutation of the numbers within each row of $\boldsymbol{X}^{(0)}$. The matrix $\boldsymbol{R}^{(1)}$ of sample rank correlation coefficients of the rows of $\boldsymbol{X}^{(1)}$ will be almost equal to $\boldsymbol{R}^{\boldsymbol{s}}$. The N columns of K sample values each, of the matrix $\boldsymbol{X}^{(1)}$, may then be used for state of knowledge propagation through the model.

4.4.2 *Latin Hypercube Sampling (LHS)*

Latin Hypercube Sampling (LHS) (McKay et al. 1979) is not a purely random sampling technique. It is partly deterministic. Instead of drawing N values at random according to the uniform distribution over the unit interval (0, 1), as in Simple Random Sampling, LHS divides the unit interval into N sections of equal length. In one version of LHS, the values at the center of each section (the conditional median values) are taken as sample values $u_1, \ldots, u_N$. These sample values cover the unit interval equally. The set of sample values

$$\left\{ \left(p_{m,n} = \mathrm{F_m}^{-1}(u_n) \right) \middle| \, m = 1, \ldots, M; n = 1, \ldots, N \right\}$$

is then used in the Monte Carlo simulation. The function $\mathrm{F_m}$ is the cumulative subjective probability distribution function specified as state of knowledge expression for the uncertain datum P_m. In another version, exactly one value is sampled at random from each of the N sections for each of the M uncertain data and according to the uniform distribution over the section. The set of sample values $\{(p_{m,n} = \mathrm{F_m}^{-1}(u_{m,n}))| \, m = 1, \ldots, M; n = 1, \ldots, N\}$ is then used in the Monte Carlo simulation.

This procedure, applied for two state of knowledge independent uncertain data, divides the unit square in a chessboard like manner into N rows and N columns. LHS requires the two sets of values $\{u_{1,n} | n = 1, \ldots, N\}$ and $\{u_{2,n} | n = 1, \ldots, N\}$ to be combined at random to pairs of values $(u_{1,in}, u_{2,jn})$. The index sets $\{i_n | n = 1, \ldots, N\}$ and $\{j_n | n = 1, \ldots, N\}$ are random permutations of the set of index values $\{n | n = 1, \ldots, N\}$. The permutations are obtained by choosing for each datum N values at random according to the uniform distribution over (0, 1) and by taking their ranks as index sets. The pairs of values $(u_{1,\mathrm{in}}, u_{2,\mathrm{jn}})$ are entered into the corresponding field of the unit square, with the index $\mathrm{i_n}$ as the row index and the index $\mathrm{j_n}$ as the column index. There is exactly one pair of values in each row and in each column of the square. This is called a "Latin Square" sample. The name "Latin Square" is attributed to the mathematician Euler who entered letters from the Latin and Greek alphabets, paired at random, into each field of a checkered array of equal numbers of rows and columns such that each letter of each alphabet appeared exactly once in each row and each column. In the case of three uncertain data, one would have to speak of a "Latin Cube" sample, while for all dimensions higher than three the sample may be rightfully called a "Latin Hypercube" sample.

4.4.2.1 Estimates of Measures of Uncertainty for Model Results

The N sample values $\{y_{j,n} | \, j = 1, \ldots, J\}$, $n = 1, \ldots, N$, for J model results, obtained using an LHS of the uncertain data, neither constitute an LHS nor an SRS. It follows from the construction principle of an LHS that they are not probabilistically independent (Hora 2003). Therefore, the distributions of estimates of measures of model result uncertainty are not known. For instance, the uncertainty measure could be the

subjective probability content between two possibly true values a and b of the model result. The estimate would then be

$$\hat{e} = \left(\frac{1}{N}\right) \sum_{n=1}^{N} \mathbf{I}_{(a<y<b)}(y_n)$$

with the value of $\mathbf{I}_{(a < y < b)}(y_n)$ equal to 1 if $a < y_n < b$ and zero otherwise.

The distribution of the estimate is unknown because the N model result values y_n are, due to the construction principle of an LHS, not independently sampled. Specifically, it is not known how to arrive at the probability for the r-th ordered sample value of a model result to be above the $u\%$ quantile of the subjective probability distribution for that model result. This precludes the computation of statistical tolerance limits and intervals. Confidence statements (confidence intervals with specified confidence level percentage) for the true mean value, variance, subjective probability content of specified intervals and two-sided or one-sided upper or lower intervals with specified subjective probability content are consequently not available for the subjective probability distributions of the model results. This is a serious drawback for uncertainty analysis using LHS since typically affordable sample sizes are only moderate making confidence statements a desirable feature of the analysis results. It has been suggested (Hansen et al. 2012) to obtain a sample of K estimates of the uncertainty measure where each estimate in the sample is derived from an LHS of size N/K independently chosen at random. The K estimates may then be used to derive confidence statements for the true value of the uncertainty measure in question.

4.4.2.2 Extensions of the Sample Size

Since confidence statements are not possible, one might want to study the convergence of estimates of uncertainty measures of the model result with increasing sample size. However, different to SRS, it is not possible to extend an LHS of sample size N to one of size $N' > N$. The construction principle of an LHS makes it necessary to start a new sample of size N'.

4.4.2.3 LHS and State of Knowledge Dependence

Sections 4.4.1.3–4.4.1.14 discuss SRS in various situations of state of knowledge dependence and for various ways of state of knowledge dependence modeling. The same situations and ways of modeling are now discussed with reference to drawing an LHS.

– State of knowledge dependence on interim results

  The set of sample values $u_{j,n}$, $n = 1, \ldots, N$ is drawn as part of the LHS.

– State of knowledge dependence on previously sampled values

The set of sample values $u_{j,n}$, $n = 1, \ldots, N$ is drawn as part of the LHS.

– Conditional distributions

Sample values $u_{j,n}$, $n = 1, \ldots, N$ are drawn for P_j as part of the LHS and $p_{j,n} = F_{jli}^{-1}(u_{j,n})$. F_{jli} is the conditional distribution for P_j given the sample value of P_i.

– Constraints

The first option is, for obvious reasons, not available for LHS. LHS for the second option proceeds just as with SRS except that the sample values $u_{j,n}$, $n = 1, \ldots, N$ are drawn for P_j as part of the LHS.

– Unknown monotone increasing or decreasing functional relationship

The set of sample values $u_{i,n}$, $n = 1, \ldots, N$ is drawn as part of the LHS.

– Known functional relationship

Sample values $u_{l,n}$, $l = 1, \ldots, L-1$, $n = 1, \ldots, N$ are independently drawn for Q_l, $l = 1, \ldots, L-1$ as part of the LHS and $q_{l,n} = F_l^{-1}(u_{l,n})$. The function F_l^{-1} is the inverse of the distribution function for Q_l.

– Correlation coefficients (Pearson) and rank correlation coefficients (Spearman)

Sample values $u_{k,n}$, $k = 1, \ldots, K$, $n = 1, \ldots, N$ are independently drawn for $z_{k,n}$, $k = 1, \ldots, K$ as part of the LHS and the subsequent steps are just as with simple random sampling (SRS).

– Copulas

The copula approach for K state of knowledge dependent uncertain data involves the independent random sampling of K values $u_{k,n}$, $k = 1, \ldots, K$ as part of the LHS. How these values are processed into sample values of the K state of knowledge dependent uncertain data is explained in Sect. 4.4.1.9 for the normal copula and in the discussion of Sect. 3.5.2.7 for the use of pair-copulas.

– Dependence modeling with elementary uncertainty contributors

The sample values $u_{i,n}$, $i = 1, \ldots, I$, $n = 1, ..., N$ (see Sect. 4.4.1.12—where I is the total number of elementary uncertainty contributors used in the dependence modeling) are drawn independently as part of the LHS.

– A multivariate simple random sample

The sample values q_n, $n = 1, ..., N$ (see Sect. 4.4.1.13) are drawn as part of the LHS.

– Spearman's and Pearson's correlation coefficients defined as sample measures of association

The sample values $u_{k,n}$, $k = 1, ..., K$; $n = 1, ..., N$ are drawn independently for each index k as part of the LHS.

4.4.3 Importance Sampling

Some uncertainty analyses are performed to estimate the subjective probability content of a critical value range C_Y of the model result. If the probability content is very small (10^{-3} or lower) large, often unaffordable sample sizes will be needed for a probability estimate with sufficiently narrow confidence interval.

The range C_Y could comprise all values that

– exceed a specified value y_{lim}

or

– do not exceed y_{lim}

or

– are not smaller than y_l and not larger than y_u.

The latter criterion can be converted to

all values y^* that do not exceed y_{lim}^*
with $y^* = | y - (y_u + y_l)/2|$ and $y_{lim}^* = (y_u - y_l)/2$.

It should be noted that the limiting values could themselves be functions of some of the uncertain data.

An estimate for the subjective probability content $u = \mathrm{sw}(Y \subset C_Y)$ is obtained as

$$\hat{u} = \left(\frac{1}{N}\right) \sum_{n=1}^{N} \mathbf{I}_{(y \epsilon C_Y)}(y_n) \tag{4.7}$$

$\mathbf{I}_{(y \epsilon C_Y)}(y_n) = 1$ if $y_n \in C_Y$ and is 0 otherwise,

N is the size of a simple random sample, and "sw" stands for "subjective probability".

An upper $v\%$ confidence limit for u is given as (Heinhold and Gaede 1968)

$$(L+1)a/((L+1)a + N - L)$$

while a lower $v\%$ confidence limit for u is given as

$$L/(L + (N - L + 1)b)$$

and a $v\%$ confidence interval is given as

$$L/(L + (N - L + 1)c), (L+1)d/((L+1)\mathrm{d} + N - L)$$
$$L = \sum_{n=1}^{N} \mathrm{I}_{(y \epsilon C_Y)}(y_n) \text{ and}$$

$$
\begin{aligned}
a &= \mathrm{q}_{v,k1,k2} && \text{with} \quad k_1 = 2(L+1), k_2 = 2(N-L) \\
b &= \mathrm{q}_{v,k1,k2} && \text{with} \quad k_1 = 2(N-L+1), k_2 = 2L \\
c &= \mathrm{q}_{(100+v)/2,k1,k2} && \text{with} \quad k_1 = 2(N-L+1), k_2 = 2L \\
d &= \mathrm{q}_{(100+v)/2,k1,k2} && \text{with} \quad k_1 = 2(L+1), k_2 = 2(N-L).
\end{aligned}
$$

Typical values for v are 95 or larger, and $\mathrm{q}_{v,k1,k2}$ is the v% quantile of the F distribution with k_1 and k_2 degrees of freedom. Tables of F distribution quantiles can be found in most textbooks on statistics (e.g. in Winkler and Hays 1975).

For instance, if $N = 100$ and none of the sample values y_n are in the critical region C_Y, then an upper 95% confidence limit for the subjective probability content $u = \mathrm{sw}(Y \subset C_Y)$ is obtained as

$$a/(a+N) = 2.9957/102,9957 = 0.029.$$

The probability is ≈ 0.905 to have none of the sample values y_n in C_Y if the subjective probability content of C_Y is 10^{-4} and the sample size is $N = 1.000$. For $N = 10{,}000$, this probability is still ≈ 0.368, i.e. on average one would have to be satisfied with 0.00029 as upper confidence limit for about every third sample of size 10,000. This upper limit might be satisfactory; however, a sample of size 10,000 is out of the question for most computer models of practical relevance.

A basic importance sampling procedure:
The idea behind importance sampling (Fishman 2000, Robert and Casella 2010) is to sample at random disproportionally much (in subjective probability) from the subspace $\boldsymbol{P}^{\boldsymbol{M}}_{\boldsymbol{crit}}$ of the value space $\boldsymbol{P}^{\boldsymbol{M}}$ of the epistemic uncertainties and to correct the obtained probability estimate for the introduced disproportionality. Two questions come immediately to mind:

- How to find the subspace $\boldsymbol{P}^{\boldsymbol{M}}_{\boldsymbol{crit}}$ where disproportionally more should be sampled than the state of knowledge would suggest?
- How to correct the resulting subjective probability estimate for the introduced disproportionality?

The simple importance sampling procedure presented below may serve as an illustration of the principle. It is based on SRS, and the discussion considers the case of no state of knowledge dependence among the involved uncertain data. The presented procedure may be adapted to account for the various state of knowledge dependence models of Sect. 3.5. As before and without loss of generality, the epistemic uncertainties (parameter values, model formulations and input data) are subsequently simply called "uncertain data".

To answer the first question, it is suggested to draw a first SRS of size N_0 (according to the subjective probability distribution specified over $\boldsymbol{P}^{\boldsymbol{M}}$) as described in Sect. 4.4.1 and to evaluate the model for each of the N_0 sample elements. The N_0 model result values $y_1, \ldots, y_{N_0}$ are then arranged in ascending order. The sample may contain none or at best one value from the critical value range C_Y, depending on how small the subjective probability content of that range is. If C_Y is the range of values

exceeding y_{lim}, then the statistical test in Sect. 6.4 may be used to identify the uncertain data that are mostly responsible for the top $w\%$ of sample values y_n where w may be chosen to be 10. The smallest value of the top $w\%$ (with none or only few values $>y_{lim}$) defines the lower boundary of a value range C_{Y_0} with $C_Y \subset C_{Y_0}$. The procedure is analogous for the other two types of critical value ranges mentioned above. Not only does this test identify the uncertain data mostly responsible, but it also suggests for each identified datum a value range that should be sampled disproportionally more often than the specified subjective probability distribution would suggest. Let the set of indices of the identified uncertain data be L. A new random sample of size N_1 is drawn according to a joint subjective probability density function that particularly emphasizes the value ranges identified by the test as mostly responsible for the $w\%$ of model result values contained in the value range C_{Y_0}. To this end, the importance sampling density functions $\mathrm{h}_{1,\mathrm{l}}(p_l)$, $l \in L$, are specified for the identified uncertain data. These density functions may be chosen such that random samples are easily drawn. Piecewise uniform density functions with $\left(1-\frac{w}{100}\right)$ probability for the range of data values identified by the test and $\frac{w}{100}$ for their complement may, for instance, serve the purpose. For all other uncertain data, the sample values are drawn according to the subjective probability distributions specified as their state of knowledge expressions, i.e. $\mathrm{h}_{1,\mathrm{m}}(p_m) = \mathrm{f}_{\mathrm{m}}(p_m)$. Should the test have identified two disconnected value ranges for an uncertain datum with the distance between these ranges larger than a chosen value (e.g. larger than the largest of the two ranges), then the probability $\left(1-\frac{w}{100}\right)$ will be proportionally shared between the two ranges. Otherwise, the smallest range that contains the two ranges receives the probability $\left(1-\frac{w}{100}\right)$.

A simple random sample of size N_1 is drawn according to the chosen importance sampling density function $\mathrm{h}_1(\boldsymbol{p}) = \prod_{m=1}^{M} h_{1,m}(p_m)$. The new set of N_1 model result values is again ordered by increasing magnitude and C_{Y_1} is defined. Again, there may be none or at best a few sample values y_n from the critical value range C_Y, and, consequently, a further iteration step is needed, i.e. the test is performed and $\mathrm{h}_2(\boldsymbol{p})$ is built according to the data and their value ranges found by the test and a sample of size N_2 is drawn according to $\mathrm{h}_2(\boldsymbol{p})$. After K iterations, with sample sizes N_1, N_2, ..., N_K very much smaller than $1/u$, hopefully many of the values in the sample of size N_K are from C_Y.

An estimate of $u = \mathrm{sw}(Y \subset C_Y)$ can be obtained as

$$u \approx \hat{u}_K = \left(\frac{1}{N_K}\right) \sum_{n=1}^{N_K} \mathbf{I}_{(Y \epsilon C_Y)}\left(y_n^{(K)}\right) \mathrm{f}\left(p_{1,n}^{(K)}, \ldots, p_{M,n}^{(K)}\right) / \mathrm{h_K}\left(p_{1,n}^{(K)}, \ldots, p_{M,n}^{(K)}\right) \tag{4.8}$$

with $\mathrm{f}(\boldsymbol{p})$ the joint density function specified as state of knowledge expression and $\mathrm{h_K}(\boldsymbol{p})$ the joint density function emphasizing the value ranges of the uncertain data mostly responsible for the model result values in the value range $C_{Y_{(K-1)}}$ while $y_n^{(K)} = \mathrm{r}\left(p_{1,n}^{(K)}, \ldots, p_{M,n}^{(K)}\right)$ are the model result values obtained with the sample of size N_K.

The correction term $c_n^{(K)} = \mathrm{f}\left(p_{1,n}^{(K)}, \ldots, p_{M,n}^{(K)}\right)/\mathrm{h_K}\left(p_{1,n}^{(K)}, \ldots, p_{M,n}^{(K)}\right)$ is the answer to the second question posed above. In the case of no state of knowledge dependence among the uncertain data, the correction term becomes

$$c_n^{(K)} = \prod_{m=1}^{M} \mathrm{f}_m\left(p_{m,n}^{(K)}\right)/\mathrm{h_{K,m}}\left(p_{m,n}^{(K)}\right) \tag{4.9}$$

or rather

$$c_n^{(K)} = \Pi_{l\in L}\mathrm{f}_l\left(p_{l,n}^{(K)}\right)/\mathrm{h}_{K,l}\left(p_{l,n}^{(K)}\right)$$

where l runs over the index set L of all uncertain data identified by the K-th test.

$\hat{u}_K$ is an estimate of

$$u = \mathrm{sw}\,(Y \subset C_Y) = \int_{p_{1,min}}^{p_{1,max}} \ldots \int_{p_{M,min}}^{p_{M,max}} \left[\mathbf{I}_{(Y\in C_Y)}(y)\mathrm{f}(p_1,\ldots,p_M)/\mathrm{h_K}(p_1,\ldots,p_M)\right]\mathrm{h_K}(p_1,\ldots,p_M)\mathrm{d}p_1\ldots\mathrm{d}p_M. \tag{4.10}$$

With

$$\mathrm{x}\left(y_n^{(K)}\right) = \mathbf{I}_{(Y\epsilon C_Y)}\left(y_n^{(K)}\right)\mathrm{f}\left(p_{1,n}^{(K)}, \ldots, p_{M,n}^{(K)}\right)/\mathrm{h_K}\left(p_{1,n}^{(K)}, \ldots, p_{M,n}^{(K)}\right)$$

$$\hat{u}_K = \left(\frac{1}{N_K}\right)\sum_{n=1}^{N_K} x\left(y_n^{(K)}\right) \text{ and } s_x^{(K)2} = (1/(N_K-1))\sum_{n=1}^{N_K}\left(x\left(y_n^{(K)}\right) - \hat{u}_K\right)^2$$

an upper v% confidence limit for the subjective probability content u of the critical range of model result values is obtained as

$$\hat{u}_K + t_{(N_K-1),v}s_x^{(K)}/N_K^{1/2}$$

and a lower v% confidence limit is obtained as

$$\hat{u}_K - t_{(N_K-1),v}s_x^{(K)}/N_K^{1/2}$$

while a v% confidence interval is obtained as

$$\left(\hat{u}_K - t_{(N_K-1),(100+v)/2}s_x^{(K)}/N_K^{1/2}, \hat{u}_K + t_{(N_K-1),(100+v)/2}s_x^{(K)}/N_K^{1/2}\right)$$

with $N_K \geq 30$ and $t_{(N_K-1),(100+v)/2}$ the [(100 + v)/2]% quantile of the Student or t distribution of $N_K - 1$ degrees of freedom.

Importance sampling will not only be of interest to estimate the very small probability content of a critical value range C_Y but also to estimate the mean value

of Y if it is predominantly determined by very large values from a range of very small subjective probability content.

The importance sampling densities used by this approach are easy to specify and to implement. However, the number of iterations needed to have a sufficient number of sample values y_n from C_Y may still be large. The literature (Ang et al. 1992; Au and Beck 2001; Morio 2011; Dai et al. 2012) therefore discusses methods that try to approximate the conditional probability density function over $\boldsymbol{P}_{\boldsymbol{crit}}^{\boldsymbol{M}}$ using a relatively small number of sample values from this subspace of $\boldsymbol{P}^{\boldsymbol{M}}$. The true conditional density function would have the values $\mathrm{h}^*(\boldsymbol{p}) = \mathrm{f}(\boldsymbol{p})/u$ for $\boldsymbol{p}$ with $y = r(\boldsymbol{p}) \in C_Y$ and zero otherwise. It is called the optimal importance sampling density function since

$$\int [\mathrm{f}(\boldsymbol{p})/\mathrm{h}^*(\boldsymbol{p})] h^*(\boldsymbol{p}) \mathrm{d}\boldsymbol{p} = u.$$

The integration runs over all $\mathrm{r}(\boldsymbol{p}) = y \in C_Y$ and the estimate obtained with this importance sampling density function

$$\hat{u} = \left(\frac{1}{N}\right) \sum\nolimits_{n=1}^{N} f(\boldsymbol{p_n})/h^*(\boldsymbol{p_n}) = u$$

would have zero variance.

Morio (2011) assumes that an initial importance sampling density function $\mathrm{h}_0(\boldsymbol{p})$ is available and provides a first number of sample values from $\boldsymbol{P}_{\boldsymbol{crit}}^{\boldsymbol{M}}$, while Au and Beck (2001) applies a Metropolis algorithm (Metropolis et al. 1953) [see (Au and Beck 2001) for application and modification] in order to populate $\boldsymbol{P}_{\boldsymbol{crit}}^{\boldsymbol{M}}$ with sample values drawn according to the conditional density function (provided the limiting distribution of the sample is in fact the conditional distribution mentioned above). A Gaussian kernel is centred at each of these sample values, and the weighted mixture is formed with the weights determined either by $\mathrm{f}(\boldsymbol{p})$ and $\mathrm{h}_0(\boldsymbol{p})$ (Morio 2011) or by density function estimation using support vectors (Dai et al. 2012; Rocco and Moreno 2002) to approximate the conditional density function over $\boldsymbol{P}^{\boldsymbol{M}}{}_{\boldsymbol{crit}}$.This approximation is then used as importance sampling density function to obtain a sample rich in model result values y_n from C_Y for the estimation of the probability content u of C_Y.

The application of a Metropolis algorithm and that of estimates of the conditional density function preclude the computation of confidence statements for u.

A general drawback of importance sampling is the danger that sampling will be locked into a once found critical region within $\boldsymbol{P}^{\boldsymbol{M}}$ while $\boldsymbol{P}_{\boldsymbol{crit}}^{\boldsymbol{M}}$ is made up of several such regions. The estimate of the conditional subjective probability density function over C_Y will be limited to the once found region and will therefore provide only an estimate of part of the subjective probability content of C_Y.

4.4.4 Subset Sampling

The motivation for this sampling technique is the same as for importance sampling. However, instead of the iterative direct estimation of the subjective probability content u of the critical value range C_Y, subset sampling obtains this probability from a sequence of subsets

$$C_{Y_0} \supset C_{Y_1} \supset C_{Y_2} \supset \ldots \supset C_{Y_K} \supset C_Y$$

as a product of estimates of the (conditional) subjective probabilities in the following expression

$$u = \mathrm{sw}(Y \subset C_Y) \leq \mathrm{sw}(Y \subset C_{Y_0}) \prod_{k=1}^{K} \mathrm{sw}\left(Y \subset C_{Y_k} \,\middle|\, Y \subset C_{Y_{(k-1)}}\right) = u_0 \prod_{k=1}^{K} u_{k|k-1}. \tag{4.11}$$

The simple subset sampling procedure presented below may serve as an illustration of the principle. It is based on SRS, and the discussion considers the case of no state of knowledge dependence among the involved uncertain data. The presented procedure may be adapted to account for the various state of knowledge dependence models of Sect. 3.5. As before and without loss of generality, the epistemic uncertainties (parameter values, model formulations and input data) are subsequently simply called "uncertain data".

Subset sampling is standard practice in risk assessments for technical facilities. Here, the sequence of subsets follows naturally from the technical context. The risk assessment starts with the estimation of the probability of an initiating event. This is followed by the estimation of the probability of safety system failure, given the initiating event. Next the probabilities of various sequences of accident progression within the facility, given the initiating event and the safety system failure, are estimated. Finally, an estimate is obtained of the probability of release of hazardous substances to the environment, given the considered accident progression sequence, system failure and initiating event. This way it is possible to estimate the very small probabilities of rare events as a product of the probability of the initiating event and of conditional probabilities.

A basic subset sampling procedure: The importance sampling procedure described in Sect. 4.4.3 is combined with subset sampling. To this end, sample values outside the value range $C_{Y_{(k-1)}}$ are discarded from each of the random samples of size N_k, $k = 1, \ldots, K$. A sequence of sets is successively defined in the course of the subset sampling process

$$C_{Y_{(-1)}} \supset C_{Y_0} \supset C_{Y_1} \supset \ldots \supset C_{Y_{(K-1)}} \supset C_{Y_K} \supset C_Y.$$

If C_Y is defined by an upper limit value y_{lim} for Y, that must not be exceeded, then C_Y contains all values y of Y with $y > \mathrm{y_{lim}}$.

The model assigns the value $y = r(\boldsymbol{p})$ to every vector $\boldsymbol{p} \in \boldsymbol{P}$ of possibly true values of the M uncertain data, and $\boldsymbol{P}$ is the carrier of the M-dimensional joint subjective probability distribution quantifying their state of knowledge. The set of all model result values, that are obtained this way, is denoted by $C_{Y_{(-1)}}$.

A first random sample of size N_0 is drawn according to the joint subjective probability distribution over $\boldsymbol{P}$. The N_0 model results are ordered by increasing value, and the smallest value of the top w% (with none or only few values $> y_{lim}$) defines the lower boundary of the subset C_{Y_0}. This provides the estimate

$$\hat{u}_0 = \frac{w}{100}$$

of the subjective probability $sw(Y \subset C_{Y_0})$. The statistical test of Sect. 6.4 may be used to identify the uncertain data that are mostly responsible for the top w% of sample values y_n, $n = 1, 2, \ldots, N_0$. The procedure is analogous for the other two types of critical value ranges mentioned in Sect. 4.3.

The test also suggests, for each identified datum, a value range that should be sampled disproportionally more often than the subjective probability distribution specified as state of knowledge expression would suggest. Let the set of indices of the identified uncertain data be L. A new random sample of size N_1 is drawn according to a joint subjective probability density function that particularly emphasizes the value ranges identified by the test as mostly responsible for the w% of model result values contained in the value range C_{Y_0}. To this end, the importance sampling density functions $h_{1,l}(p_l)$, $l \in L$, are specified for the identified uncertain data. These density functions may be chosen such that random samples are easily drawn. Piecewise uniform density functions with $\left(1 - \frac{w}{100}\right)$ probability for the range of data values identified by the test and $w/100$ for their complement may serve the purpose. For all other uncertain data P_k, $k \notin L$, $h_{1,k}(p_k) = f_k(p_k)$, *i. e.*the sample values are drawn according to their state of knowledge expressions. Should the test have identified two disconnected value ranges for one or more uncertain data with the distance between these ranges being larger than a chosen value (e.g. larger than the largest of the two ranges) then the probability $\left(1 - \frac{w}{100}\right)$ may be proportionally shared between the two ranges.

A simple random sample of size N_1 is drawn according to the chosen importance sampling density function $h_1(\boldsymbol{p})$. Different to the importance sampling procedure of Sect. 4.4.3, all model result values outside C_{Y_0} are discarded. The remaining number of sample values is $N_1' \leq N_1$. The new set of N_1' values is ordered by increasing magnitude. Again, there may be none or at best a few sample values y_n from the critical value range C_Y, and, consequently, a further iteration step is needed. The test is performed to identify those uncertain data that are mostly responsible for the top w% of the N_1' results in C_{Y_0} where the smallest of such values now defines the lower boundary of C_{Y_1}.

$$u_1 = \mathrm{sw}(Y \subset C_{Y_1}) = \mathrm{sw}(Y \subset C_{Y_1} | Y \subset C_{Y_0}) \mathrm{sw}(Y \subset C_{Y_0}) = u_{1|0} u_0$$

$$u_{1|0} = \mathrm{sw}(Y \subset C_{Y_1} | Y \subset C_{Y_0}) = \int_{p_{1,min}}^{p_{1,max}} \dots \int_{p_{M,min}}^{p_{M,max}} \left[\left(\mathbf{I}_{(y \in C_{Y_1})}(y) \mathrm{f}(p_1, \dots, p_M | Y \subset C_{Y_0}) \right. \right.$$
$$\left. \left. / \mathrm{h}_1(p_1, \dots, p_M | Y \subset C_{Y_0}) \right] \mathrm{h}_1(p_1, \dots, p_M | Y \subset C_{Y_0}) \mathrm{d}p_1 \dots \mathrm{d}p_M \right.$$

and

$y = r(p_1, \dots, p_M)$ is the model result.

$$u_{1|0} = E\left\{ \left(\mathbf{I}_{(Y \in C_{Y_1})}(Y) | Y \subset C_{Y_0} \right) \right\} = \mathrm{sw}(Y \subset C_{Y_1} | Y \subset C_{Y_0}),$$
$$\mathrm{f}(p_1, \dots, p_M | Y \subset C_{Y_0}) = \mathrm{f}(p_1, \dots, p_M) / u_0 \text{ and } u_0 = \mathrm{sw}(Y \subset C_{Y_0}).$$

An estimate $\hat{u}_{1|0}$ of $u_{1|0}$ is obtained from

$$\tilde{u}_{1|0} = \left(\frac{1}{N_1'} \right) \sum_{n=1}^{N_1'} \left(\mathbf{I}_{(y \in C_{Y_1})} \left(y_n^{(1)} \right) \mathrm{f}\left(p_{1,n}^{(1)}, \dots, p_{M,n}^{(1)} | Y \subset C_{Y_0} \right) \right) / \mathrm{h}_1\left(p_{1,n}^{(1)}, \dots, p_{M,n}^{(1)} | Y \subset C_{Y_0} \right)$$

with $\mathrm{f}\left(p_{1,n}^{(1)}, \dots, p_{M,n}^{(1)} | Y \subset C_{Y_0} \right) = \mathrm{f}\left(p_{1,n}^{(1)}, \dots, p_{M,n}^{(1)} \right) / u_0$ and $\mathrm{h}_1\left(p_{1,n}^{(1)}, \dots, p_{M,n}^{(1)} | Y \subset C_{Y_0} \right)$
$= \mathrm{h}_1\left(p_{1,n}^{(1)}, \dots, p_{M,n}^{(1)} \right) / \int \dots \int \mathrm{h}_1(p_1, \dots, p_M) \mathrm{d}p_1 \dots \mathrm{d}p_M$ where the integration runs over all $(p_1, \dots, p_M)$ with $y = \mathrm{r}(p_1, \dots, p_M) \in C_{Y_0}$.

Using $u_0 \approx \hat{u}_0 = \mathrm{w}/100$, $\int \dots \int \mathrm{h}_1(p_1, \dots, p_M) \mathrm{d}p_1 \dots \mathrm{d}p_M \approx N_1'/N_1$

$$\hat{u}_{1|0} = \left(\frac{1}{N_1'} \right) \sum_{n=1}^{N_1'} \left(\mathbf{I}_{(y \in C_{Y_1})} \left(y_n^{(1)} \right) \left[\mathrm{f}\left(p_{1,n}^{(1)}, \dots, p_{M,n}^{(1)} \right) / \hat{u}_0 \right] / \left[\mathrm{h}_1\left(p_{1,n}^{(1)}, \dots, p_{M,n}^{(1)} \right) / \left(N_1'/N_1 \right) \right] \right.$$
$$u_1 \approx \hat{u}_1 := \hat{u}_{1|0} \hat{u}_0. \tag{4.12}$$

The importance sampling density function $\mathrm{h}_2(\boldsymbol{p})$ is built according to the uncertain data and their value ranges found by the test as mostly responsible for the top w% model result values in C_{Y_1}. $\left(1 - \frac{w}{100}\right)$ is the probability content assigned over the identified ranges while their complement receives $w/100$. A sample of size N_2 is drawn according to $\mathrm{h}_2(\boldsymbol{p})$. All model result values outside C_{Y_1} are discarded, and an estimate $\hat{u}_2$ of u_2 is obtained from

$$u_2 = \mathrm{sw}(Y \subset C_{Y_2}) = \mathrm{sw}(Y \subset C_{Y_2} | Y \subset C_{Y_1}) \mathrm{sw}(Y \subset C_{Y_1} | Y \subset C_{Y_0}) \mathrm{sw}(Y \subset C_{Y_0})$$
$$= u_{2|1} u_{1|0} u_0$$

as

$$\hat{u}_2 := \hat{u}_0 \prod_{k=1}^{2} \hat{u}_{k|k-1}$$

$$= \hat{u}_0 \prod_{k=1}^{2} \left(1/N_k' \right) \sum_{n=1}^{N_k'} \mathbf{I}_{(y \in C_{Y_k})} \left(y_n^{(k)} \right) \left[\mathrm{f}\left(p_{1,n}^{(k)}, \dots, p_{M,n}^{(k)} \right) / \hat{u}_{k-1} \right] / \left[\mathrm{h}_k\left(p_{1,n}^{(k)}, \dots, p_{M,n}^{(k)} \right) / \left(N_k'/N_k \right) \right].$$

After K iterations, with each of the sample sizes N_0, N_1, N_2, ..., N_K very much smaller than $1/u$ and with $C_{Y_0} \supset C_{Y_1} \supset \ldots \supset C_{Y_{(K-1)}} \supset C_Y$, hopefully many of the values in the sample of size N_K' are from C_Y. Defining $C_{Y_K} := C_Y$ (i.e. all values that are larger than y_{lim}), an estimate of $u = \text{sw}(Y \subset C_Y)$ can be obtained from

$$u = \text{sw}(Y \subset C_Y) = u_0 \prod_{k=1}^{K} u_{k|k-1} \approx \hat{u}_K = \hat{u}_0 \prod_{k=1}^{K} \hat{u}_{k|k-1} \text{ with}$$

$$\hat{u}_{k|k-1} = \left(\frac{1}{N'_k}\right) \sum_{n=1}^{N'_k} \left(\mathbf{I}_{\left(y \in C_{Y_k}\right)}\left(y_n^{(k)}\right)\right) \left[\frac{\text{f}\left(p_{1,n}^{(k)}, \ldots, p_{M,n}^{(k)}\right)}{\hat{u}_0 \prod_{j=1}^{k-1} \hat{u}_{j|j-1}}\right] / \left[\text{h}_\text{k}\left(p_{1,n}^{(k)}, \ldots, p_{M,n}^{(k)}\right) / \left(N'_k / N_k\right)\right] \tag{4.13}$$

$$\hat{u}_K := \hat{u}_0 \prod_{k=1}^{K} \left(1/N_k \hat{u}_{k-1}\right) \sum_{n=1}^{N'_k} \mathbf{I}_{\left(y \in C_{Y_k}\right)}\left(y_n^{(k)}\right) \left[\text{f}\left(p_{1,n}^{(k)}, \ldots, p_{M,n}^{(k)}\right) / \text{h}_\text{k}\left(p_{1,n}^{(k)}, \ldots, p_{M,n}^{(k)}\right)\right]$$

Approaches to subset sampling are, for instance, presented in Au and Beck (2001) and Song et al. (2009). In both papers, the estimate $\hat{u}_0$ of the probability content of C_{Y_0} is derived from the initial simple random sample of size N_0 and the definition of C_{Y_0}. In order to generate sample values $y_n^{(k)}$ from subset $C_{Y_{(k-1)}}$, $k = 1, \ldots, K$, additional to those contained in the sample of size N_k (Au and Beck 2001), apply a modified Metropolis algorithm (Metropolis et al. 1953; Calvetti and Somersalo 2007). The distribution of the sample values (original plus those generated by the Metropolis algorithm) will tend (with increasing number of Markov Chain steps of the algorithm) towards the conditional probability distribution over $C_{Y_{(k-1)}}$. Due to the generation of additional sample values by the Metropolis algorithm (using already existing sample values), the samples are not independent and confidence statements for u_k are not available. Song et al. (2009), on the other hand, generate additional sample values from subset $C_{Y_{(k-1)}}$ by locating the centre of an importance sampling density function at the sample value $y_n^{(k)}$ in $C_{Y_{(k-1)}}$ that has the highest probability density value among those from $C_{Y_{(k-1)}}$. The computation in Song et al. (2009) of the variance of the estimate $\hat{u}_K$ of u_K assumes independence of the estimates $\hat{u}_k$, $k = 1, \ldots, K-1$. Both (Au and Beck 2001; Song et al. 2009) do without probability density function estimation. Song et al. (2009) compares the subset/Metropolis and the subset/importance sampling combination for three sample problems. In these problems, the efficiency (total sample size required for acceptable accuracy) is shown to be high for both with the combination subset/importance sampling being slightly more efficient.

References

Ang, G. L., Ang, A. H.-S., & Tang, W. H. (1992). Optimal importance-sampling density estimator. *Journal of Engineering Mechanics, ASCE, 118*(6), 1146–1163.

Au, S.-K., & Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics, 16*, 263–277.

Calvetti, D., & Somersalo, E. (2007). *Introduction to Bayesian scientific computing*. New York: Springer Science+Business Media.

Dai, H., Zhang, H., & Wang, W. (2012). A support vector density-based importance sampling for reliability assessment. *Reliability Engineering and System Safety, 106*, 86–93.

Fishman, G. S. (2000). *Monte Carlo – Concepts, algorithms and applications*, *Springer series in operations research*. New York: Springer.

Hansen, C. W., Helton, J. C., & Sallaberry, C. J. (2012). Use of replicated Latin hypercube sampling to estimate sampling variance in uncertainty and sensitivity analysis results for the geologic disposal of radioactive waste. *Reliability Engineering and System Safety, 107*, 139–148.

Heinhold, J., & Gaede, K.-W. (1968). *Ingenieur-Statistik*. München: R. Oldenbourg.

Hora, S. C. (2003). A distribution-free test for the relationship between model input and output when using Latin hypercube sampling. *Reliability Engineering and System Safety, 79*, 333–339.

Kruskal, W. H. (1958). Ordinal measures of association. *Journal of the American Statistical Association*, *53*, 814–861.

Iman, R. L., & Conover, W. J. (1982). A distribution-free approach to inducing rank correlations among input variables. *Communications in Statistics – Simulation and Computation, 11*(3), 311–334.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics, 21*(2), 239–245.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics, 21*, 1087–1092.

Morio, J. (2011). Non-parametric adaptive importance sampling for the probability estimation of a launcher impact position. *Reliability Engineering and System Safety, 96*, 178–183.

Robert, C. P., & Casella, G. (2010). *Monte Carlo statistical methods*, *Springer texts in statistics*. New York: Springer Science+Business Media.

Rocco, C. M., & Moreno, J. A. (2002). Fast Monte Carlo reliability evaluation using support vector machine. *Reliability Engineering and System Safety, 76*, 237–243.

Song, S., Lu, Z., & Qiao, H. (2009). Subset simulation for structural reliability sensitivity analysis. *Reliability Engineering and System Safety, 94*, 658–665.

Storlie, C. B., Swiler, L. P., Helton, J. C., & Sallaberry, C. J. (2009). Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering and System Safety, 94*, 1735–1763.

Winkler, R. L., & Hays, W. L. (1975). *Statistics: Probability, inference, and decision* (2nd ed.). New York: Holt, Rinehart and Winston.

Chapter 5
Step 4: Estimate Uncertainty

5.1 Introduction

Step 2 quantified and probabilistically expressed the states of knowledge at the level of parameter values, models and input data,[1] and Step 3 propagated this probabilistic expression through the arithmetic and logic instructions of the computer model. From this propagation follows a joint subjective probability distribution for the model results in a logically consistent way. It is the aim of Step 4 to derive, from this distribution, quantitative uncertainty statements for individual model results. If the joint subjective probability distribution is known, the marginal distribution can be obtained for each model result and uncertainty statements can be read directly from it. The marginal subjective probability distribution for a model result Y_j shall subsequently be simply called the subjective probability distribution for Y_j. It quantifies the combined effect of all uncertain data taken into account by the uncertainty analysis. The following analysis results are of specific interest:

(a) The $u\%$ quantile $y_{j,\,u\%}$ or the $(100-u)\%$ quantile $y_{j,\,(100\,-\,u)\%}$ of the subjective probability distribution for Y_j;
The $u\%$ quantile says that $y_{j,\,u\%}$ is not exceeded by Y_j with subjective probability $u/100$ or, in other words, $u\%$ of the population of model result values, which follow from the state of knowledge quantifications at the level of the uncertain data, do not exceed $y_{j,\,u\%}$. The $(100-u)\%$ quantile says that Y_j exceeds $y_{j,\,(100\,-\,u)\%}$ with subjective probability $u/100$.

(b) A $u\%$ subjective probability interval, delimited by two values such that $u\%$ of the population of model result values for Y_j, do neither exceed the upper value nor fall below the lower value (i.e. are contained).

[1] In what follows, uncertain parameters, models and input data are simply referred to as uncertain data since model uncertainties are represented by uncertain parameters and uncertain parameters are categorized as uncertain data.

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_5

(c) The percentage of the population of model result values for Y_j that exceed (do not exceed) a given upper (lower) limit value, i.e. do not comply with the limit.
(d) The subjective probability for a model result Y_j to be larger (smaller) than another model result Y_k (comparison of two or more model results).

The previous chapters used the term "possibly true" in connection with uncertain data of the computer model application. "Possibly true" stands for "cannot be excluded at the present state of knowledge". The purpose of the application of the model is to use the result (for instance, Y_j) as an answer to an assessment question. The population of model result values for Y_j, which follow from the combined effect of the state of knowledge quantifications at the *level of the uncertain data, is a population of possibly true answers to the assessment question under the condition that the following assumptions hold*:

– The (encoded) computer model is not seriously flawed.
– The most important contributors to the uncertainty of its results have been accounted for by the analysis (the combined effect of all those not accounted for may be neglected).
– The states of knowledge, at the level of the uncertain data, have been appropriately expressed by a joint subjective probability distribution.
– Any changes to the interpretation of "subjective probability" (see Sect. 3.4.1) are of only minor consequence.

If these assumptions hold, according to the judgement of the analyst, then the analysis results discussed above may be interpreted as follows:

– The $u\%$ quantile $y_{j,\ u\%}$ of the subjective probability distribution of the model result Y_j says that $u\%$ of the population of possibly true values for Y_j (i.e. $u\%$ of the population of possibly true answers to the assessment question) do not exceed this quantile value or, in other words, the true answer does not exceed $y_{j,\ u\%}$ with subjective probability $u/100$.
– A $u\%$ subjective probability interval, delimited by two values of the model result, is an interval that contains $u\%$ of the population of possibly true values for Y_j (i.e. $u\%$ of the population of possibly true answers to the assessment question).
– The percentage of the population of possibly true values for Y_j (i.e. of the population of possibly true answers to the assessment question) that exceed (do not exceed) the limit value, i.e. that do not comply with the upper (lower) limit.

So far it was assumed that the subjective probability distribution for Y_j is known. However, the methods of uncertainty propagation presented in the previous chapter do not provide the subjective probability distributions of the model results. They provide sets of model result values permitting only the estimation of some or all of the uncertainty statements mentioned above. Section 5.2 explains how these estimates are derived from the sets of model result values obtained by uncertainty propagation using simple random sampling (SRS), while those based on uncertainty propagation using Latin Hypercube sampling (LHS) are discussed in Sect. 5.3.

Estimates of the very small subjective probability content of a critical value range of the model result Y_j are discussed in Sects. 4.4.3 and 4.4.4.

5.2 Uncertainty Statements Available from Uncertainty Propagation Using SRS

An uncertainty analysis, using Monte Carlo simulation with SRS, provides in Step 3 a random input sample with each sample element independently drawn according to the joint subjective probability distribution of the uncertain data. Each input sample element is then transformed into the corresponding output sample element by the model. This way each element of the output sample is independently drawn at random according to the unknown joint subjective probability distribution of the model results (Table 5.1).

A random sample drawn according to the unknown subjective probability distribution of any individual model result Y_j is therefore immediately available as ($y_{j,1}$, $y_{j,2}$, . . . , $y_{j,N}$) and can be used to estimate mean value, variance, quantile values, intervals of given subjective probability content as well as subjective probabilities of violation of limit values. All of these estimates can be supplemented by confidence intervals and confidence limits. They quantify the possible influence of the sampling error on the estimates. This error is due to the random variation from sample to sample which is pronounced by the limited sample size affordable in practice. The following Sect. 5.2.1 explains the meaning of the confidence as well as tolerance confidence limits and intervals. Sections 5.2.2 to 5.2.5 derive estimates of mean value, quantile values, intervals of given subjective probability content as well as subjective probabilities of violation of limit values and supplement them with confidence limits and intervals to specified confidence levels. Section 5.2.6 investigates the variability of statistical tolerance limits. Sections 5.2.7 and 5.2.8 show how to compare different model results on the basis of their random output sample values. Finally, ways of graphical presentation of uncertainty estimates are shown in subsection 1 of Sect. 5.4.1 for single value model results and in subsection 2 of

Table 5.1 Output of the uncertainty propagation by Monte Carlo simulation using SRS

Computer model run no.	Input sample element no.	Uncertainty no.				Output sample element no.	Model result no.			
		1	2	. . .	M		1	2	. . .	J
		Sample values					Sample			
1	1	$p_{1,1}$	$p_{2,1}$	. . .	$p_{M,1}$	1	$y_{1,1}$	$y_{2,1}$	. . .	$y_{J,1}$
2	2	$p_{1,2}$	$p_{2,2}$	. . .	$p_{M,2}$	2	$y_{1,2}$	$y_{2,2}$	. . .	$y_{J,2}$
.	.			.		.			.	
.	.			.		.			.	
.	.			.		.			.	
N	N	$p_{1,N}$	$p_{2,N}$	. . .	$p_{M,N}$	N	$y_{1,N}$	$y_{2,N}$	. . .	$y_{J,N}$

Sect. 5.4.1 for model results that are a function of some independent variable such as time or space.

5.2.1 The Meaning of Confidence and Tolerance Confidence Limits and Intervals

– Confidence Limits, Intervals and Levels
It often does not suffice to provide just a single estimate (a so-called point value) for a fixed but unknown quantity θ. Intervals $[g_1, g_2]$ are preferred, particularly if it is possible to provide the degree of belief for θ to be in the interval $[g_1, g_2]$.

In statistics, θ is generally a characteristic value (like a parameter value) of the distribution F_X of a random variable X and is to be estimated from a simple random sample of size N, i.e. from a realization of the sample variables (X_1, X_2, …, X_N) where each X_n ($n = 1, 2 \ldots, N$) is independently and identically distributed according to F_X. In other words: The sample variables X_n will vary from sample to sample of size N according to the distribution F_X, and therefore, the estimate θ will also vary from sample to sample. Because of this variability, the value G of a function $h(X_1, X_2, \ldots, X_N)$ is also a random variable. If there are functions $h_1(X_1, X_2, \ldots, X_N) = G_1$ and $h_2(X_1, X_2, \ldots, X_N) = G_2$, such that $G_1 \leq \theta \leq G_2$ with probability $v/100$, then $[G_1, G_2]$ is a confidence interval at a confidence level of $v\%$ (a $v\%$ confidence interval). G_3 is an upper confidence limit at a level of $v\%$ (an upper $v\%$ confidence limit) if there is a function $h_3(X_1, X_2, \ldots, X_N) = G_3$ such that $\theta \leq G_3$ with probability $v/100$. The confidence interval and limit at a given level of $v\%$ are random variables because $(X_1, X_2, \ldots, X_N)$ is not a specific simple random sample but an N-dimensional random variable. This is why there is a probability $v/100$ for the fixed value θ to be within $[G_1, G_2]$. The practical meaning of this probability $v/100$ is that there is a chance of $v\%$ to draw a simple random sample $(x_1, x_2, \ldots, x_N)$ such that $g_1 \leq \theta \leq g_2$ where $g_1 = h_1(x_1, x_2, \ldots, x_N)$ and $g_2 = h_2(x_1, x_2, \ldots, x_N)$. So, once a simple random sample of size N is drawn and g_1 and g_2 are computed, the interval $[g_1, g_2]$ may or may not contain θ. However, because of what has been said above about the practical meaning of the associated confidence level, the degree of belief is $v\%$ that the interval $[g_1, g_2]$ does in fact contain θ**.** The argumentation is analogous for confidence limits. The interval $[g_1, g_2]$ is only a realization of the confidence interval $[G_1, G_2]$ but is also called confidence interval in practice. Clearly, if k simple random samples of size N are to be drawn and the corresponding k intervals $[g_1, g_2]$ computed, it is to be expected that $kv/100$ of these intervals will in fact contain the fixed but unknown value θ.

– Tolerance Confidence Limits, Intervals and Levels
The confidence statements given above are concerned with an unknown fixed value. Tolerance confidence statements, on the other hand, are concerned with a random variable X and in particular with the unknown limiting value(s) of a

one- (or two-) sided interval of values of X the probability content of which is at least as large as specified.

If there are functions $h_1(X_1, X_2, \ldots, X_N) = G_1$ and $h_2(X_1, X_2, \ldots, X_N) = G_2$, such that at a confidence level of $v\%$ $G_1 \leq X \leq G_2$ with at least probability $u/100$, then $[G_1, G_2]$ is a $(u\%, v\%)$ tolerance confidence interval for X. G_3 is an upper $(u\%, v\%)$ tolerance confidence limit if there is a function $h_3(X_1, X_2, \ldots, X_N) = G_3$ such that at a confidence level of $v\%$ $X \leq G_3$ with at least probability $u/100$. The tolerance confidence interval and limit are random variables because $(X_1, X_2, \ldots, X_N)$ is not a specific simple random sample but an N-dimensional random variable. This is why there is a probability $v/100$ for X to be within $[G_1, G_2]$ with at least probability $u/100$. The practical meaning of the probability $v/100$ is that there is a chance of $v\%$ to draw a simple random sample $(x_1, x_2, \ldots, x_N)$ such that $g_1 \leq X \leq g_2$ with at least probability $u/100$ where $g_1 = h_1(x_1, x_2, \ldots, x_N)$ and $g_2 = h_2(x_1, x_2, \ldots, x_N)$. So, once a simple random sample of size N is drawn and g_1 and g_2 are computed, the interval $[g_1, g_2]$ may or may not contain X with at least probability $u/100$ (may or may not contain at least $u\%$ of the population of values of X). However, because of what has been said above about the practical meaning of the associated confidence level of $v\%$, the degree of belief is $v/100$ that the interval $[g_1, g_2]$ does in fact contain X with at least probability $u/100$. The argumentation is analogous for tolerance confidence limits. The interval $[g_1, g_2]$ is only a realization of the tolerance confidence interval $[G_1, G_2]$ but, in practice, is also called tolerance confidence interval or more precisely statistical tolerance interval. Clearly, if k simple random samples of size N are to be drawn and the corresponding k intervals $[g_1, g_2]$ computed, it is to be expected that $kv/100$ of these intervals will in fact contain X with at least probability $u/100$ (or in other words, $kv/100$ of these intervals will in fact contain at least $u\%$ of the population of values of X).

An upper $(u\%, v\%)$ statistical tolerance limit is also an upper $v\%$ confidence limit for the $u\%$ quantile of X, and a lower $(u\%, v\%)$ statistical tolerance limit is a lower v % confidence limit for the $(100 - u)\%$ quantile of X. A two-sided $(u\%, v\%)$ statistical tolerance limit is a $v\%$ confidence interval delimiting a value range that contains X with at least probability $u/100$ (the range contains at least $u\%$ of the population of values of X at a confidence level of $v\%$).

5.2.2 The Mean Value of the Model Result

The mean value of the subjective probability distribution of a model result Y_j is estimated from a simple random sample (SRS) of size N as

$$m_{N,Y_j} = \left(\frac{1}{N}\right) \sum_{n=1}^{N} y_{j,n}.$$

The values $y_{j,\,n}$ are those in the output sample of the uncertainty propagation by Monte Carlo simulation. The vector $(y_{j,\,1}, y_{j,\,2}, \ldots, y_{j,\,N})$ of sample values is a realization of the random vector $(Y_{j,\,1}, Y_{j,\,2}, \ldots, Y_{j,\,N})$. Each component of the random vector has the subjective probability distribution of the model result Y_j. The larger the sample size N, the closer comes the distribution of the sum

$$M_{N,Y_j} = \left(\frac{1}{N}\right) \sum_{n=1}^{N} Y_{j,n}$$

to a normal distribution with mean value $\mathrm{E}\{Y_j\}$ and variance $\sigma_{Y_j}^2/N$. This follows from the central limit theorem. As a rule of thumb, this approximation by the normal distribution is considered satisfactory for sample sizes $N \geq 30$, provided the variance $\sigma_{Y_j}^2$ of the subjective probability distribution of the model result Y_j exists. Given the variance $\sigma_{Y_j}^2$ is known, an approximate v% confidence interval for the mean value of the model result is delimited by

$$\left(m_{N,Y_j} - a\right) \text{ and } \left(m_{N,Y_j} + a\right) \text{ with}$$
$$a = z_{(100+v)/2}\sigma_{Y_j}/N^{1/2}.$$

$z_{(100\,+\,v)/2}$ is the $[(100+v)/2]$% quantile of the standard normal distribution. In practice however, σ_{Y_j} needs to be estimated from the sample values and the estimate is given by

$$s_{N,Y_j} = \left[\left(\frac{1}{N-1}\right) \sum_{n=1}^{N} \left(y_{j,n} - m_{N,Y_j}\right)^2\right]^{\frac{1}{2}}.$$

As a consequence of the replacement of σ_{Y_j} by s_{N,Y_j}, an approximate v% confidence interval for the mean value of Y_j is given by

$$\left(m_{N,Y_j} - b\right) \text{ and } \left(m_{N,Y_j} + b\right) \text{ with}$$
$$b = t_{(N-1),(100+v)/2}\, s_{N,Y_j} N^{-1/2}$$

$t_{(N-1),(100+v)/2}$ the $[(100 + v)/2]$% quantile of the Student or t-distribution with $(N - 1)$ degrees of freedom [see Eqs. (3.4) to (3.6)].

The mean value of the subjective probability distribution of the model result Y_j is the so-called best estimate (see subsection 4 of Sect. 3.6.1). It is generally different from the model result value obtained with the mean values ("best estimates") of the subjective probability distributions for the uncertain data. The rare case where the model result is a linear function of the latter is an exception.

The mean value is not suitable as an uncertainty measure for Y_j. However, mean values of the relative frequencies of sample elements above, below or between specified values of Y_j are suitable as uncertainty measures. For this reason, the following sections deal with the estimation of quantile values of the subjective probability distribution of the model result.

5.2.3 A Quantile Value of the Model Result

The $u\%$ quantile $y_{j,\,u\%}$ of the subjective probability distribution of the model result Y_j is a quantitative expression of the uncertainty of Y_j (see also the interpretation under a) in the introduction to this chapter). The model result does not exceed $y_{j,\,u\%}$ with subjective probability $u/100$, and it exceeds $y_{j,\,u\%}$with subjective probability $1 - u/100$.

To obtain an estimate of the $u\%$ quantile, the values in the simple random sample $(y_{j,\,1}, y_{j,\,2}, \ldots, y_{j,\,N})$ are ordered by increasing magnitude and the i-th ordered value is used as an estimate. The index i is the smallest integer value larger than or equal to $Nu/100$. Confidence statements are required for estimates of distribution quantiles if the sample size N is relatively small and particularily if u is large (90, 95, etc.) or small (10, 5, etc.). These quantiles are in the tail ends of the distributions where sample values tend to be spread out over wide ranges and estimates exhibit considerable variability from sample to sample of size N. Confidence statements are given in the form of $(u\%, v\%)$ tolerance confidence limits (see Sect. 5.2.1). They are known in the literature (Wilks 1941; Wald 1943; Conover 1980; Guenther 1985) as one-sided (lower or upper) or two-sided $(u\%, v\%)$ tolerance confidence or statistical tolerance limits and intervals.

Tolerance confidence limits are obtained from the order statistics of the sample variables. The order statistic of a given rank is the sample variable with the value of that rank in the ordered sample. The order statistic of rank 1 is the sample variable with the smallest, and the order statistic of rank N is the sample variable with the largest value in the sample of size N. The probability that at most $m_u - 1$ sample values lie above $y_{j,\,u\%}$ is given by the binomial distribution expression:

$$\sum_{j=0}^{m_u-1} \frac{N!}{j!(N-j)!}\left(1-\frac{u}{100}\right)^{j}\left(\frac{u}{100}\right)^{N-j}.$$

Consequently, if m_u satisfies

$$\sum_{j=0}^{m_u-1} \frac{N!}{j!(N-j)!}\left(1-\frac{u}{100}\right)^{j}\left(\frac{u}{100}\right)^{N-j} \le 1 - v/100 \tag{5.1}$$

then the probability that at least m_u sample values are larger than $y_{j,\,u\%}$ is given by

$$1 - \sum_{j=0}^{m_u-1} \frac{N!}{j!(N-j)!}\left(1-\frac{u}{100}\right)^{j}\left(\frac{u}{100}\right)^{N-j} \ge v/100. \tag{5.2}$$

From (5.2), one concludes that the probability to have m_u or more sample values above $y_{j,\,u\%}$ is at least v/100. Consequently, the smallest of the m_u sample values from the top, i.e. the $(N + 1 - m_u)$-th order statistic, is an upper $(u\%, v\%)$ tolerance confidence limit saying: "The value of the model result Y_j lies below the

($N + 1 - m_u$)-th order statistic at least with subjective probability $u/100$ at a confidence level of at least $v\%$."

The binomial distribution is discrete, and therefore, it is unlikely to have equality in (5.1) and (5.2).

Two situations are of interest:

1. N, u and v are given and

   (a) m_u is to be determined such that the ($N + 1 - m_u$)-th ordered sample value (i.e. the m_u-th value from the top of the ordered sample) is an upper ($u\%$, $v\%$) statistical tolerance limit

   or

   (b) m_l is to be determined such that the m_l-th value from the bottom of the ordered sample is a lower ($u\%$, $v\%$) statistical tolerance limit.

The solutions are found as follows:

(1a) Find the largest integer m_u that satisfies inequality (5.1). To simplify matters, the following approximate relationship (Scheffé and Tukey 1944) may be used:

$$N \approx 0.25\chi^2_{(2m_u, v\%)}\left(1 + \frac{u}{100}\right)/\left(1 - \frac{u}{100}\right) + 0.5(m_u - 1) \tag{5.3}$$

$\chi^2_{(2m_u, v\%)}$ is the $v\%$ quantile of the χ^2 distribution with degree of freedom $2m_u$ (Table III in Winkler and Hays (1975)). The largest value obtained for m_u, for which the right-hand side of (5.3) does not exceed N, may be used in (5.1) to check whether it is the largest integer satisfying the inequality. The m_u-th value from the top of the ordered sample is then an upper ($u\%$, $v\%$) statistical tolerance limit.

(1b) Find the largest integer m_l that satisfies

$$\sum_{j=0}^{m_l - 1} \frac{N!}{j!(N-j)!}\left(1 - \frac{u}{100}\right)^j \left(\frac{u}{100}\right)^{N-j} \leq 1 - v/100. \tag{5.4}$$

The largest value found for m_l, such that in

$$N \approx 0.25\chi^2_{(2m_l, v\%)}\left(1 + \frac{u}{100}\right)/\left(1 - \frac{u}{100}\right) + 0.5(m_l - 1) \tag{5.5}$$

the right-hand side does not exceed N, may be used in (5.4) to check whether it is the largest integer satisfying the inequality. The m_l-th value from the bottom of the ordered sample is then a lower ($u\%$, $v\%$) statistical tolerance limit.

If N, u and v are the same for (a) and (b), then $m_u = m_l$ and the difference lies only in the rank of the order statistic that is to be used as limit. There may be no solution in situation (1) if N is not large enough. This leads to the second situation:

2. u, v and m_l or m_u are given and the smallest sample size N is to be determined such that in the ordered sample

(a) the m_u-th value from the top is an upper (u%, v%) statistical tolerance limit
or
(b) the m_l-th value from the bottom is a lower (u%, v%) statistical tolerance limit.

The solutions are found as follows:

(2a) Use the approximate formula (5.3) and find the smallest value for N that is larger or equal to the right-hand side of (5.3) and check with the inequality (5.1).
(2b) Use the approximate formula (5.5) and find the smallest value for N that is larger or equal to the right-hand side of (5.5) and check with the inequality (5.4).

Clearly, if u and v are the same and $m_l = m_u$, then N is the same for (2a) and (2b).

For instance, if the sample is of size $N = 59$, then the maximum model result value in this sample ($m_u = 1$) is already a one-sided upper (95%, 95%) statistical tolerance limit [see Table A5 in Conover (1980)]. It says: At a confidence level of at least 95% the model result value lies below this limit at least with subjective probability 0.95.

$N = 45$ is the minimum sample size for a one-sided (95%, 90%) statistical tolerance limit (i.e. at a confidence level of at least 90%). The largest sample value of the model result is then a one-sided upper (95%, 90%) statistical tolerance limit, i.e. at a confidence level of at least 90% the model result lies at least with subjective probability 95/100 below this limit. The minimum required sample sizes are obtained as the smallest integer number N such that for $m_u = 1$, the inequality

$$\left(\frac{u}{100}\right)^N \leq 1 - v/100 \tag{5.6}$$

is just satisfied (u is the quantile percentage and v the confidence level percentage). The term $(u/100)^N$ is the probability that none of the N model result values exceed the u% quantile of the unknown subjective probability distribution. This probability should not exceed $1 - v/100$. The probability that at least one of the N model result values exceeds the u% quantile is then $\geq v/100$. The largest value ($m_u = 1$) will always be among those that do exceed the u% quantile.

The required sample size N is obviously independent of the number of uncertain data considered in the Monte Carlo simulation and depends only on the values chosen for the two percentages u and v.

If one can afford more than the minimum sample size, then one may be able to use a value below the maximum value in the sample as upper statistical tolerance limit. The approximate formula (5.3) tells which of the ordered values in the sample can then serve as the one-sided upper (u%, v%) statistical tolerance limit.

N is the given sample size and the largest integer m_u is to be determined from (5.3) such that the right-hand side is still smaller or equal to N. The m_u-th (from the top) ordered (by increasing magnitude) element in the sample of model result values is then an upper (u%, v%) statistical tolerance limit. For instance, if $N = 100$ and $u = v = 95$, then $m_u = 2$, i.e. the second largest value in the sample is an upper (95%,

95%) statistical tolerance limit. The subjective probability for model result values to lie below this value is at least 0.95 at a confidence level of at least 95%. At least 95% of the population of possibly true values of the model result lie below this value at a confidence level of at least 95% (given the assumptions mentioned in the introduction to this chapter hold).

5.2.4 A Subjective Probability Interval for the Model Result

Again, two situations are of interest:

1. N, u and v are given and
   m_l and m_u are to be determined such that the m_l-th value from the bottom and the m_u-th value from the top of the ordered sample delimit a two-sided (u%, v%) statistical tolerance interval.
   or
2. u, v, m_l and m_u are given and the smallest sample size N is to be determined such that in the ordered sample the m_l-th value from the bottom and the m_u-th value from the top delimit a two-sided (u%, v%) statistical tolerance interval.

The solutions are found as follows:

1. Find the largest integer m that satisfies

$$\sum_{j=0}^{m-1} \frac{N!}{j!(N-j)!} \left(1 - \frac{u}{100}\right)^j \left(\frac{u}{100}\right)^{N-j} \leq 1 - v/100. \tag{5.7}$$

The approximate relationship

$$N \approx 0.25\chi^2_{(2m,v\%)} \left(1 + \frac{u}{100}\right) / \left(1 - \frac{u}{100}\right) + 0.5(m-1) \tag{5.8}$$

may be used ($\chi^2_{(2m,v\%)}$ is the v% quantile of the χ^2 distribution with degree of freedom $2m$) to find the largest integer m for which the right-hand side of (5.8) does not exceed N. This value may then be checked whether it is the largest integer satisfying the inequality (5.7). Any pair of values (m_l, m_u) with $m_l + m_u = m$ has the following property: The m_l-th value from the bottom and the m_u-th value from the top of the ordered sample delimit a two-sided (u%, v%) statistical tolerance interval. The difference between inequalities (5.1), (5.4) and (5.7) lies only in the ranks of the order statistics that are to be used as the statistical tolerance limit.There may be no solution in situation (1) if N is not large enough. This leads to the second situation with the following solution:

2. Use the approximate formula (5.8) with $m = m_l + m_u$ and find the smallest value for N that is larger or equal to the right-hand side of (5.8) and check with the inequality (5.7).

Inequality (5.7) says that with probability of at least v/100 at least m of the N sample values are outside an interval that is not explicitly specified and that contains Y_j with probability u/100. The order statistics of rank m_l and of rank (N + 1 − m_u) with $m_l + m_u = m$ delimit an interval such that, at a confidence level of at least v%, Y_j lies in the interval with at least subjective probability u/100. Different to (5.1) and (5.4), this interval is only implicitly specified, namely through the condition $m_l + m_u = m$.

Uncertainty is often to be quantified in the form of an interval that contains the model result value at least with subjective probability $u = 0.9$, for instance. The smallest and the largest value in a sample of size $N = 46$ [see Table A6 in Conover (1980)] enclose a value range of at least 0.9 subjective probability at a confidence level of at least 95%. This interval is called a two-sided (90%, 95%) statistical tolerance limit. It contains at least 90% of the population of possibly true model result values at a confidence level of at least 95% (given the assumptions in the introduction to this chapter hold).

If a sample size larger than $N = 46$ can be afforded, then the approximate relationship

$$N \approx 0.25\chi^2_{(2(m_l+m_u),v\%)}\left(1+\frac{u}{100}\right)/\left(1-\frac{u}{100}\right)+0.5(m_l+m_u-1)$$

where N is the given sample size and $\chi^2_{(2(m_l+m_u),v\%)}$ is the v% quantile of the χ^2 distribution with degree of freedom $2(m_l + m_u)$ [see Table A2 in Conover (1980)] can be used to determine sample values such that they define an approximate two-sided (u%, v%) statistical tolerance limit. The integer numbers m_l and m_u are to be determined such that the right-hand side of the above approximate relationship is just smaller or equal to N. The m_l-th value from the bottom and m_u-th value from the top of the ordered (by increasing magnitude) sample determine an approximate (u%, v%) statistical tolerance interval. The subjective probability for the model result to lie within this interval is at least u/100 at a confidence level of at least v%. This interval is also called a two-sided (u%, v%) statistical tolerance limit. It contains at least u% of the population of possibly true values of the model result at a confidence level of at least v% (given the assumptions in the introduction to this chapter hold).

5.2.5 *Compliance of the Model Result with a Limit Value*

Let y_{lim} be the limit value, Y_j the model result to be compared to y_{lim}, N the sample size (i.e. the total number of replications of the computer model application), m the number of model result values $y_{j,\,n}$, $n = 1, \ldots, N$, in the sample with $y_{j,\,n} > (<)\, y_{lim}$, v % the confidence level and

$a = F_{v,k1,k2}$ the v% quantile of the F-distribution with $k_1 = 2(m + 1)$, $k_2 = 2(N - m)$ degrees of freedom [tabulated, for instance, in Winkler and Hays (1975)]

$b = F_{v,k1,k2}$ the v% quantile of the F-distribution with $k_1 = 2(N–m + 1)$, $k_2 = 2m$ degrees of freedom

$c = F_{(100 + v)/2,k1,k2}$ the $(100 + v)/2$% quantile of the F-distribution with $k_1 = 2(N – m + 1)$, $k_2 = 2m$ degrees of freedom

$d = F_{(100 + v)/2,k1,k2}$ the $(100 + v)/2$% quantile of the F-distribution with $k_1 = 2(m + 1)$, $k_2 = 2(N – m)$ degrees of freedom

then (Heinhold 1968):

– a v% confidence interval for the subjective probability $sw(Y_j >(<) y_{lim})$ is given as:

$$\frac{m}{[m + (N - m + 1)c]} \leq sw\left(Y_j > (<)y_{lim}\right) \leq \frac{(m + 1)d}{[(m + 1)d + N - m]};$$

– an upper confidence limit for $sw(Y_j >(<) y_{lim})$ is given as:

$$sw\left(Y_j > (<)y_{lim}\right) \leq \frac{(m + 1)a}{[(m + 1)a + N - m]};$$

– a lower confidence limit for $sw(Y_j >(<) y_{lim})$ is given as:

$$\frac{m}{[m + (N - m + 1)b]} \leq sw\left(Y_j > (<)y_{lim}\right).$$

5.2.6 *The Sample Variability of Statistical Tolerance Limits*

How variable are statistical tolerance limits? In other words: Would replicated sampling with equal sample size N provide widely differing statistical tolerance limits? The maximum value of a simple random sample of size N is distributed according to $[F_Y(y)]^N$ since none of the sample values exceed y. In the context of uncertainty analysis, F_Y denotes the unknown subjective probability distribution for the model result Y. The minimum sample size for an upper (95%, 95%) statistical tolerance limit is 59 as can be seen from (5.6) or from Table A5 in Conover (1980). The statistical tolerance limit is the maximum model result value in the sample of size $N = 59$. If many samples of size $N = 59$ are drawn according to a continuous distribution F_Y, roughly 45% of the maximum values will lie between the 95.05% and the 98.85% quantile of the subjective probability distribution for the model result and roughly another 45% between the 98.85% and the 99.92% quantile. If the

subjective probability distribution for the model result has a long tail, extending to large values, the latter range may be quite wide, implying considerable variability of the statistical tolerance limits and therefore possibly considerable conservatism if the statistical tolerance limit is compared to an upper safety limit. This is the price to be paid for trying to infer relatively extreme quantiles from a small sample drawn according to a distribution that is only indirectly specified. If the matter is important enough to justify some effort in reducing the chance for a considerably conservative compliance check, then it will be important enough to justify a larger sample size N. If the sample size is 100, for instance, the 99th ordered value is an upper (95%, 95%) statistical tolerance limit, as can be seen from (5.2) with $m_u = 2$. The 99th order statistic is distributed according to

$$(\mathrm{F_Y}(y))^{100} + 100(\mathrm{F_Y}(y))^{99}(1 - \mathrm{F_Y}(y)) \tag{5.9}$$

The expression (5.9) gives the probability that either all $N = 100$ model results do not exceed y or that at most one result does exceed y. Consequently,

$$1 - \left\{(\mathrm{F_Y}(y))^{100} + 100(\mathrm{F_Y}(y))^{99}(1 - \mathrm{F_Y}(y))\right\}$$

is the probability that at least two model results in the sample of size $N = 100$ do exceed the value y. Among those two results is the 99th ordered sample value (i.e. the second largest value).

Roughly 45% of the 99th ordered values from many simple random samples of size $N = 100$ will lie between the 95.05% and the 98.35% quantile since (5.9) says that the two top ranked model results are with probability of about 0.95 above the 95.05% quantile of the model results and with probability of about 0.493 above the 98.35% quantile. Roughly another 45% of the 99th ordered values are between the 98.35% and the 99.65% quantile, while they are above the 99.65% quantile with probability of about 0.05. Therefore, this limit exhibits already less variability if compared to the statistical tolerance limit obtained from samples of size $N = 59$.

5.2.7 *Comparison of Two Model Results*

The comparison can only be probabilistic, i.e. the subjective probability is determined for model result Y_j to be larger (smaller) than model result Y_i, $j \neq i$, j, $i \in (1, \ldots, J)$. An estimate of this subjective probability is available from the number of pairs of values $(y_{j,\,n}, y_{i,\,n})$ in the sample of size N with $y_{j,\,n} > (<)\, y_{i,\,n}$, i.e.

$$\mathrm{sw_{est}}\left(Y_j > (<) Y_i\right) = N^{-1}m$$

$$\text{with} \quad m = \sum_{n=1}^{N} m_n$$

and

$$m_n = 1 \text{ if } y_{j,n} > (<)\ y_{i,n} \text{ and 0 otherwise.}$$

A v% confidence interval for $\text{sw}(Y_j > Y_i)$ is obtained as (Heinhold and Gaede 1968):

$$\frac{m}{[m + (N - m + 1)c]} \leq \text{sw}\big(Y_j > (<)Y_i\big) \leq \frac{(m+1)d}{[(m+1)d + N - m]}$$

c is the $(100 + v)/2$% quantile of the F-distribution (tabulated in, for instance, Winkler and Hays 1975) with degrees of freedom $2(N - m + 1)$ and $2m$ while d is the $(100 + v)/2$% quantile of the F-distribution with degrees of freedom $2(m + 1)$ and $2(N - m)$.

5.2.8 Comparison of More than Two Model Results

If $L > 2$ model results are to be compared, the following approach may be useful:

For $l = 1, \ldots, L$ determine how often (out of the N model runs) model result Y_l was the largest (smallest) among the L results and set

$$g_l = n_{l,1}$$

$n_{l,\,1}$ is the number of model runs with Y_l largest (smallest) among the L results. Then set

$$g_l = g_l + 2n_{l,2}$$

$n_{l,\,2}$ is the number of runs with Y_l second largest (smallest), and so forth. Finally,

$$g_l = N^{-1} \sum_{j=1}^{L} j n_{l,j}$$

is compared for $l = 1, \ldots, L$ and the model results are ranked according to g_l for $l = 1, \ldots, L$, i.e. the result Y_l with the smallest g_l receives the top rank. Since the g_l are estimates of mean values, an approximate 90% confidence interval can be derived as follows from random samples (SRS) of size $N > 30$:

lower endpoint: $g_l - b$
upper endpoint: $g_l + b$
with

$$b = t_{(N-1),(100+v)/2}\, s/N^{1/2}$$

$t_{(N-1),(100+v)/2}$ is the [(100+v)/2]% quantile of the Student or t-distribution with $(N-1)$ degrees of freedom [see Eqs. (3.4) to (3.6)] and

$$s = \left[\left(\frac{1}{N-1}\right)\sum_{n=1}^{N}(j_n - g_l)^2\right]^{1/2}.$$

The value of j_n is either 1, 2, ..., L depending on whether the n-th sample value for Y_l is the largest (smallest), second largest (smallest) and so forth among the n-th set of sample values of the L model results.

5.3 Uncertainty Statements Available from Uncertainty Propagation Using LHS

An uncertainty analysis, using Monte Carlo simulation with LHS, provides in Step 3 an input sample that is partly deterministic (see Sect. 4.4.2). The M-dimensional unit hypercube $\mathbf{I}^M$ is divided into N^M M-dimensional elemental hypercubes or cells $C_{n1,n2,\ldots,nM}$, $n_m = 1, \ldots, N$, $m = 1, \ldots, M$ each of probability content $1/N^M$. Once the first input sample element $\boldsymbol{e_1} = (p_{1,1(1)}, p_{2,1(2)}, \ldots, p_{M,1(M)})$ has been drawn, where $p_{m,1(m)} = F_m^{-1}(u_{m,1(m)})$, $m = 1, \ldots, M$, and the numbers 1(1), 1(2), ..., 1(M) are the first elements in M random permutations of the numbers 1, 2, ..., N, the next sample element can only be drawn from $(N-1)^M$ of the elemental hypercubes namely from all of those that do not have any of the numbers 1(m), $m = 1, \ldots, M$, as their m-th index (and so forth). The input sample variables $\boldsymbol{E_1}, \ldots, \boldsymbol{E_N}$ are therefore dependent but identically distributed.

Each input sample element $\boldsymbol{e_n} = (p_{1,n(1)}, p_{2,n(2)}, \ldots, p_{M,n(M)})$, $n = 1, ..., N$, is converted into the corresponding output sample element $\boldsymbol{a_n} = (y_{1,n}, y_{2,n}, \ldots, y_{J,n})$ by the computer model. While the first output sample element $\boldsymbol{a_1}$ is drawn according to the mixture of the conditional joint distributions of the model results over each of the N^M elemental hypercubes, the second sample element $\boldsymbol{a_2}$ is drawn according to the mixture over only $(N-1)^M$ elemental hypercubes (they depend on the elemental hypercube of the first sample element), the third sample element $\boldsymbol{a_3}$ is drawn according to the mixture over each of $(N-2)^M$ elemental hypercubes (they depend on the elemental hypercubes of the first and the second sample element) and so forth. It follows that the output sample variables $\boldsymbol{A_1}, \ldots, \boldsymbol{A_N}$ are not independent but identically distributed according to the unknown joint subjective probability distribution of the model results.

A sample drawn according to the unknown subjective probability distribution of any individual model result Y_j is therefore immediately available as $(y_{j,1}, y_{j,2}, \ldots, y_{j,N})$. The value of the density function is given as (Table 5.2)

Table 5.2 Output of the uncertainty propagation by Monte Carlo simulation using LHS

Computer model run no.	Input sample element no.	Uncertainty no.				Output sample element no.	Model result no.			
		1	2	...	M		1	2	...	J
		Sample values					Sample			
1	1	$p_{1,1(1)}$	$p_{2,1(2)}$	...	$p_{M,1(M)}$	1	$y_{1,1}$	$y_{2,1}$	...	$y_{J,1}$
2	2	$p_{1,2(1)}$	$p_{2,2(2)}$	...	$p_{M,2(M)}$	2	$y_{1,2}$	$y_{2,2}$	...	$y_{J,2}$
.	.			.		.			.	
.	.			.		.			.	
.	.			.		.			.	
N	N	$p_{1,N(1)}$	$p_{2,N(2)}$	...	$p_{M,N(M)}$	N	$y_{1,N}$	$y_{2,N}$	...	$y_{J,N}$

$$f_j(y_{j,1}, y_{j,2}, \ldots, y_{j,N}) = f_j(y_{j,1}) f_{j|1}(y_{j,2}) \ldots f_{j|1,\ldots,N-1}(y_{j,N}). \tag{5.10}$$

The indices j|1, j|1,2 and so forth denote the mixture density function under the condition of the elemental hypercube of the first, the first and the second sample element and so forth. In the simple random sampling case

$$f_j(y_{j,1}, y_{j,2}, \ldots, y_{j,N}) = \prod_{n=1}^{N} f_j(y_{j,n}).$$

5.3.1 Estimates of Mean Values of Functions of the Model Result

The mean value of a sum of dependent variables equals the sum of the mean values of the individual variables. The sample can therefore be used to estimate mean values. Mean values of interest are those of functions $g(Y_j)$, such as

- The mean value of Y_j (with $g(Y_j) = Y_j$)
- The variance as the difference of the mean value of ${Y_j}^2$ and the square of the mean value of Y_j (with $g(Y_j) = {Y_j}^2$)
- The subjective probability of violation of a limit value ($g(Y_j) = 1$ if $Y_j > (<) y_{lim}$ and is 0 otherwise).

However, all of these estimates, due to the dependence of the sample values, cannot be supplemented by confidence intervals and confidence limits.

The variances of the estimates of mean values of functions $g(Y_j)$ are those of $M_{g(Y_j),SRS} = N^{-1} \sum_{n=1}^{N} g(Y_{j,n})$ in the case of simple random sampling and $M_{g(Y_j),LHS} = N^{-1} \sum_{i=1}^{f(N)} W_i g(Y_{j,i})$ in the case of Latin Hypercube sampling.

f(N) = N^M the number of elemental hypercubes or cells and $W_i = 1$ if the cell number i has been chosen by the Latin Hypercube sampling and $W_i = 0$ otherwise. If all uncertain data $P_1, \ldots, P_M$ are state of knowledge independent, then the variances relate to each other as shown in McKay et al. (1979).

Sections 5.3.2 to 5.3.5 derive estimates of mean value, quantile values, intervals of given subjective probability content as well as of the subjective probability of violation of a limit value from sample values of Y_j obtained with a Latin Hypercube sample of the uncertain data. Section 5.3.8 discusses the use of K replicated Latin Hypercube samples to obtain estimates with confidence statements. Sections 5.3.6 and 5.3.7 show how to compare different model results on the basis of their random output sample values. Finally, ways of graphical presentation of uncertainty estimates are discussed in Sect. 5.4.2.

5.3.2 The Mean Value of the Model Result

The mean value of the subjective probability distribution of a model result Y_j can be estimated from a sample of model output values $y_{j,\,n}$ obtained using an LHS of size N of the uncertain data, as

$$m_{Y_j} = \left(\frac{1}{N}\right) \sum_{n=1}^{N} y_{j,n}.$$

The vector $(y_{j,\,1}, y_{j,\,2}, \ldots, y_{j,\,N})$ of sample values is a realization of the random vector $(Y_{j,\,1}, Y_{j,\,2}, \ldots, Y_{j,\,N})$. Each component of the random vector has the subjective probability distribution of the model result Y_j. The sample variables $(Y_{j,\,1}, Y_{j,\,2}, \ldots, Y_{j,\,N})$ are, however, not independent, and the central limit theorem can therefore not be applied. Consequently, the distribution of M_{Y_j} cannot be approximated by a normal distribution. Confidence limits and intervals for the mean value of Y_j are therefore not available. See Sect. 5.3.8 for an attempt to circumvent this handicap.

5.3.3 A Quantile Value

The u% quantile $y_{j,\,u\%}$ of the subjective probability distribution of the model result Y_j is a quantitative expression of the uncertainty of Y_j. The model result does not exceed $y_{j,\,u\%}$ with subjective probability u/100, and it exceeds $y_{j,\,u\%}$ with subjective probability 1 – u/100.

To obtain an estimate of the u% quantile, the values in the model output sample $(y_{j,\,1}, y_{j,\,2}, \ldots, y_{j,\,N})$ are ordered by increasing magnitude and the i-th ordered value is used as an estimate. The index i is the smallest integer value larger than or equal to Nu/100. Confidence statements are required for estimates of distribution quantiles if

the sample size N is relatively small and particularily if u is large (90, 95, etc.) or small (10, 5, etc.). These quantiles are in the tail ends of the distributions where sample values tend to be spread out over wide ranges and estimates exhibit considerable variability from sample to sample of size N. While confidence statements are given in the form of (u%, v%) tolerance confidence limits (see Sect. 5.2.1) in the case of simple random sampling (SRS), they are not available from the model output sample obtained with an LHS of the uncertain data due to the dependence of the sample variables. See Sect. 5.3.8 for an attempt to circumvent this handicap.

5.3.4 A Subjective Probability Interval

An estimate of an interval containing q% subjective probability of Y_j is delimited by two quantile value estimates $\widehat{y}_{j,u\%}$ and $\widehat{y}_{j,w\%}$ such that $(w - u) = q$. Confidence statements are not available for this interval estimate since the sample variables are not independent. For the same reason, a two-sided (q%,v%) statistical tolerance limit is also not available. See Sect. 5.3.8 for an attempt to circumvent this handicap.

5.3.5 Compliance with a Limit Value

Let y_{lim} be an upper limit value, Y_j the model result to be compared to y_{lim}, N the sample size and i the number of model output values $y_{j,n}$ in the sample with $y_{j,\,n} \leq y_{lim}$. The mean value estimate

$$m_{\left(Y_j \leq y_{lim}\right)} = \frac{i}{N} \quad \text{with } i = \sum_{n=1}^{N} \mathbf{I}_{\left(Y_j \leq y_{lim}\right)}\left(y_{j,n}\right)$$

$$\mathbf{I}_{\left(Y_j \leq y_{lim}\right)}\left(y_{j,n}\right) = 1 \text{ if } y_{j,n} \leq y_{lim} \text{ and } \mathbf{I}_{\left(Y_j \leq y_{lim}\right)}\left(y_{j,n}\right) = 0 \text{ otherwise}$$

is an estimate of the subjective probability that the model result Y_j does not violate the limit value y_{lim}. Again, confidence intervals and limits, as in the case of simple random sampling (SRS) (see Sect. 5.2.5), are not available since the sample variables are not independent. See, however, Sect. 5.3.8 for an attempt to circumvent this handicap.

5.3.6 Comparison of Two Model Results

The comparison can only be probabilistic, i.e. the subjective probability is determined for model result Y_j to be larger (smaller) than model result Y_i, $j \neq i$, j, $i \in$

$(1, \ldots, J)$. An estimate of this subjective probability is given by the number of pairs of values $(y_{j,\,n}, y_{i,\,n})$ in the sample of size N with $y_{j,\,n} > (<)\, y_{i,\,n}$, i.e.

$$\mathrm{sw}_{\mathrm{est}}\left(Y_j > (<) Y_i\right) = N^{-1} m$$

$$\text{with } m = \sum_{n=1}^{N} m_n, m_n = 1 \text{ if } y_{j,n} > (<)\, y_{i,n} \text{ and 0 otherwise.}$$

Confidence intervals for $\mathrm{sw}(Y_j > (<) Y_i)$ are not available unless K replicated LH samples of size N/K are used (see Sect. 5.3.8) instead of only one sample of size N.

5.3.7 Comparison of More than Two Model Results

If $L > 2$ model results are to be compared, the following approach may be useful:

For $l = 1, \ldots, L$ determine how often (out of the N model runs) model result Y_l was the largest (smallest) among the L results and set

$$g_l = n_{l,1}$$

$n_{l,\,1}$ is the number of model runs with Y_l largest (smallest) among the L results. Then set

$$g_l = g_l + 2n_{l,2}$$

$n_{l,\,2}$ is the number of runs with Y_l second largest (smallest), and so forth. Finally, the estimated mean ranks

$$g_l = N^{-1} \sum_{j=1}^{L} j n_{l,j}$$

are compared for $l = 1,\ldots,L$ and the model results are ordered accordingly.

5.3.8 Estimates from Replicated Latin Hypercube Samples

The model output sample variables $(Y_{j,\,1}, Y_{j,\,2}, \ldots, Y_{j,\,N})$ are identically distributed but are not independent. This fact precludes the computation of confidence limits and intervals for estimates using model output values from a Latin Hypercube sample (LH sample) of the uncertain data. In order to circumvent this handicap, it was suggested (McKay et al. 1979; Hansen et al. 2012) to use K independently drawn LH samples of size N/K instead of one sample of size N. The K estimate variables are independent and identically distributed, and their sample values can therefore be used to obtain approximate confidence limits and intervals relying on the central limit theorem. Condition is, however, that the LH samples are drawn at

random from within the equal subjective probability intervals of the uncertain data, as opposed to using the conditional mean or median values. The approximation is considered satisfactory for $K \geq 30$.

– The mean value of Y_j
  The mean value estimates $m_{Y_{j,k}}$, $k = 1, \ldots, K$, from K replicated LH samples of size N/K are used as follows:

$$e_{Y_j} = K^{-1} \sum_{k=1}^{K} m_{Y_{j,k}} \quad \text{with} \quad m_{Y_{j,k}} = KN^{-1} \sum_{i=1}^{\frac{N}{K}} y_{j,k,i}.$$

  The vector of mean value estimates $\left(m_{Y_{j,1}}, m_{Y_{j,2}}, \ldots, m_{Y_{j,K}}\right)$ is a realization of the random vector $\left(M_{Y_{j,1}}, M_{Y_{j,2}}, \ldots, M_{Y_{j,K}}\right)$. The components of this vector are independent and identically distributed with mean value $\mathrm{E}\{Y_j\}$. Consequently, an approximate v% confidence interval for the mean value of Y_j is obtainable as

$$\left(e_{Y_j} - b, e_{Y_j} + b\right)$$

$$b = t_{(K-1),(100+v)/2}\, s_{K,Y_j} K^{-1/2} \text{ and } s_{K,Y_j} = \left[\left(\frac{1}{K-1}\right) \sum_{k=1}^{K} \left(m_{Y_{j,k}} - \ e_{Y_j}\right)^2\right]^{1/2}$$

  with $t_{(K\ -\ 1),\ (100\ +\ v)/2}$ the [(100 + v)/2]% quantile of the Student or t-distribution with (K – 1) degrees of freedom [see Eqs. (3.4) to (3.6)]. The approximation is considered satisfactory for K ≥ 30.

– A quantile value of Y_j
  The quantile estimates $\widehat{y}_{j,u\%,k}$, $k = 1, \ldots, K$, from K replicated LH samples of size N/K are used as follows:

$$e_{Y_j,u\%} = K^{-1} \sum_{k=1}^{K} \widehat{y}_{j,u\%,k} \text{ with } \widehat{y}_{j,u\%,k} = y_{j,k,i} \text{ and } i \text{ is the smallest integer larger}$$

  than or equal to $NK^{-1}u/100$.

  The vector of quantile estimates $\left(\widehat{y}_{j,u\%,1}, \widehat{y}_{j,u\%,2}, \ldots, \widehat{y}_{j,u\%,K}\right)$ is a realization of a random vector, the components of which are independent and identically distributed with mean value larger than $y_{j,\ u\%}$. This may be avoided by taking for every second index k the $y_{j,\ k,\ i}$ with i the largest integer smaller than or equal to $NK^{-1}u/100$.

  An approximate v% confidence interval for the quantile value $y_{j,\ u\%}$ is obtained as

$$\left(e_{Y_j,u\%} - b, e_{Y_j,u\%} + b\right)$$

$$b = t_{(K-1),(100+v)/2}\, s_{K,Y_j,u\%} K^{-1/2} \quad \text{and} \quad s_{K,Y_j,u\%}$$

$$= \left[\left(\frac{1}{K-1}\right) \sum_{k=1}^{K} \left(\widehat{y}_{j,u\%,k} - e_{Y_j,u\%}\right)^2\right]^{1/2}$$

with $t_{(K-1),(100+v)/2}$ the [(100 + *v*)/2]% quantile of the Student or t-distribution with (*K*-1) degrees of freedom [see Eqs. (3.4) to (3.6)] and $K \geq 30$.

– A subjective probability interval of Y_j
From quantile value estimates $e_{Y_j,u\%}$ and $e_{Y_j,w\%}$, obtained as shown above with $(w - u) = q$, the interval

$$e_{Y_j,u\%} - t_{(K-1),\frac{100+v}{2}}\ s_{K,Y_j,u\%}\ K^{-\frac{1}{2}} \quad \text{and} \quad e_{Y_j,w\%} + t_{(K-1),\frac{100+v}{2}}\ s_{K,Y_j,w\%}\ K^{-\frac{1}{2}}$$

is obtained. It delimits an approximate *v*% confidence interval for a *q*% subjective probability interval of Y_j.

– Compliance with a limit value
Let y_{lim} be the limit value, Y_j the model result to be compared to y_{lim}, *K* the number of sample replicates, *N*/*K* their sample size and i_k, $k = 1, .., K$, the numbers of the model output values in each sample with $y_j \leq y_{lim}$. The mean value estimate

$$e_{\left(Y_j \leq y_{lim}\right)} = K^{-1} \sum_{k=1}^{K} m_{\left(Y_j \leq y_{lim}\right),k}$$
$$m_{\left(Y_j \leq y_{lim}\right),k} = KN^{-1} i_k$$

is an estimate of the subjective probability that the model result Y_j does not violate the limit value y_{lim}.

The vector of mean value estimates $\left(m_{\left(Y_j \leq y_{lim}\right),1}, m_{\left(Y_j \leq y_{lim}\right),2}, \ldots, m_{\left(Y_j \leq y_{lim}\right),K}\right)$ is a realization of a random vector. The components of this random vector are independent and identically distributed with the subjective probability of compliance as mean value. Consequently, given $K \geq 30$, an approximate *v*% confidence interval for this probability is available using the central limit theorem

$$\left(e_{\left(Y_j \leq y_{lim}\right)} - b, e_{\left(Y_j \leq y_{lim}\right)} + b\right)$$
$$b = t_{(K-1),(100+v)/2}\ s_{K,\left(Y_j \leq y_{\lim}\right)} K^{-1/2}$$
$$\text{and} \quad s_{K,\left(Y_j \leq y_{\lim}\right)} = \left[\left(\frac{1}{K-1}\right) \sum_{k=1}^{K} \left(m_{\left(Y_j \leq y_{lim}\right),k} - e_{\left(Y_j \leq y_{lim}\right)}\right)^2\right]^{\frac{1}{2}}$$

with $t_{(K-1),(100+v)/2}$ the [(100 + v)/2]% quantile of the Student or t-distribution with $(K - 1)$ degrees of freedom [see Eqs. (3.4) to (3.6)].

5.4 Graphical Presentation of Uncertainty Analysis Results

Computer model results may be scalar values, such as table entries, matrix elements, vector components and so forth, or functions of some independent variables like time and/or space. The uncertainty information discussed above together with information on state of knowledge dependence between model results is contained in the sample from the unknown joint subjective probability distribution of the model results. Various ways of presenting this information are useful for differing reasons.

5.4.1 Graphical Presentation of Uncertainty Analysis Results Obtained Using SRS

Subsection 1 discusses the graphical presentation of scalar model results. The various forms of illustration shown in this subsection are suitable for individual table entries, matrix elements, vector components and even for function values at selected points in time and/or space. Forms of graphical presentation of analysis results for function values over a sequence of argument values (time, space, etc.) are presented in subsection 2.

1. Graphical Presentation for Scalar Model Results
   - The empirical cumulative distribution function (cdf) of sample values
     The empirical cdf (see Fig. 5.1) shows immediately an estimate of the subjective probability for the model result Y to be smaller or equal to y. This format is

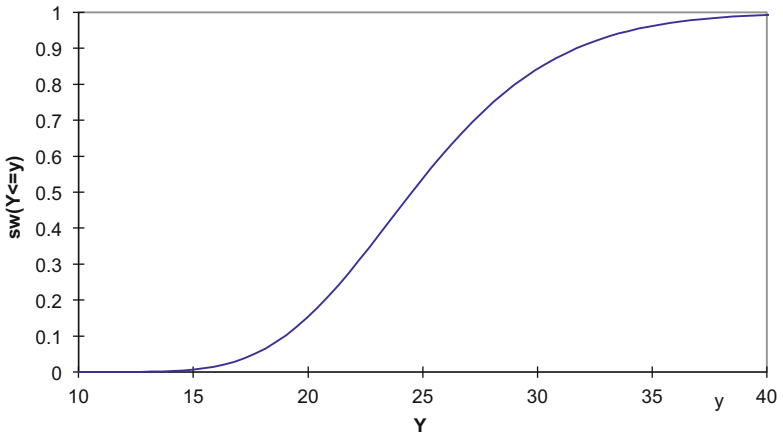


Fig. 5.1 Empirical cumulative subjective probability distribution function (cdf) of the sample values y_n, $n = 1,\ldots, N$, of a computer model result Y; sample size N is 200

of particular interest for the comparison of the model result Y to a lower limit value, i.e. a safety limit that the true value must exceed. The subjective probability for violation of the limit y_{lim} is the cumulative subjective probability of Y at y_{lim}.

– The empirical complementary cumulative distribution function (ccdf)
The empirical ccdf (see Fig. 5.2) shows immediately an estimate of the subjective probability for the model result Y to exceed y. This format is of particular interest for the comparison of the model result Y to an upper limit value, i.e. a safety limit that the true value must not exceed. The subjective probability for violation of the limit y_{lim} is the complementary cumulative subjective probability of Y at y_{lim}.
Figure 5.3 shows the usual way of how the uncertainty information from a subjective probability distribution for a model result is presented including point values and statistical tolerance limits.
Figures 5.4, 5.5 and 5.6 compare the uncertainty information from the subjective probability distributions for two model results.

– The histogram of sample values
The histogram gives immediately an impression of the range of values of the model result Y with the highest subjective probability. In other words, it depicts the range where the estimated subjective probability is concentrated. This way it also facilitates the comparison in one graph of the uncertainty of several model results $Y_1, \ldots, Y_K$ that may have similar 90% intervals or 90% quantile values, while the subjective probability within these ranges is concentrated at significantly different locations.

– The "bar code"
This representation shows each individual sample value of the model result Y by a horizontal stroke within a vertical box that contains 90% of the sample

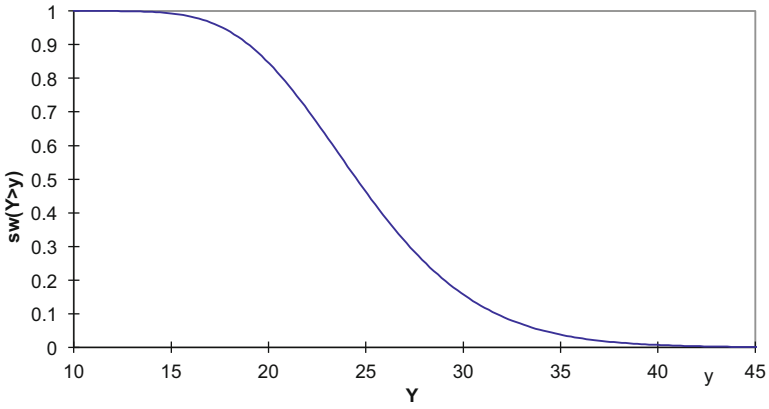


Fig. 5.2 Empirical complementary cumulative subjective probability distribution function (ccdf) of the sample values y_n, $n = 1,\ldots, N$, of a computer model result Y; sample size N is 200

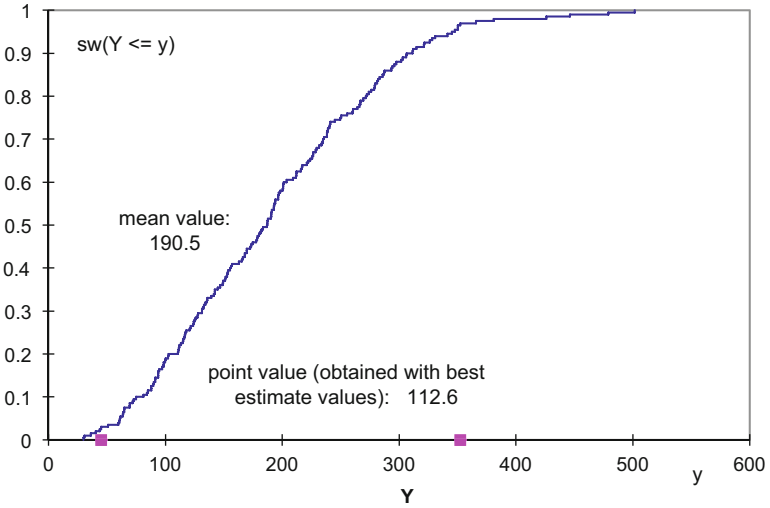


Fig. 5.3 Empirical cumulative subjective probability distribution function of the sample values y_n, $n = 1,\ldots, N$, of a computer model result Y; sample size N is 200; a two-sided (90%, 95%) statistical tolerance limit is indicated on the abscissa; an estimate of the mean value and the point result obtained with best estimate values of the uncertain data are given in the legend

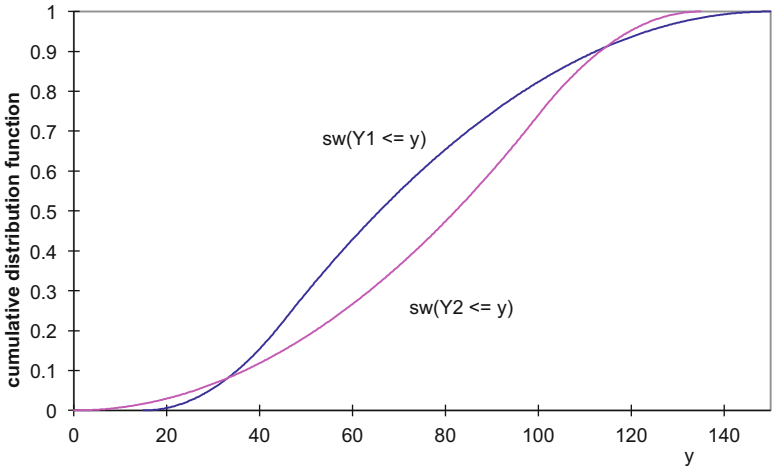


Fig. 5.4 Comparison of the cumulative subjective probability distribution functions for two computer model results

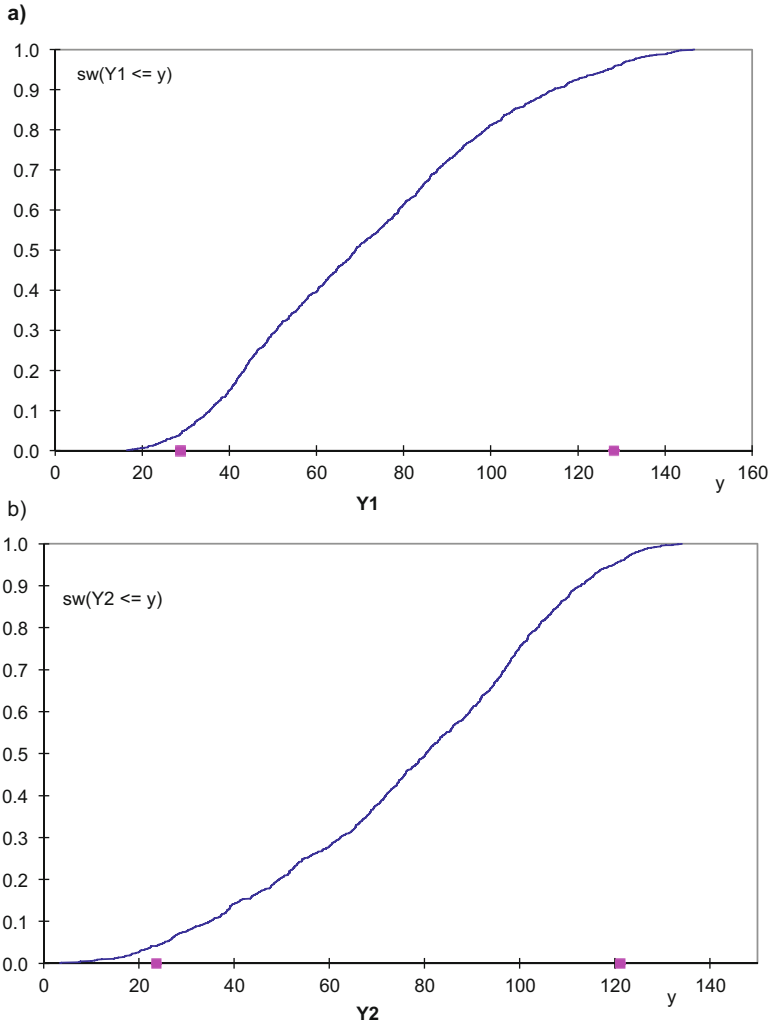


Fig. 5.5 (**a**, **b**) Comparison of the (90%, 95%) two-sided statistical tolerance limits shows that, at a confidence level of at least 95%, the (symmetric in probability) range that contains at least 90% of the population of possibly true values of model result Y_1 is shifted to higher values if compared to that of Y_2 in figure (**b**)

values. It needs less space than a histogram and still allows the identification of those ranges where the subjective probability is concentrated.

– Box and Whisker plot
Figures 5.7 and 5.8 show another way of facilitating the comparison of the subjective probability distributions for several model results $Y_1, \ldots, Y_J$. It is often most informative to compare estimates of the inner quartile ranges (the boxes), i.e. the ranges between the 25% and 75% quantile values of the

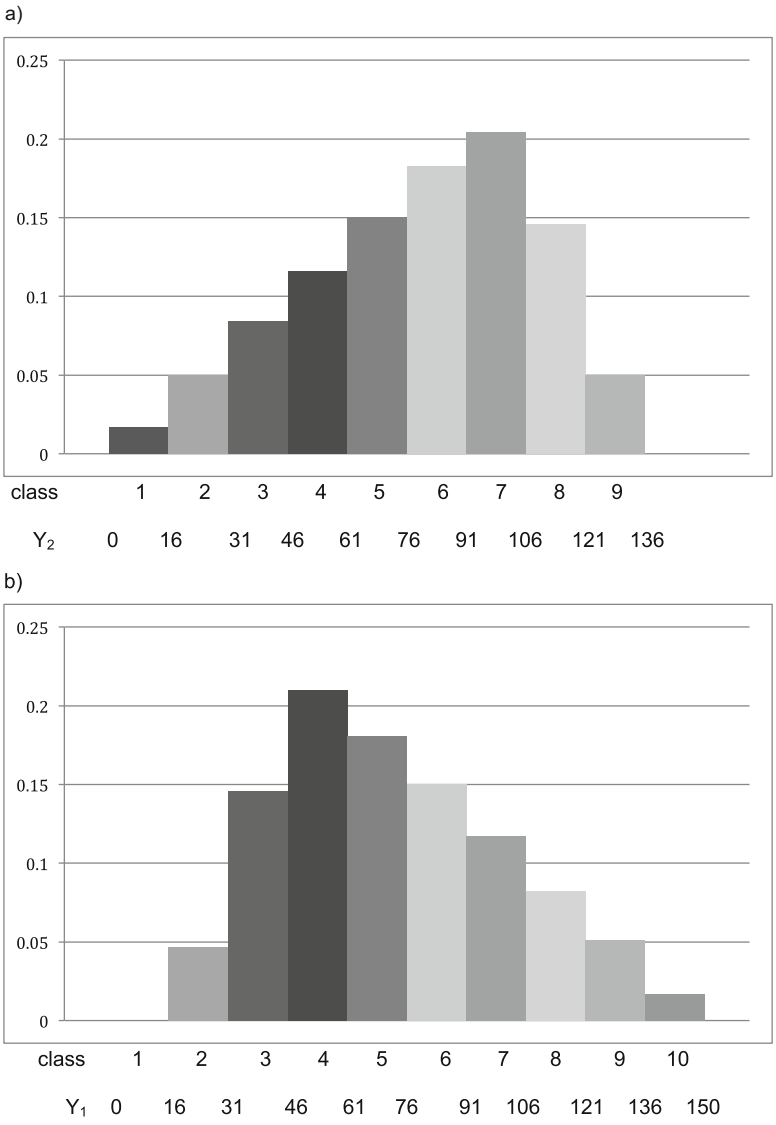


Fig. 5.6 (**a**, **b**) The relative frequencies of sample values within the classes indicated on the abscissa are shown for two computer model results. The comparison of Figures 5.5 and 5.6 says that although the (90%, 95%) two-sided statistical tolerance limit for Y_1 is shifted to larger values, the estimated subjective probability is concentrated at smaller values than in the case of Y_2. The relative frequency of sample values in class l, $l = 1, \ldots, 10$, is shown on the vertical axis

subjective probability distribution while simultaneously looking at estimates of the inner 90% ranges (between the 5% and 95% quantile values) and while not ignoring the extreme sample values (usually indicated by a dot) at both ends as well as the positions of the sample median and mean value relative to the limits of the inner quartile range.

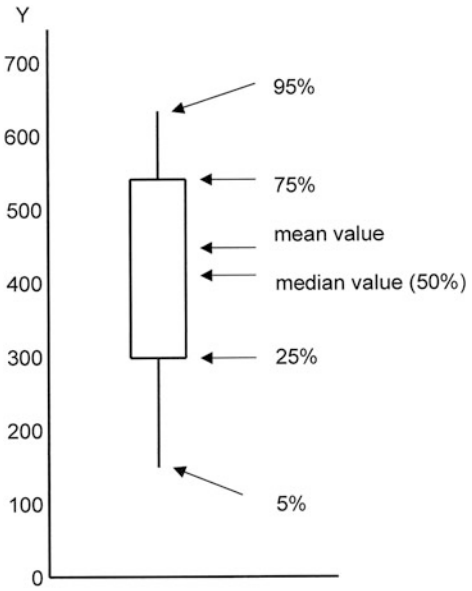


Fig. 5.7 Box and Whisker plot; the largest sample value could be indicated by a dot above the top end of the vertical bar, while the smallest sample value could be shown by a dot below the bottom end of the vertical bar

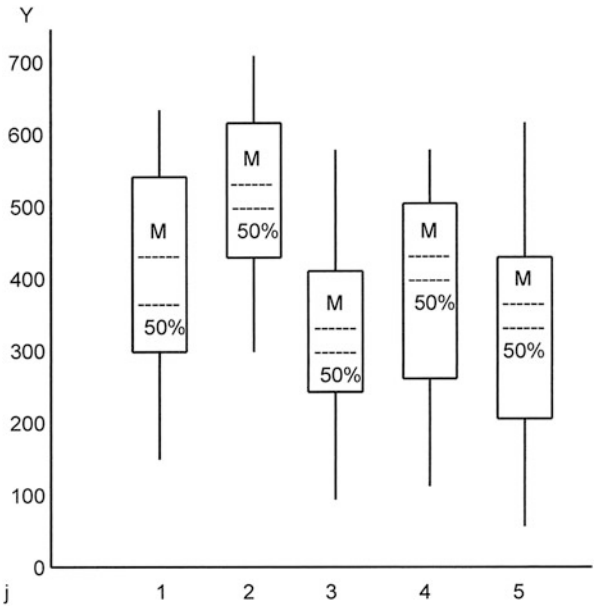


Fig. 5.8 Comparison of uncertainty analysis results for $J = 5$ different computer model results where the horizontal lines indicate the location of the mean value (M) and of the median value (50%)

A comparison of uncertainties of model results at 11 different locations and for 3 case specifications by Box and Whisker plots is shown in Hanna et al. (1998); A thin line, which extends from the box to the estimates of the 97.5%

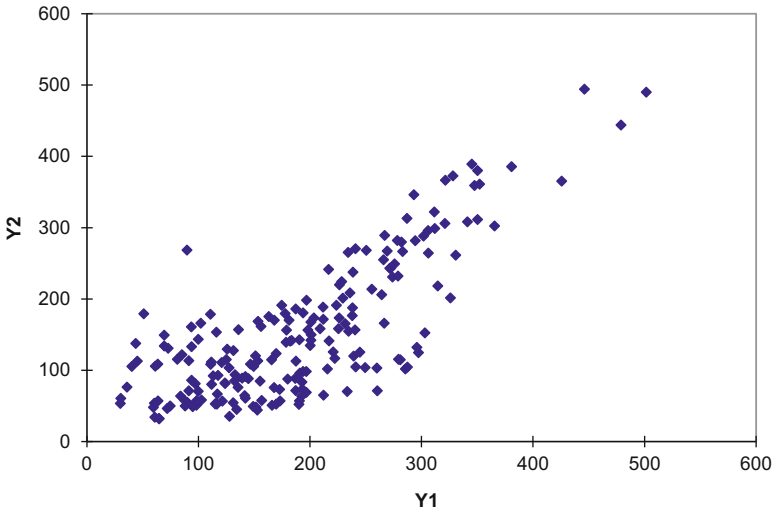


Fig. 5.9 Scatter diagram of a pair of computer model results. The plot indicates that the values computed for the model result Y_2 tend to be high if high values are computed for model result Y_1, and they tend to be low in the case of low values for Y_1. It may be informative to use the number of the corresponding run of the computer model as marker, instead of the diamond

and 2.5% quantile values, indicates the largest and smallest sample values in this example.

– Scatter plot of sample values
  The presentation format of Fig. 5.9 helps to depict state of knowledge dependences between pairs of model results.

2. Graphical Presentation for Functions
   Figures 5.10 and 5.11 illustrate the presentation, over a sequence of discrete points in time, of estimated subjective probability quantile values and of two-sided statistical tolerance limits of a model result that is a function of the independent variable "time".

The temporal evolution of the mean value of the locally (at the given points in time) obtained subjective probability distributions of the model result "population of guano-birds" as well as of the model result value obtained with the reference or best estimate values of the uncertain data could also be shown in Figures 5.10 and 5.11.

Figure 5.10 shows continuous connections of the quantile estimates of local subjective probability distributions of the model result, i.e. obtained at the specified points in time.

Figure 5.11 presents continuous connections of the local statistical tolerance limits. Since these limits are obtained locally, it cannot be said that, at a confidence level of at least v%, at least u% of the population of temporal evolutions do not cross the two lines connecting the endpoints of the local statistical tolerance limits.

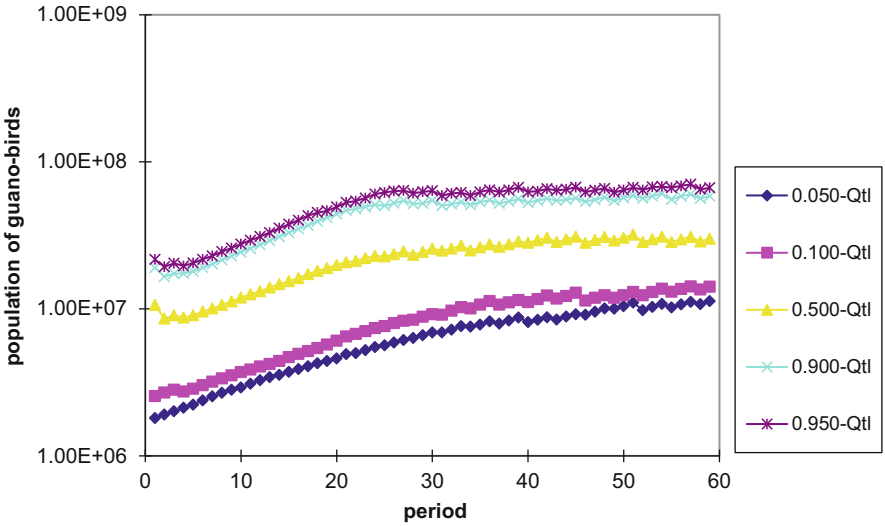


Fig. 5.10 The continuous connections of selected quantile values of the locally (at the endpoints of 60 intervals of four months each) obtained subjective probability distributions for the model result "population of guano-birds"; the quantile values are estimated using a random sample of size $N = 1000$ (1000 model runs each over 60 periods of four months)

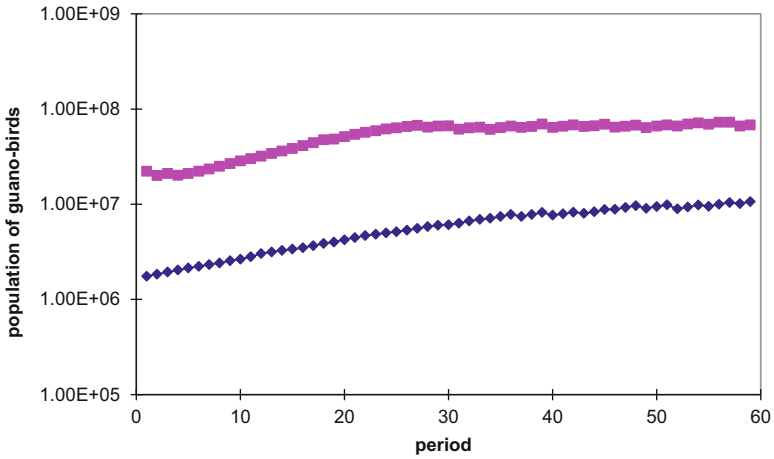


Fig. 5.11 The continuous connections of the locally (at the endpoints of 60 intervals of four months each) computed (90%, 95%) two-sided statistical tolerance limits of the model result "population of guano-birds"; the tolerance limits are obtained using a random sample of size $N = 1000$. The limits contain at least 90% subjective probability at a confidence level of at least 95% at each of the discrete points of time (time interval endpoints)

Locally, at least $u\%$ run between the two endpoints at a confidence level of at least $v\%$. However, they need not be the same temporal evolutions at every point in time.

For more ways of graphical presentation see also Ibrekk and Morgan (1987) and Morgan and Henrion (1990).

5.4.2 *Graphical Presentation of Uncertainty Analysis Results Obtained Using LHS*

The output sample values obtained with the values of a Latin Hypercube input sample may be used for graphical presentation in the same way as those obtained using SRS. The only exception is that confidence and statistical tolerance limits and intervals are not available for the uncertainty measures of the model results.

Confidence limits and intervals may, however, be derived from the estimates computed with the output sample values obtained using K replicated Latin Hypercube input samples (see Sect. 5.3.8).

References

Conover, W. J. (1980). *Practical nonparametric statistics* (2nd ed.). New York: Wiley.

Guenther, W. C. (1985). Two-sided distribution-free tolerance intervals and accompanying sample size problems. *Journal of Quality Technology, 17*(1), 40–43.

Hanna, S. R., Chang, J. C., & Fernau, M. E. (1998). Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atmospheric Environment, 32*(21), 3619–3628.

Hansen, C. W., Helton, J. C., & Sallaberry, C. J. (2012). Use of replicated Latin hypercube sampling to estimate sampling variance in uncertainty and sensitivity analysis results for the geologic disposal of radioactive waste. *Reliability Engineering and System Safety, 107*, 139–148.

Heinhold, J., & Gaede, K.-W. (1968). *Ingenieur-Statistik*. Wien: R. Oldenbourg München.

Ibrekk, H., & Morgan, M. G. (1987). Graphical communication of uncertain quantities to nontechnical people. *Risk Analysis, 7*(4), 519–529.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics, 21*(2), 239–245.

Morgan, M. G., & Henrion, M. (1990). *Uncertainty—A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge: Cambridge University Press.

Scheffé, H., & Tukey, J. W. (1944). A formula for sample sizes for population tolerance limits. *The Annals of Mathematical Statistics, 15*, 217.

Wald, A. (1943). An extension of Wilks method for setting tolerance limits. *The Annals of Mathematical Statistics, 14*, 45–55.

Wilks, S. S. (1941). Determination of sample sizes for setting tolerance limits. *The Annals of Mathematical Statistics, 12*, 91–96.

Winkler, R. L., & Hays, W. L. (1975). *Statistics: Probability, inference, and decision* (2nd ed.). New York: Holt, Rinehart and Winston.

Chapter 6
Step 5: Rank Uncertainties

6.1 Introduction

The primary goal is the quantification of the uncertainty of results from the application of a computer model. The analysis therefore focuses on the combined influence of all potentially important uncertain data[1] on the model results. To this end, the state of knowledge with respect to model formulations, model parameters, application-specific input data and parameters of numerical solution algorithms (in short: for all potentially important uncertain data) is quantified and expressed by subjective probability distributions. State of knowledge dependences, which might be influential, are suitably quantified and modelled (see Sect. 3.5).

As a consequence of these quantifications and of the logic encoded in the computer model, a subjective probability distribution follows for each of the model results. Quantitative uncertainty statements in the form of, for instance, 5% and 95% quantiles of these distributions could be immediately obtained if the distributions were known. In practice, they are not known and quantile values have to be estimated from a random sample drawn according to the unknown distributions.

To obtain such a random sample of size N (output sample), Monte Carlo simulation is performed (see Chap. 4). It proceeds by drawing a random sample (input sample) according to the marginal subjective probability distributions and state of knowledge dependence quantifications specified for the uncertain data and by subsequently performing a model run with each set of sampled values. N such sets are drawn, where all uncertain data are varied simultaneously, and the computer model is evaluated for each of these N sets. Since models are often computationally demanding (several hours or tens of hours on a modern processor), one can only afford samples of relatively small to moderate size N. The N sets of model results, obtained through running the model for each of the N sets of sampled input values,

[1] As in previous chapters, the term "uncertain data" is used since model uncertainties are represented by uncertain parameters and uncertain parameters are categorized as uncertain data.

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_6

constitute a random sample (output sample) drawn according to their unknown joint subjective probability distribution. The desired quantile values can be estimated from this sample by standard statistical techniques.

For almost all practically relevant model applications, uncertainty analysis proceeds by Monte Carlo simulation. It grants the analyst all the freedom needed in the quantification and probabilistic modelling of the states of knowledge as well as of any state of knowledge dependences. The affordable sample size, i.e. the number of evaluations of the model (for short called "model runs"), is usually small for computationally demanding models. Consequently, confidence intervals or limits are to be determined together with the estimates of model result uncertainty (see Chap. 5).

For consistency and efficiency reasons, uncertainty importance analysis cannot afford a separate set of model runs, performed with specifically chosen sets of values for the uncertain data, but has to use those executed for the purpose of uncertainty analysis. Consequently, correlation coefficients, correlation ratios (square root of the approximate first-order importance indices from variance decomposition) and standardized regression coefficients, the latter often obtained from stepwise regression, are with or without transformation into ranks, a reasonable choice of importance measures. They are available from the output sample of the Monte Carlo simulation in combination with the input sample.

Of course, if the subjective probability distribution of a model result is long-tailed, which may frequently be the case, estimates of quantile values situated in the tail regions will be highly variable from sample to sample of size N. Therefore, one needs to quantify the possible impact of the estimation error (also called "sampling error") by computing one-sided or two-sided (u%, v%) statistical tolerance limits. With, for instance, $u = 90$ and $v = 95$, a two-sided statistical tolerance limit contains, at a confidence level of at least 95%, at least 90% of the population of model result values that follow from the combined effect of the state of knowledge quantifications for all uncertain data considered by the uncertainty analysis.

Quantile value estimates from a Latin Hypercube sample (LHS) can be expected to show less variability from sample to sample of size N, yet statistical tolerance limits cannot be computed from an LHS as was explained in Chap. 4. The sample estimates of the 95% quantile, for instance, may still be significantly below the true 95% quantile for long-tailed distributions. Without statistical tolerance limits, there is no indication of how likely this shortcoming is. This constitutes a handicap, if quantile value estimates are to be compared to limit values in order to check compliance with safety standards. As a consequence of this, either replicated LHS (see Sects. 4.4.2.1 and 5.2.8) or simple random sampling (SRS) needs to be used.

Usually, the model application is to provide input to some important decision. Once the quantitative uncertainty statements are derived, the question arises whether uncertainties are too large for meaningful decision-making. If the uncertainty is judged to be too large, one needs to know its main contributors. To this end, uncertainty importance analysis is performed. It tells where to improve the state of knowledge in order to reduce the uncertainty of the model result most effectively. In other words, it provides guidance as to whether further model development is primarily needed or improved knowledge of parameter and input data values. This

is a choice between two very different directions of further activities with differing costs and chances for success. For instance, the importance analysis may suggest that additional tests, experiments and theoretical investigations leading to improved model development are required or that further data collection has priority.

The multivariate random (input) sample of N sets of values, each set consisting of one value each for the M uncertain data, will show the effects of the specified state of knowledge dependences and most likely those of spurious correlations. Due to the presence of correlations within the multivariate sample, it will not be sufficient to look at the uncertainty importance ranking derived from correlation coefficients alone. Differences between this ranking and one obtained from standardized regression coefficients need to be understood if significant. The correlation ratio (square root of the first-order uncertainty importance measure from variance decomposition) is an indispensable measure whenever model uncertainty is expressed by more than two alternatives with their indices used as values of a substitute uncertain datum (see Sect. 3.4.1) or when measures, quantifying the extent of linear or monotone relationship between an uncertain datum and a model result, are not adequate. Due to the small sample size, the computation of the sample correlation ratio is affordable only in approximate form. An excellent compilation of additional uncertainty importance measures is to be found in Helton et al. (2006).

Figure 6.1 presents results from an uncertainty analysis of the damage index associated with a planned potentially hazardous industrial operation, computed as a function of the distance. The uncertainty is quantified by intervals obtained as two-sided (95%, 95%) statistical tolerance limits from an SRS with $N = 1000$ model runs. The figure shows the continuous connections of the upper (and lower) endpoints of these intervals computed at various distances from the operation. Close to the facility, the uncertainty of the model result "damage index" is too large to permit meaningful decision-making.

Importance measures from correlation and regression (see Sect. 6.3), derived from an SRS of relatively small size N, may still exhibit considerable variability from sample to sample of size N. This variability can be expected to be smaller in the case of an LHS of the same size. However, because of what was said above, an LHS is often not an option and using a separate random sample for uncertainty importance analysis is usually unaffordable because of processor-time (and calendar time)

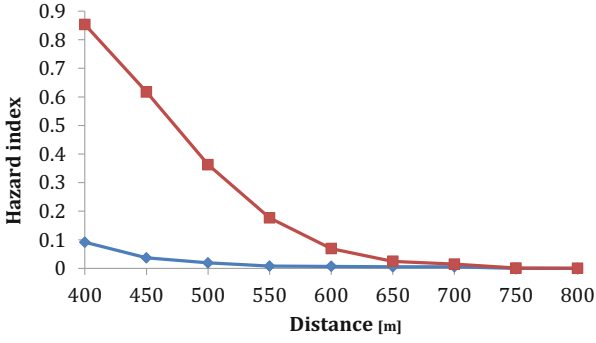


Fig. 6.1 Continuous connections of the upper (and lower) endpoints of the two-sided (95%, 95%) statistical tolerance limits for the damage index computed at various distances from the location of the planned potentially hazardous operation; importance measures to this example are discussed in Sect. 3.3.

requirements. It is not only cost-effective (and in fact the only practically feasible way) but also natural and mathematically consistent, to exploit the same random sample used in uncertainty analysis for the purpose of identifying the main contributors to model result uncertainty. The effort involved in determining the importance measures of Sect. 6.3 is comparatively low. These measures are almost a by-product of the uncertainty analysis.

6.2 Differential Sensitivity and "One-at-a-Time" Analysis

Difference or differential quotients of the model result with respect to each of the uncertain data tell by how much the model result changes if the uncertain datum is changed from nominal by a small amount. It will be impossible to compute these quotients with respect to uncertain data that can have only a finite number of possibly true values and in particular for those that are indices of alternatives like alternative model formulations used to express the state of knowledge for model uncertainties.

The best estimate or reference values (see Sect. 3.6.1.4), or any values chosen at random according to the marginal subjective probability distribution specified as state of knowledge expression, may be used as the nominal values of the uncertain data. With Y denoting the model result, $P_1, \ldots, P_M$ the uncertain data and f the encoded logic of the computer model, the relationship between model result and uncertain data is formally represented by $Y = \mathrm{f}(P_1, \ldots, P_M)$.

At the point $\boldsymbol{p}_0 = p_{0,1}, \ldots, p_{0,M}$ of nominal values in the space spanned by the $\boldsymbol{M}$ uncertain data, the first-order Taylor series expansion of the model result is

$$\widehat{Y} = y_0 + \sum\nolimits_{m=1}^{M} \left[(\partial \mathrm{f}/\partial p_m)\Big|_{\boldsymbol{p}_0} \right] (P_M - p_{0,m}) \tag{6.1}$$

with $y_0 = \mathrm{f}(\boldsymbol{p}_0)$ and provided the function f is differentiable with respect to P_m, $m = 1, \ldots, M$. As is indicated by $|_{\boldsymbol{p}_0}$, the partial derivatives are taken at $\boldsymbol{p}_0$. They are the differential quotients at this point.

The mean value and variance of the subjective probability distribution of the first-order Taylor series expansion of Y about $\boldsymbol{p}_0$ are

$$\mathrm{E}\left\{\widehat{Y}\right\} = y_0 + \sum\nolimits_{m=1}^{M} \left[(\partial \mathrm{f}/\partial p_m)\Big|_{\boldsymbol{p}_0} \right] (\mathrm{E}\{P_M\} - p_{0,m}) \tag{6.2}$$

and

$$\mathrm{Var}\left\{\widehat{Y}\right\} = \sum_{m=1}^{M} \left[(\partial \mathrm{f}/\partial p_m)\big|_{\boldsymbol{p}_0} \right]^2 \mathrm{Var}\{P_M\}$$

$$+2 \sum\nolimits_{m=1}^{M} \sum\nolimits_{l=m+1}^{M} \rho_{m,l} (\{\mathrm{Var} P_m\} \mathrm{Var}\{P_l\})^{\frac{1}{2}} [(\partial \mathrm{f}/\partial p_m)(\partial \mathrm{f}/\partial p_l)]\big|_{\boldsymbol{p}_0}. \tag{6.3}$$

A correlation coefficient $\rho_{m,\,l} \neq 0$ between P_m and P_l is due to state of knowledge dependence between these uncertain data.

Equation (6.3) can be used to judge the uncertainty contribution of each of the uncertain data to the uncertainty of $\widehat{Y}$ as measured by the variance. However, the representation of $\mathrm{Var}\{Y\}$ by $\mathrm{Var}\{\widehat{Y}\}$ for the purpose of uncertainty importance ranking suffers from a number of shortcomings:

– The derivatives are local (at $\boldsymbol{p}_0$) and often apply only in a small neighbourhood about the selected nominal point, i.e. the ranking will not be suitable in the case of:

  Moderate to strong non-linearities of the function f over the carrier of the joint subjective probability distribution of the uncertain data;
  Thresholds for Y (discontinuities);
  Contributions to model result uncertainty that are large in absolute value only if the values of two or more of the uncertain data are changed simultaneously from nominal.

– The ranking ignores the state of knowledge, as expressed by the full marginal subjective probability distributions for the uncertain data.
– It will be very expensive to compute the partial derivatives [even if automatic differentiation tools are applied (Bischof et al. 2008)] for functions f represented by computationally demanding complex models and for realistic numbers M of uncertain data.

The partial derivatives of a model result with respect to the uncertain data, computed at selected points in the parameter space, are local sensitivity measures in the usual interpretation. They are independent of the subjective probability distributions expressing the state of knowledge for the uncertain data. The selected points may be those of the multivariate sample generated by the Monte Carlo simulation for the purpose of uncertainty analysis. The mean value of the partial derivatives taken over these points could then be used as a global sensitivity measure that is indeed dependent on the subjective probability distributions (Kucherenko et al. 2009). Computing those derivatives at all N points for all M, uncertain data will, however, prove to be unaffordable for practically relevant model applications even if commercially available software for automatic differentiation (Bischof et al. 2008) can be applied or adjoint methods are used (Cacuci 2003).

Model developers often resort to so-called one-at-a-time variations in order to get an impression of how sensitive a model result is to individual uncertain data. They proceed as follows:

– Instead of a small change, as in the case of differential or difference quotients, each uncertain datum is individually changed from nominal by $\pm$ one or two standard deviations of its marginal subjective probability distribution.
– An uncertain model is varied from its standard formulation (an alternative model formulation is used).
– A phenomenon of uncertain relevance is included or excluded from the computer model application.
– The description of an uncertain scenario is modified (an alternative scenario is considered).

A model run is then performed with each individual change.
These "one-at-a-time" changes are not satisfactory because:

- They require model runs in addition to those needed for the uncertainty analysis.
- They provide sensitivity information that depends on the chosen nominal values and on the magnitudes of the chosen value changes.
- The selected changes of the uncertain data are somewhat arbitrary and do not account for the range of possibly applicable values.
- The states of knowledge, expressed by the subjective probability distributions of the uncertain data, are not fully accounted for and state of knowledge dependences are ignored.
- The effect of thresholds of the model result and of strong model non-linearities may not be captured.
- The procedure is unaffordable for moderate to large numbers of uncertain data unless it is restricted to a small subset selected by expert judgement.

6.3 Affordable Measures for Uncertainty Importance Ranking

The objective of an uncertainty importance analysis is to rank uncertain data with respect to their contribution to the uncertainty of the model result. To this end, the analysis must be global and not local, i.e. the validity of the provided ranking must not be limited to a small neighbourhood around a specific point in the value space of the uncertain data. Furthermore, the importance analysis must account for the state of knowledge as expressed by the subjective probability distributions specified for the uncertain data and it must observe state of knowledge dependences. Finally, the analysis must have the potential of accounting for discontinuities of the model result and for effects of varying uncertain data simultaneously and over their full uncertainty range. Last but not least, the computational effort of evaluating the uncertainty importance measures must be affordable.

Affordable measures are those that can be derived from the same data that were used for the uncertainty analysis. These data are the random sample of N times M values of the uncertain data (input sample) and the corresponding random sample of N values of the model result (output sample). The latter sample having been obtained from the first through evaluation of the computer model for each of the N sets of M data values, i.e.

$$\begin{aligned} y_1 &= \mathrm{f}\left(p_{1,1}, \ldots, p_{M,1}\right) \\ y_2 &= \mathrm{f}\left(p_{1,2}, \ldots, p_{M,2}\right) \\ &\vdots \\ y_N &= \mathrm{f}\left(p_{1,N}, \ldots, p_{M,N}\right) \end{aligned} \tag{6.4}$$

6.3.1 *Uncertainty Importance Measures Computed from Raw Data*

This section deals with uncertainty importance measures that are obtained directly from the data in the input sample and in the output sample of the Monte Carlo uncertainty analysis. In particular, measures from regression (Sects. 6.3.1.1 and 6.3.1.3), correlation (Sect. 6.3.1.2) and variance decomposition (Sects. 6.3.1.4 and 6.3.1.5) are discussed. This discussion is followed by recommendations and necessary checks in Sect. 6.3.1.6.

6.3.1.1 Uncertainty Importance Measures from Regression Analysis

The clearest picture of the uncertainty importance of the individual uncertain data can be obtained from a function g that is linear in these data and approximates the model result $Y = \mathrm{f}(P_1, \ldots, P_M)$ sufficiently well. The meaning of "sufficiently" will be explained later. The coefficients $\beta_0, \beta_1, \ldots \beta_M$ of

$$\beta_0 + \sum_{m=1}^{M} \beta_m P_m + E = \mathrm{g}(P_1, \ldots, P_M) + E = \widehat{Y} + E = \mathrm{f}(P_1, \ldots, P_M) = Y, \quad (6.5)$$

where $\widehat{Y}$ denotes the linear approximation to Y and E is the approximation error, are estimated from the random sample (6.4) by fitting the following expressions:

$$\widehat{y}_n = \widehat{\mathrm{g}}\left(p_{1,n}, \ldots, p_{M,n}\right) = b_0 + \sum_{m=1}^{M} b_m p_{m,n}, \quad n = 1, \ldots, N \quad (6.6)$$

to the sample values y_n such that the approximation error satisfies a prescribed condition. The aim is to derive uncertainty importance measures. Uncertainty is here measured by the variance of the subjective probability distribution quantifying the state of knowledge of the model result. Therefore, it seems only reasonable to compute the estimates $b_0, b_1, \ldots, b_M$ such that the portion of the sample variance of the model result, which is not explained by the approximation (6.6), is minimal. To this end, $b_0, b_1, \ldots, b_M$ are obtained according to the following condition:

$$\sum_{n=1}^{N} \left(y_n - \widehat{y}_n\right)^2 = \sum_{n=1}^{N} \left(y_n - b_0 - \sum_{m=1}^{M} b_m p_{m,n}\right)^2 =^{!} \textit{minimal} \quad (6.7)$$

The optimization problem (6.7) is solved through setting the partial derivatives of

$$h(b_0, b_1, \ldots, b_M) = \sum_{n=1}^{N} \left(y_n - b_0 - \sum_{m=1}^{M} b_m p_{m,n}\right)^2 \quad (6.8)$$

to zero with respect to $b_0, b_1, \ldots, b_M$:

$$\partial h / \partial b_0 = -2 \sum_{n=1}^{N} \left(y_n - b_0 - \sum_{m=1}^{M} b_m p_{m,n}\right) = 0 \quad (6.9)$$

and for $k = 1, \ldots, M$

$$\partial h/\partial b_k = -2\sum_{n=1}^{N}\left(y_n - b_0 - \sum_{m=1}^{M} b_m p_{m,n}\right)p_{k,n} = 0. \tag{6.10}$$

From (6.9) follows

$$b_0 = \left[\sum_{n=1}^{N} y_n - \sum_{n=1}^{N}\sum_{m=1}^{M} b_m p_{m,n}\right]/N = m_y - \sum_{m=1}^{M} b_m m_m \tag{6.11}$$

$$m_m = \left(\frac{1}{N}\right)\sum_{n=1}^{N} p_{m,n}$$

is the sample mean of the uncertain datum P_m.
From (6.6) and (6.11) follows

$$m_{\hat{Y}} = b_0 + \sum_{m=1}^{M} b_m m_m = m_Y \tag{6.12}$$

with m_Y the sample mean of Y.

To simplify the presentation of the following calculations, the customary matrix and vector notation is being used. Matrices are denoted by bold capital letters and vectors by bold lower case letters. Without this notation, it would be very cumbersome and awkward to follow through, and concisely present, the multitude of arithmetic operations involved. For those readers who are not familiar with this notation, the main operations are briefly explained. For a more detailed explanation, Bowerman et al. (2005) may be consulted.

The N times M sample values of the uncertain data (input sample) are arranged in an $N \times M$ array. This array is extended by a column vector carrying the coefficient value 1 of b_0 from the system of equations (6.6). The result is an $N \times (M + 1)$ array called matrix $\boldsymbol{P}$ with N rows and $M + 1$ columns. The N sample values of the model result (output sample) are arranged in an $N \times 1$ array, called column vector $\boldsymbol{y}$, while the estimates $b_0, b_1, \ldots, b_M$ are arranged in an $(M + 1) \times 1$ array called column vector $\boldsymbol{b}$.

$$\boldsymbol{P} = \begin{bmatrix} 1 & p_{1,1} & p_{2,1} & \cdots & p_{M,1} \\ 1 & p_{1,2} & p_{1,2} & \cdots & p_{M,2} \\ & & & \cdot & \\ & & & \cdot & \\ & & & \cdot & \\ 1 & p_{1,N} & p_{2,N} & \cdots & p_{M,N} \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_1 \\ \cdot \\ \cdot \\ b_M \end{bmatrix} \tag{6.13}$$

The $(M + 1) \times N$ matrix $\boldsymbol{P}'$ has the columns of $\boldsymbol{P}$ as rows and the rows of $\boldsymbol{P}$ as columns.

$$\boldsymbol{P}' = \begin{bmatrix} 1 & 1 & \dots 1 \\ p_{1,1} & p_{1,2} & \cdots p_{1,N} \\ p_{2,1} & p_{2,2} & \cdots p_{2,N} \\ & & \cdot \\ & & \cdot \\ & & \cdot \\ p_{M,1} & p_{M,2} & \cdots p_{M,N} \end{bmatrix}, \boldsymbol{P}'\boldsymbol{P} = \begin{bmatrix} N & \sum_{n=1}^{N} p_{1,n} & \sum_{n=1}^{N} p_{2,n} & \cdots \sum_{n=1}^{N} p_{M,n} \\ \sum_{n=1}^{N} p_{1,n} & \sum_{n=1}^{N} p_{1,n}p_{1,n} & \sum_{n=1}^{N} p_{1,n}p_{2,n} & \cdots \sum_{n=1}^{N} p_{1,n}p_{M,n} \\ \sum_{n=1}^{N} p_{2,n} & \sum_{n=1}^{N} p_{2,n}p_{1,n} & \sum_{n=1}^{N} p_{2,n}p_{2,n} & \cdots \sum_{n=1}^{N} p_{2,n}p_{M,n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \sum_{n=1}^{N} p_{M,n} & \sum_{n=1}^{N} p_{M,n}p_{1,n} & \sum_{n=1}^{N} p_{M,n}p_{2,n} & \cdots \sum_{n=1}^{N} p_{M,n}p_{M,n} \end{bmatrix} \tag{6.14}$$

The $(M + 1) \times (M + 1)$ matrix $\boldsymbol{P}'\boldsymbol{P}$ results as the product of the matrices $\boldsymbol{P}'$ and $\boldsymbol{P}$. The element in row i and column j of $\boldsymbol{P}'\boldsymbol{P}$ is the sum of the products of the elements in row i of $\boldsymbol{P}'$ with those in column j of $\boldsymbol{P}$ where the summation runs from 1 to N. This product immediately demonstrates the convenience of the matrix/vector notation as do the products of matrix $\boldsymbol{P}'$ and vector $\boldsymbol{y}$ as well as of matrix $\boldsymbol{P}'\boldsymbol{P}$ and vector $\boldsymbol{b}$, both providing a column vector with $M + 1$ components. The elements of these products are obtained as explained for those of $\boldsymbol{P}'\boldsymbol{P}$, the only difference being that there is only one column to be multiplied with the matrix rows.

$$\boldsymbol{P}'\boldsymbol{y} = \begin{bmatrix} \sum_{n=1}^{N} y_n \\ \sum_{n=1}^{N} p_{1,n}y_n \\ \sum_{n=1}^{N} p_{2,n}y_n \\ \cdot \\ \cdot \\ \cdot \\ \sum_{n=1}^{N} p_{M,n}y_n \end{bmatrix} = (\boldsymbol{y}'\boldsymbol{P})', \quad \boldsymbol{P}'\boldsymbol{P}\boldsymbol{b} = \begin{bmatrix} Nb_0 + \sum_{m=1}^{M} b_m \sum_{n=1}^{N} p_{m,n} \\ b_0 \sum_{n=1}^{N} p_{1,n} + \sum_{m=1}^{M} b_m \sum_{n=1}^{N} p_{1,n}p_{m,n} \\ b_0 \sum_{n=1}^{N} p_{2,n} + \sum_{m=1}^{M} b_m \sum_{n=1}^{N} p_{2,n}p_{m,n} \\ \cdot \\ \cdot \\ b_0 \sum_{n=1}^{N} p_{M,n} + \sum_{m=1}^{M} b_m \sum_{n=1}^{N} p_{M,n}p_{m,n} \end{bmatrix} \tag{6.15}$$

$$\boldsymbol{y}' = (y_1, y_2, \dots, y_N)$$

the product of the row vector $\boldsymbol{y}'$ and the column vector $\boldsymbol{y}$ is a single number

$$\boldsymbol{y}'\boldsymbol{y} = \sum_{n=1}^{N} y_n^2. \tag{6.16}$$

Using the above notation, the operations in (6.10) and (6.9) can simply be written as

$$\boldsymbol{P}'\boldsymbol{y}\text{-}\boldsymbol{P}'\boldsymbol{P}\boldsymbol{b}=0 \tag{6.17}$$

and therefore

$$\boldsymbol{b} = (\boldsymbol{P}'\boldsymbol{P})^{-1}\boldsymbol{P}'\boldsymbol{y}. \tag{6.18}$$

$(\boldsymbol{P}'\boldsymbol{P})^{-1}$ is the inverse of the matrix product $\boldsymbol{P}'\boldsymbol{P}$. It too is an $(M + 1) \times (M + 1)$ matrix, and its elements are determined (standard subroutines are available) such that

$$(\boldsymbol{P}'\boldsymbol{P})^{-1}(\boldsymbol{P}'\boldsymbol{P}) = \boldsymbol{I} = \begin{bmatrix} 1 & 0 & \ldots & . & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ & & & & . \\ & & & & . \\ 0 & \ldots & . & 0 & 1 \end{bmatrix}$$

i.e. the unit matrix with all elements outside the diagonal equal to 0, while the values along the diagonal are 1.

The elements of the vector $\boldsymbol{b}$ are sample estimates of the coefficients $\beta_0, \beta_1, \ldots, \beta_M$. They are obtained from a least squares linear approximation (6.6) of the sample values $y_1, \ldots, y_N$ in the uncertain data. b_m is the amount of units by which the linear approximation

$$\widehat{Y} = b_0 + \sum_{m=1}^{M} b_m P_m$$

changes per unit change of the uncertain datum P_m, all other uncertain data remaining constant.

Some cautionary remarks are in place at this point:

- The sample size N must not be smaller than $M + 1$. For smaller N, condition (6.7) would mean fitting a function linear in $M + 1$ variables (i.e. the components of vector $\boldsymbol{b}$) to $N \leq M$ points in the $(M + 1)$-dimensional space. This problem has infinitely many solutions (i.e. not a unique one).
- The system of linear equations

  $b_0 + \sum_{m=1}^{M} b_m p_{m,n} = y_n, \quad n = 1, \ldots, N$

  in $M + 1$ unknowns can be solved exactly if $N = M + 1$ and $\boldsymbol{P}'\boldsymbol{P}$ is invertible. The linear function in the uncertain data therefore reproduces the N sample values of the model result Y, i.e. $\widehat{y}_n = y_n$ for $n = 1, \ldots, N$. Values $y_{N+1}, \ldots, y_{N+K}$ of Y for further K sampled sets of data values $p_{1,N+k}, \ldots, p_{M,N+k}, k = 1, \ldots, K$ can then be predicted exactly by the linear function if $Y = \mathrm{f}(P_1, \ldots, P_M)$ is indeed linear in the uncertain data. $\widehat{y}_n = y_n$, for $n = 1, \ldots, N$, means over-fitting the sample values $y_1, \ldots, y_N$ if $Y = \mathrm{f}(P_1, \ldots, P_M)$ is non-linear in one or more of the M uncertain data.

In this case, nothing can be said about the errors in the predictions $\widehat{y}_{N+1}, \ldots, \widehat{y}_{N+K}$ obtained from the linear function. However, if N is larger than $M + 1$, the least squares linear approximations $\widehat{y}_n$ will not be equal to the sample values $y_1, \ldots, y_N$ and the N error terms will, under certain assumptions, permit quantitative statements about the possible error in the predictions $\widehat{y}_{N+1}, \ldots \widehat{y}_{N+K}$. There may still be over-fitting of the sample data to some extent as long as N is not significantly larger than $M + 1$.

– A column k of sample values of an uncertain datum, which is a linear combination of the columns of sample values of other uncertain data, must not be included in the matrix $\boldsymbol{P}$, i.e. the corresponding datum P_k must not be part of the list of arguments of the approximating function in (6.5) because this linear combination property carries over to the matrix $\boldsymbol{P'P}$. In other words: In $\boldsymbol{P'P}$ the column k is also a linear combination of $\boldsymbol{P'P}$-columns with the same numbers as those columns in $\boldsymbol{P}$ that are combined in column k of $\boldsymbol{P}$. Consequently, the columns of $\boldsymbol{P'P}$ span only an M-dimensional space. However, solving the equation

$$(\boldsymbol{P'P})\boldsymbol{b} = \boldsymbol{P'y} \tag{6.19}$$

means to attempt finding $M + 1$ coefficient values $b_0, b_1, \ldots, b_M$ such that the vector $\boldsymbol{P'y}$ in the $(M + 1)$-dimensional space can be assembled as a linear combination of the $M + 1$ columns of $\boldsymbol{P'P}$. Since the latter span only an M-dimensional space, this task cannot be accomplished [see "multicollinearity" in Draper and Smith (1998), Freund and Minton (1979)], i.e. $\boldsymbol{P'P}$ is not invertible. If the column, which is a linear combination of others, is left out of matrix $\boldsymbol{P}$, then the columns of $\boldsymbol{P'P}$ still span only an M-dimensional space but $\boldsymbol{P'y}$ is then also only of dimension M and Eq. (6.19) can be solved. Obviously, there will be no coefficient b_k if the column of the sample values of P_k is left out. A practical case is given by K uncertain data that are fractions of a total (see Sect. 3.5.2.3). The sample value of P_k, with $k = K$, is determined as 1 minus the sum of the sample values of the other $K - 1$ uncertain data and P_K must therefore not be in the argument list of the approximating function in (6.5). The same applies to pairs of uncertain data that are completely state of knowledge dependent (see Sect. 3.5.2.6). Only the so-called free uncertain datum of such a pair may be among the arguments of the approximating function.

– A column of $\boldsymbol{P}$ may, by chance, be almost equal to a linear combination of other columns (near multicollinearity). The situation is close to the one discussed above. The information on some of the $M + 1$ dimensions of the solution space will only be weak in $\boldsymbol{P'P}$ and will be almost overridden by rounding effects during the attempt to solve (6.19) for $\boldsymbol{b}$. The solution vector obtained for $\boldsymbol{b}$ will be severely contaminated by the effect of rounding errors.
– If $(M + 1) < N$ but still close to N, one or more pairs of columns of the matrix $\boldsymbol{P}$ may have a sample correlation coefficient of above 0.5 or below -0.5 that is not due to specified state of knowledge dependence but due to chance (spurious correlation). The effect of these correlations on the solution of (6.19) is less pronounced yet closely related to near multicollinearity. To reduce the probability

of a spurious correlation > 0.5 and < −0.5 in the sample of MN data values, a sample size significantly larger than M would need to be drawn (see, for instance, Tables 6.4 and 6.5).

– The formulae used for the computation of the estimates $b_0, b_1, \ldots, b_M$ are those of least squares linear regression. Yet, this section has so far avoided the term "regression coefficients". The reason lies in the different situation of uncertainty importance analysis as compared to the usual problem setting of applied regression analysis. In the case of uncertainty importance analysis, there is only one unique value y_n to each set of data values $p_{1,n}, \ldots, p_{M,n}$ since $Y = \mathrm{f}(P_1, \ldots, P_M)$ is deterministic while in regression analysis the value of Y at $\boldsymbol{p}'_{\boldsymbol{n}} = (p_{1,n}, \ldots, p_{M,n})$ is a random variable for $n = 1, \ldots, N$. In uncertainty importance analysis, the error term E of (6.5) is therefore deterministic at every point in the M-dimensional space of the uncertain data, while in applied regression analysis it is assumed to follow a normal distribution with zero mean value and equal standard deviation.

  Further in this section, approximating functions are considered that have only $M' < M$ uncertain data in their argument list. In this case, the value of Y is indeed a random variable at $\boldsymbol{p}'_{\boldsymbol{n}} = (p_{1,n}, \ldots, p_{M',n}), n = 1, \ldots, N$ due to the variability of the uncertain data that are not in the list (not counting those that are left out due to complete state of knowledge dependence). It may therefore be rightfully considered to call $b_0, b_1, \ldots, b_{M'}$ estimates of regression coefficients. For this reason and as a matter of simplicity, the components of the vector $\boldsymbol{b}$ shall from now on be called estimates of regression coefficients. The regression coefficients, obtained from the sample values (input and output sample of the uncertainty analysis), are influenced by the states of knowledge specified for the uncertain data, as one would expect from uncertainty importance measures. The subjective probability distributions determine the region in the M-dimensional space where most sample values will be concentrated and the slopes of the approximating plane will accordingly change if the state of knowledge changes for some of the uncertain data.

The aim of uncertainty importance analysis is not to exactly represent the computer model result over the value space of the uncertain data but to place emphasis on that range in the value space where the state of knowledge is concentrated. The estimates of the regression coefficients and therefore the slopes of the least squares approximation plane are influenced by the states of knowledge of the uncertain data and consequently also by the resulting state of knowledge of the model result. Unfortunately, they are not yet uncertainty importance measures suitable for ranking. The reasons are:

– The value of the estimate b_m of the regression coefficient is the amount of units by which the linear approximation $\widehat{Y}$ changes per unit change of the uncertain datum P_m, all other uncertain data $P_l, l \neq m$, remaining constant. The sign of b_m indicates the direction of the change of $\widehat{Y}$. Positive sign indicates that P_m and Y tend to change in the same direction (both increase or both decrease), while negative sign means that they tend to change in opposite directions. The value of b_m depends on the units used in measuring the model result and the uncertain datum P_m. For

instance, b_m will be different if P_m is measured in grams instead of kilograms. In the case of grams, the range of units of P_m will be expanded by orders of magnitude, if compared to P_m measured in kg, and the slope along the P_m-axis will therefore be far less pronounced (b_m will be much smaller) than in the case where P_m is measured in kg. If $P_j, j \neq m$, is measured in kg and P_m in grams, then the comparison of b_m to b_j for ranking purposes will be misleading.

– Even if the uncertain data were transformed into unit-less quantities (for instance, by dividing their value by the mean value of their subjective probability distribution), a change of their unit-less value by one could still mean very different changes in terms of the variance of the subjective probability distributions of the unit-less uncertain data.

These shortcomings suggest using the sample values in their standardized form and to compute estimates of the so-called standardized regression coefficients for ranking. To this end, the sample values y_n, $p_{m,\,n}$, $n = 1, \ldots, N$; $m = 1, \ldots, M$ are transformed into

$$\begin{aligned} y_n^* &= (y_n - m_Y)/s_Y \\ p_{m,n}^* &= \left(p_{m,n} - m_m\right)/s_m \end{aligned} \tag{6.20}$$

with

$$\begin{aligned} s_Y^2 &= \frac{1}{N-1}\sum_{n=1}^{N}(y_n - m_Y)^2 \\ s_m^2 &= \frac{1}{N-1}\sum_{n=1}^{N}\left(p_{m,n} - m_m\right)^2 \end{aligned}$$

and s_Y, s_m denoting the sample standard deviations of Y and P_m.

Estimates of standardized least squares linear regression coefficients $\beta_0^*, \beta_1^*, \ldots, \beta_M^*$ of the regression model

$$Y^* = \beta_0^* + \sum\nolimits_{m=1}^{M} \beta_m^* P_m^* + D$$

are then obtained by using the transformed values instead of the original sample values. The condition to be satisfied by the estimates of the standardized regression coefficients is

$$h_s\left(b_0^*, \ldots, b_M^*\right) = \sum\nolimits_{n=1}^{N}\left(y_n^* - b_0^* - \sum\nolimits_{m=1}^{M} b_m^* p_{m,n}^*\right)^2 = \textit{minimal!} \tag{6.21}$$

or

$$\partial h_s / \partial b_0^* = -2\sum\nolimits_{n=1}^{N}\left(y_n^* - b_0^* - \sum\nolimits_{m=1}^{M} b_m^* p_{m,n}^*\right) = 0 \tag{6.22}$$

and for $k = 1, \ldots, M$

$$\partial h_s/\partial b_k^* = -2\sum_{n=1}^{N}\left(y_n^* - b_0^* - \sum_{m=1}^{M} b_m^* p_{m,n}^*\right)p_{k,n}^* = 0 \tag{6.23}$$

From (6.22) follows the condition:

$$\sum_{n=1}^{N}(y_n - m_Y)/s_Y - Nb_0^* - \sum_{n=1}^{N}\sum_{m=1}^{M} b_m^*(p_{m,n} - m_m)/s_m = 0. \tag{6.24}$$

It is satisfied by

$$b_0^* = \left(\frac{1}{N}\right)\left(\sum_{n=1}^{N} y_n - m_Y\right)/s_Y - \sum_{m-1}^{M} b_m^*\left(\frac{1}{N}\right)\left(\sum_{n=1}^{N} p_{m,n} - m_m\right)/s_m$$

and since all sums over sample values vanish, it follows that

$$b_0^* = 0.$$

The same matrix and vector notation as before is used in the computation of $b_1^*, \ldots, b_M^*$ but with

$$\boldsymbol{y_s} = \begin{bmatrix} y_1{}^* \\ y_2{}^* \\ \cdot \\ \cdot \\ \cdot \\ y_N{}^* \end{bmatrix} \quad \boldsymbol{b_s} = \begin{bmatrix} b_1{}^* \\ b_2{}^* \\ \cdot \\ \cdot \\ b_M{}^* \end{bmatrix} \quad \boldsymbol{P_s} = \begin{bmatrix} p_{1,1}{}^* & p_{2,1}{}^* & \cdots & p_{M,1}{}^* \\ p_{1,2}{}^* & p_{2,2}{}^* & \cdots & p_{M,2}{}^* \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ p_{1,N}{}^* & p_{2,N}{}^* & \cdots & p_{M,N}{}^* \end{bmatrix} \tag{6.25}$$

From (6.23), (6.25) and with $b_0^* = 0$ it follows

$$\boldsymbol{P_s}'\boldsymbol{P_s}\boldsymbol{b_s} = \boldsymbol{P_s}'\boldsymbol{y_s} \tag{6.26}$$

which gives in analogy to (6.18)

$$\boldsymbol{b_s} = (\boldsymbol{P_s}'\boldsymbol{P_s})^{-1}\boldsymbol{P_s}'\boldsymbol{y_s}.$$

The components $b_1^*, \ldots b_M^*$ of $\boldsymbol{b_s}$ are called estimates of standardized regression coefficients. Positive sign of b_m^* indicates that P_m and Y tend to change in the same direction (both increase or both decrease), while negative sign means that they tend to change in opposite directions. b_m^* is the amount of standard deviations s_Y by which the linear approximation $\widehat{Y}$ changes per standard deviation change of the uncertain datum P_m, all other uncertain data P_l, $l \neq m$, remaining constant.

This follows from

$$\widehat{Y} = b_0 + \sum_{m=1}^{M} b_m P_m \text{ and } \widehat{Y}_{(k)} = b_0 + \sum_{m=1}^{M} b_m P_m + b_k s_k$$

so that $\widehat{Y}_{(k)} - \widehat{Y} = \Delta\widehat{Y} = b_k s_k = b_k^* s_Y$ *since* $b_k^* = b_k s_k/s_Y$. This relationship between b_k and b_k^* will be shown using (6.16) and the following notation:

$$\boldsymbol{W} = \begin{bmatrix} (p_{1,1} - m_1)(p_{2,1} - m_2) \ldots (p_{M,1} - m_M) \\ (p_{1,2} - m_1)(p_{2,2} - m_2) \ldots (p_{M,2} - m_M) \\ . \\ . \\ . \\ (p_{1,N} - m_1)(p_{2,N} - m_2) \ldots (p_{M,N} - m_M) \end{bmatrix} \quad \boldsymbol{v} = \begin{bmatrix} y_1 - m_y \\ y_2 - m_y \\ . \\ . \\ . \\ y_N - m_y \end{bmatrix} \tag{6.27}$$

$$\boldsymbol{S} = \begin{bmatrix} s_1 0 \ \ldots \ 0 \\ 0 \ s_2 \ 0 \ .. \ 0 \\ . \\ . \\ 0 \ldots 0 \ s_M \end{bmatrix}$$

Condition (6.7) can be written as

$$F(\boldsymbol{b}, \boldsymbol{W}, \boldsymbol{v}) = \sum_{n=1}^{N} \left(v_n - \sum_{m=1}^{M} b_m w_{m,n} \right)^2 = \textit{minimal}! \tag{6.28}$$

For $k = 1, \ldots, M$ the condition

$$\partial F / \partial b_k = -2 \sum_{n=1}^{N} \left(v_n - \sum_{m=1}^{M} b_m w_{m,n} \right) w_{k,n} =^{!} 0 \tag{6.29}$$

or, in matrix and vector notation,

$$(\boldsymbol{v} - \boldsymbol{W}\boldsymbol{b})' \boldsymbol{W} =^{!} 0 \tag{6.30}$$

or

$$\boldsymbol{W}'(\boldsymbol{v} - \boldsymbol{W}\boldsymbol{b}) =^{!} 0 \tag{6.31}$$

leads to

$$\boldsymbol{W}'\boldsymbol{v} = \boldsymbol{W}'\boldsymbol{W}\boldsymbol{b} \tag{6.32}$$

and therefore

$$\boldsymbol{b} = (\boldsymbol{W}'\boldsymbol{W})^{-1} \boldsymbol{W}'\boldsymbol{v}. \tag{6.33}$$

$\boldsymbol{v} = \boldsymbol{W}\boldsymbol{b}$ cannot be used in (6.31) to compute $\boldsymbol{b}$ since $\boldsymbol{W}$ is an $N \times M$ matrix (i.e. the coefficient matrix of N linear equations in $M < N$ unknowns) and therefore $\boldsymbol{W}$ is not invertible.

With

$$\boldsymbol{P}_s = \boldsymbol{W}\boldsymbol{S}^{-1}, \boldsymbol{y}_s = s_Y^{-1}\boldsymbol{v}$$

and following (6.26)

$$\boldsymbol{b}_s = \left(\boldsymbol{S}^{-1}\boldsymbol{W}'\boldsymbol{W}\boldsymbol{S}^{-1}\right)^{-1}\boldsymbol{S}^{-1}\boldsymbol{W}'\left(s_Y^{-1}\boldsymbol{v}\right) \tag{6.34}$$

or

$$\boldsymbol{b}_s = \boldsymbol{S}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\left(s_Y^{-1}\boldsymbol{v}\right). \tag{6.35}$$

Using (6.33), the components of $\boldsymbol{b}_s$ are

$$\boldsymbol{b}_s = \boldsymbol{S}\boldsymbol{b}s_Y^{-1} = \begin{bmatrix} b_1 s_1/s_y \\ b_2 s_2/s_y \\ \cdot \\ \cdot \\ b_M s_M/s_y \end{bmatrix}. \tag{6.36}$$

The estimate b_m^* of the standardized regression coefficient equals the estimate b_m of the regression coefficient times the sample standard deviation of the uncertain datum P_m and divided by the sample standard deviation of the computer model result.

The standardized regression coefficients are suitable for ranking the uncertain data with respect to their contribution to the uncertainty of the least squares linear approximation $\widehat{Y}$. Is this ranking also useful with respect to their uncertainty importance for the model result Y ? To answer this question, one needs to find the fraction of the sample variance s_Y^2 of the model result that is represented by the sample variance of the linear approximation $\widehat{Y}$. This fraction is called "coefficient of determination". It is the square of the multiple correlation coefficient, i.e. the square of the correlation coefficient $\rho\left(Y, \widehat{Y}\right)$ between the model result and its least squares linear approximation.

The so-called empirical coefficient of determination R^2 is the square of the sample value of $\rho\left(Y, \widehat{Y}\right)$. The sample value $r_{Y,\widehat{Y}}$ is computed from the sample values of Y and $\widehat{Y}$ as follows:

$$r_{Y,\hat{Y}} = \left[\sum_{n=1}^{N}\left(\widehat{y}_n - m_{\hat{Y}}\right)(y_n - m_Y)\right] / \left[\sum_{n=1}^{N}\left(\widehat{y}_n - m_{\hat{Y}}\right)^2 \sum_{n=1}^{N}(y_n - m_Y)^2\right]^{1/2}. \tag{6.37}$$

Using (6.12) and the notation (6.27):

$$\sum_{n=1}^{N}\left(\widehat{y}_n - m_{\hat{Y}}\right)(y_n - m_Y) = \sum_{n=1}^{N}\left(\sum_{m=1}^{M} b_m w_{m,n}\right)(y_n - m_Y) = (\boldsymbol{W}\boldsymbol{b})'\boldsymbol{v} \tag{6.38}$$

$$r_{Y,\hat{Y}}{}^2 = (\boldsymbol{W}\boldsymbol{b})'\boldsymbol{v}(\boldsymbol{W}\boldsymbol{b})'\boldsymbol{v} / \boldsymbol{W}\boldsymbol{b}'(\boldsymbol{W}\boldsymbol{b})\boldsymbol{v}'\boldsymbol{v} \tag{6.39}$$

or

$$r_{Y,\hat{Y}}{}^2 = (\boldsymbol{b}'\boldsymbol{W}'\boldsymbol{v})^2 / \boldsymbol{b}'\boldsymbol{W}'\boldsymbol{W}\boldsymbol{b}(\boldsymbol{v}'\boldsymbol{v}).$$

From (6.32) follows

$$r_{Y,\hat{Y}}{}^2 = \frac{(\boldsymbol{b}'\boldsymbol{W}'\boldsymbol{v})^2}{\boldsymbol{b}'\boldsymbol{W}'\boldsymbol{v}(\boldsymbol{v}'\boldsymbol{v})} = \frac{\boldsymbol{b}'\boldsymbol{W}'\boldsymbol{v}}{(\boldsymbol{v}'\boldsymbol{v})} = \frac{\boldsymbol{b}'\boldsymbol{W}'\boldsymbol{W}\boldsymbol{b}}{\boldsymbol{v}'\boldsymbol{v}} \tag{6.40}$$

$$\frac{1}{N-1}\boldsymbol{b}'\boldsymbol{W}'\boldsymbol{W}\boldsymbol{b} = \frac{1}{N-1}\sum\nolimits_{n=1}^{N}\left(\widehat{y}_n - m_Y\right)^2 \tag{6.41}$$

Equation (6.41) uses (6.12) and represents that portion of the sample variance of the model result that is represented by the sample variance of the linear approximation $\widehat{Y}$, while

$$\frac{1}{N-1}\boldsymbol{v}'\boldsymbol{v} = \frac{1}{N-1}\sum\nolimits_{n=1}^{N}(y_n - m_Y)^2$$

is the sample variance of the model result Y.
Since

$$\begin{aligned}\sum_{n=1}^{N}(y_n - m_Y)^2 &= \sum_{n=1}^{N}\left[\left(y_n - \widehat{y}_n\right) + \left(\widehat{y}_n - m_Y\right)\right]^2 \\ &= \sum\nolimits_{n=1}^{N}\left(y_n - \widehat{y}_n\right)^2 - 2\sum\nolimits_{n=1}^{N}\left(y_n - \widehat{y}_n\right)\left(\widehat{y}_n - m_Y\right) + \sum\nolimits_{n=1}^{N}\left(\widehat{y}_n - m_Y\right)^2 \\ &= (\boldsymbol{v} - \boldsymbol{W}\boldsymbol{b})^2 - 2(\boldsymbol{v} - \boldsymbol{W}\boldsymbol{b})'\boldsymbol{W}\boldsymbol{b} + (\boldsymbol{W}\boldsymbol{b})'\boldsymbol{W}\boldsymbol{b}\end{aligned}$$

and with $(\boldsymbol{v} - \boldsymbol{W}\boldsymbol{b})'\boldsymbol{W}$ the null vector of (6.30), it follows that

$$(N-1)s_Y^2 = \sum\nolimits_{n=1}^{N}(y_n - m_Y)^2 = \sum\nolimits_{n=1}^{N}\left(y_n - \widehat{y}_n\right)^2 + \sum\nolimits_{n=1}^{N}\left(\widehat{y}_n - m_Y\right)^2 \tag{6.42}$$

The sums in (6.42) are known in regression analysis as "total sum of squares" (SST), "residual sum of squares" (SSE) and "regression sum of squares" (SSR) so that

$$\mathrm{SST} = \mathrm{SSE} + \mathrm{SSR}.$$

It follows from (6.40)

$$r_{Y,\hat{Y}}{}^2 = \sum\nolimits_{n=1}^{N}\left(\widehat{y}_n - m_Y\right)^2 / \sum\nolimits_{n=1}^{N}(y_n - m_Y)^2 = \mathrm{SSR}/\mathrm{SST}. \tag{6.43}$$

R^2 is the fraction of the sample variance of the model result Y that is represented by the sample variance of the least squares linear approximation $\widehat{Y}$.
The closer the value of R^2 is to 1, the smaller is SSE/SST, i.e. the variability of the y_n about the regression hyperplane, relative to their total variability, since

$$\mathrm{SSR}/\mathrm{SST} = 1 - \mathrm{SSE}/\mathrm{SST} \tag{6.44}$$

and SSE measures the variability that is not explained by the linear least squares approximation in the uncertain data.

If Y is a linear function in the uncertain data P_m ($m = 1, \ldots, M$) only and all P_m are in the regression model, then $R^2 = 1$ since SSE = 0.

The value of R^2 will be equal to 1 even in the case of a non-linear function $Y = f(P_1, \ldots, P_M)$ if the sample size N is equal to $M + 1$ and M is the number of uncertain data in the linear least squares approximation. The sample values of the model result are reproduced exactly in this case because the linear least squares approximation is over-fitting the sample data.

The ranking, obtained from the estimates b_m^* of the standardized regression coefficients, refers to the fraction R^2 of the sample variance of the model result Y. $R^2 < 0.5$ means that the ranking of the uncertain data refers to less than half of the variance of the sample values y_n and may therefore not be suitable to indicate where the main contributions to the uncertainty of the model result come from.

The variance of the estimate b_m of the regression coefficient β_m stems from the variance of the difference between Y and its least squares linear approximation and is computed as (Heinhold and Gaede 1968; Freund and Minton 1979; Draper and Smith 1998):

$$S_{b_m}^2 = \left(\frac{1}{1 - R_m^2}\right)\left(\frac{1 - R^2}{N - M - 1}\right)\left(\frac{s_Y^2}{s_m^2}\right). \tag{6.45}$$

${R_m}^2$ is the empirical coefficient of determination of a least squares linear approximation for P_m in the other ($M - 1$) uncertain data and $1/(1 - {R_m}^2)$ is known as the "variance inflation factor". Obviously, strong sample correlations of P_m with other uncertain data (no matter whether intended or spurious) increase the variance inflation factor as the value of R_m will be close to 1 in this case. R^2 is the empirical coefficient of determination of the least squares linear approximation of Y in the M uncertain data. Clearly, the more N exceeds M the smaller the variance of the estimate of the regression coefficient. Following (6.36), the variance of the estimate of the standardized regression coefficient is then given as

$$S_{b_m^*}^2 = S_{b_m}^2 s_m^2 / s_Y^2.$$

6.3.1.2 Uncertainty Importance Measures from Correlation

Could the sample correlations $r_{Y,m}$, $m = 1, \ldots, M$, be used for uncertainty importance ranking? The sample correlation coefficient $r_{Y,m}$ of Y and P_m is computed as

$$r_{Y,m} = \left[\sum_{n=1}^{N} (y_n - m_Y)(p_{m,n} - m_m)\right] / \left[\sum_{n=1}^{N} (y_n - m_Y)^2 \sum_{n=1}^{N} (p_{m,n} - m_m)^2\right]^{1/2} \tag{6.46}$$

while the sample correlation coefficient between the uncertain data P_m and P_l is

$$r_{m,l} = \left[\sum_{n=1}^{N} \left(p_{m,n} - m_m\right)\left(p_{l,n} - m_l\right)\right] / \left[\sum_{n=1}^{N} \left(p_{m,n} - m_m\right)^2 \sum_{n=1}^{N} \left(p_{l,n} - m_l\right)^2\right]^{1/2}. \tag{6.47}$$

The correlation coefficient $r_{Y,m}$ measures the extent of linear relationship between the sample values of the model result Y and those of the uncertain datum P_m.

However, using (6.20) and the matrix/vector notation (6.25) together with (6.26)

$$\boldsymbol{P_s}'\boldsymbol{P_s} = \boldsymbol{R} = \begin{bmatrix} 1 & r_{1,2} & & \ldots r_{1,M} \\ r_{2,1} & 1 & r_{2,3} & \ldots r_{2,M} \\ \cdot & & & \\ \cdot & & & \\ r_{M,1} r_{M,2} & & \ldots & 1 \end{bmatrix} \quad \boldsymbol{P_s}'\boldsymbol{y_s} = \boldsymbol{P_s}'\boldsymbol{P_s}\boldsymbol{b_s} = \boldsymbol{r_Y} = \begin{bmatrix} r_{Y,1} \\ r_{Y,2} \\ \cdot \\ \cdot \\ r_{Y,M} \end{bmatrix}$$

and therefore

$$r_{Y,m} = \sum_{l=1}^{M} r_{m,l} b_l^*, \quad m = 1, \quad \ldots, M. \tag{6.48}$$

The sample correlation coefficient $r_{Y,\,m}$ is not only an estimate of the amount of standard deviation changes s_Y of $\widehat{Y}$ due to one standard deviation change s_m of P_m, all P_l, $l \neq m$, remaining constant. It also includes estimates of the amounts of standard deviation changes s_Y of $\widehat{Y}$ that are due to the fractions $r_{m,\,l}$ of standard deviation changes of those uncertain data P_l, $l \neq m$, which are correlated with P_m and are therefore affected by the standard deviation change of P_m. The sign of $r_{Y,\,m}$ indicates the direction of the change of $\widehat{Y}$ (positive: same direction as the change of P_m; negative: opposite direction).

In the case of zero correlation (specified and spurious) between the sample values $(p_{l,n}, p_{m,n})$, $n = 1, \ldots, N$; $m = 1, \ldots, M$, $l \neq m$

$$r_{Y,\hat{Y}}{}^2 = \sum_{m=1}^{M} b_m^{*2} \tag{6.49}$$

and following (6.48)

$$r_{Y,\hat{Y}}{}^2 = \sum_{m=1}^{M} r_{Y,m}^2. \tag{6.50}$$

(6.49) and (6.50) follow from (6.40) and (6.36) since, in the case of zero correlation, the off-diagonal elements of $\boldsymbol{P_s}'\boldsymbol{P_s}$ and of $\boldsymbol{W}'\boldsymbol{W}$ are zero and therefore

$$\boldsymbol{b_s'}\boldsymbol{b_s} = \boldsymbol{b}'\boldsymbol{S}'\boldsymbol{S}\boldsymbol{b}/s_y{}^2 = \boldsymbol{b}'\boldsymbol{W}'\boldsymbol{W}\boldsymbol{b}/\boldsymbol{v}'\boldsymbol{v} = (\boldsymbol{W}\boldsymbol{b})'\boldsymbol{W}\boldsymbol{b}/\boldsymbol{v}'\boldsymbol{v} = r_{Y,\hat{Y}}{}^2$$
$$\boldsymbol{W}'\boldsymbol{W} = (N-1)\ \boldsymbol{S}^2, \boldsymbol{v}'\boldsymbol{v} = (N-1)s_y{}^2 \ \text{ and } \ \sum_{m=1}^{M} b_m^{*2} = \boldsymbol{b_s'}\boldsymbol{b_s}.$$

The matrix $\boldsymbol{R} = \boldsymbol{P_s}'\boldsymbol{P_s}$ is the sample correlation matrix of the uncertain data. Its diagonal elements $r_{m,\,m}$ have the value 1, and its off-diagonal elements $r_{m,l} = r_{l,m}$, $m \neq l$, $m,l = 1, \ldots, M$ are the sample correlation coefficients of the data pairs (P_m, P_l). If all off-diagonal elements are zero, then $\boldsymbol{R} = \boldsymbol{I}$, the unit matrix, and the estimates b_m^* of the

standardized regression coefficients are equal to the sample correlation coefficients $r_{Y,m}$ since

$$\boldsymbol{R}\boldsymbol{b}_s = \boldsymbol{r}_Y. \tag{6.51}$$

However, if $r_{m,l} \neq 0$ for some indices m and l, then (6.48) shows the relationship between $r_{Y,m}$ and the involved sample correlation coefficients $r_{m,l} \neq 0$. The estimates b_m^*, $m = 1, \ldots, M$ are obtained from

$$\boldsymbol{b}_s = (\boldsymbol{P}_s{}'\boldsymbol{P}_s)^{-1}\boldsymbol{P}_s{}'\boldsymbol{y}_s = \boldsymbol{R}^{-1}\boldsymbol{r}_Y \tag{6.52}$$

A non-zero sample correlation coefficient $r_{m,\,l}$ may be intended, i.e. due to state of knowledge dependence specified between the uncertain data P_m and P_l. It may, however, also be non-zero due to the chance mechanism of the sampling process. In this case, the sample correlation between P_m and P_l is unintended and is called "spurious correlation".

Spurious correlation "contaminates" the value $r_{Y,m}$ as can be concluded from (6.48). Even if b_m^* is zero, $r_{Y,m}$ may be non-zero either justifiably so because of state of knowledge dependence between uncertain data or unjustifiably since contaminated by spurious correlation among the sample values of the uncertain data. The closer M is to N and in particular the more M exceeds N, the larger the spurious correlations will be in magnitude and in number. For instance, if the states of knowledge of P_m and P_l are independent and quantified by a normal probability distribution, then the empirical correlation coefficient $R_{m,l}$ is distributed according to Rosner (1995):

$$R_{m,l} \sim T_{N-2}/\left(N - 2 + T_{N-2}{}^2\right)^{1/2}$$

$T_{N\,-\,2}$ is Student distributed with degree of freedom $N - 2$.

The probability to have a spurious correlation $r_{m,l} > 0.5$ or $r_{m,l} <$ - 0.5 between at least two of M state of knowledge independent uncertain data is the larger the smaller N and the larger M is. Table 6.1 shows the entries of the matrix $\boldsymbol{P}_s{}'\boldsymbol{P}_s$ for a small number M of state of knowledge independent uncertain data and for a small sample size N (Tables 6.2, 6.3, 6.4, and 6.5).

In situations where the state of knowledge expressions are not standard normal distributions, the sample sizes in the tables may be taken as an indication of those required so that spurious correlations larger or equal to r_{lim} occur with sufficiently low probability.

Spurious correlations as well as unacceptable differences between the specified correlations and actual sample correlations can, however, be eliminated before evaluation of the model for each set of values in the input file of the uncertain data. This may be achieved by an iterative application of the method mentioned in Sect. 4.4.1.14. The resulting matrix has the same column entries as matrix $\boldsymbol{P}$, but they are permuted such that the sample correlations are as specified. Unfortunately, the input sample and consequently also the output sample of the uncertainty analysis are then no longer simple random samples. The confidence statements derived for

Table 6.1 Sample correlation coefficients of a sample of size $N = 100$ with $M = 6$ and no state of knowledge dependence specified

	1	2	3	4	5	6
1	1.0000E+00	−3.2249E-02	4.3691E-02	4.4895E-02	2.6887E-02	−1.7620E-01
2	−3.2249E-02	1.0000E+00	1.2624E-01	−1.7446E-02	−1.2475E-01	1.1281E-01
3	4.3691E-02	1.2624E-01	1.0000E+00	4.5395E-02	−6.8870E-02	2.1102E-01
4	4.4895E-02	−1.7446E-02	4.5395E-02	1.0000E+00	−7.3201E-02	4.4130E-02
5	2.6887E-02	−1.2475E-01	−6.8870E-02	−7.3201E-02	1.0000E+00	−1.1567E-01
6	−1.7620E-01	1.1281E-01	2.1102E-01	4.4130E-02	−1.1567E-01	1.0000E+00

All off-diagonal values are spurious correlations with $r_{6,3} = r_{3,6}$ the largest

Table 6.2 Minimum sample size N required so that the probability is at least 0.9 for the maximum sample correlation value $r_{Y,m}$ to be less than r_{lim}

	M							
r_{lim}	10	20	30	40	50	60	90	120
0.2	165	195	215	220	230	245	255	265
0.3	75	90	90	100	105	105	115	120

The states of knowledge of Y and of the M uncertain data are expressed by standard normal distributions and are specified as independent (i.e. Y is not a function of any of the uncertain data)

Table 6.3 Minimum sample size N required so that the probability is at least 0.95 for the maximum sample correlation value $r_{Y,m}$ to be less than r_{lim}

	M							
r_{lim}	10	20	30	40	50	60	90	120
0.2	195	220	240	255	260	280	295	305
0.3	85	100	110	115	120	120	125	135

The states of knowledge of Y and of the M uncertain data are expressed by standard normal distributions and are specified as independent (i.e. Y is not a function of any of the M uncertain data)

Table 6.4 Minimum sample size N required so that the probability is at least 0.9 for the maximum sample correlation value $r_{m,l}$ to be less than r_{lim}

	M							
r_{lim}	10	20	30	40	50	60	90	120
0.2	225	295	345	355	385	410	445	460
0.3	105	130	150	165	165	175	200	200
0.4	60	75	80	85	90	100	105	110

The states of knowledge of the M uncertain data are expressed by standard normal distributions and are specified as independent

Table 6.5 Minimum sample size N required so that the probability is at least 0.95 for the maximum sample correlation value $r_{m,l}$ to be less than r_{lim}

	M							
r_{lim}	10	20	30	40	50	60	90	120
0.2	255	330	370	390	425	435	470	490
0.3	115	145	160	175	185	195	205	215
0.4	65	80	90	95	100	105	110	115

The states of knowledge of the M uncertain data are expressed by standard normal distributions and are specified as independent

computer model results in Chap. 5 require that the output sample be a simple random sample.

The partial correlation coefficient (Freund and Minton 1979) is not suggested as an uncertainty importance measure. The square of the estimate of the partial correlation coefficient of Y and P_m is the square of the correlation coefficient between the residuals of the least squares linear regressions of Y and P_m on the P_k, $k = 1, \ldots, M$, $k \neq m$. The reason why the partial correlation coefficient is not suggested for uncertainty importance ranking is best understood in the case where Y is a linear function of state of knowledge independent uncertain data. In this case, all partial correlation coefficients have unity as their absolute value although uncertainty contributions to Y can be substantially different. If that part of the variability of the $p_{m,n}$, which is not explained by the $p_{l,n}$, $l \neq m$, explains all the variability of the residuals of a linear regression model constructed from the y_n and the $p_{l,n}$ $l \neq m$, then the absolute value of the estimated partial correlation coefficient for P_m will be at its maximum (unity) although the variability of the residuals may only be a minute fraction of the total variability of the y_n and other individual parameters may explain a much larger fraction. This shortcoming of the partial correlation coefficient is also observed in Janssen (1994).

The question remains whether one should use the sample correlation coefficients $r_{Y,\,m}$ or the estimates b_m^*, $m = 1, \ldots, M$ of the standardized regression coefficients for ranking. The $r_{Y,m}$ are much easier to obtain as their computation requires only the product $\boldsymbol{P_s}'\boldsymbol{y_s}$, and they are available even if $M > N - 1$. The computation of the b_m^* necessitates the inversion of $\boldsymbol{P_s}'\boldsymbol{P_s}$, and a unique solution to $\boldsymbol{P_s}'\boldsymbol{P_s}\boldsymbol{b_s} = \boldsymbol{P_s}'\boldsymbol{y_s}$ is only available for $M \leq N - 1$.

The decision between the two measures for ranking will depend on whether $N > M$ can be afforded and whether the uncertain data P_m are to be ranked by the uncertainty contribution of P_m alone, all P_l, $l \neq m$, remaining constant, or whether those uncertainty contributions that are due to state of knowledge dependence between P_m and any of the other uncertain data should be included. In the latter case, it should be noted that even if the model result Y does not functionally depend on P_m (P_m is not an argument of the function that provides the value of Y), $r_{Y,m}$ may be non-zero. It may assign a significant rank to P_m due to non-zero correlation coefficients $r_{m,l}$ brought about by state of knowledge dependence between P_m and any of the other uncertain data P_l,$l \neq m$, and/or by spurious correlations as (6.48) shows.

If spurious correlations were negligible, and an uncertainty importance measure for P_m that includes the effect of those uncertain data that are state of knowledge dependent with P_m were sufficient, then the sample correlation coefficient $r_{Y,m}$ would certainly be a computationally cheap option. However, in practical situations spurious correlations are rarely negligible, and therefore, the standardized regression coefficients will need to be used as uncertainty importance measures not forgetting to also compute their variance inflation factors [the first quotient in (6.45)].

6.3.1.3 Uncertainty Importance Measures from Stepwise Regression Analysis

But what to do, if M is so large that sample sizes N, which keep spurious correlations and variance inflation factors small, are simply unaffordable? The solution is to perform the regression in a stepwise fashion following, for instance, a so-called forward selection method. In this method, least squares linear regression starts with a model in only one uncertain datum, namely the P_m with the largest coefficient of determination $R^2_{(1)}$ (see (6.43)) for

$$\widehat{Y}_{(1)} = b_0^{(1)} + b_m^{(1)} P_m.$$

The index in brackets indicates the number of uncertain data included in the least squares linear approximation. The next step selects the $P_l, l \neq m$, with the largest difference $\Delta R^2_{(1,2)} = R^2_{(2)} - R^2_{(1)}$ among the $M - 1$ uncertain data left. $R^2_{(2)}$ is the coefficient of determination for

$$\widehat{Y}_{(2)} = b_0^{(2)} + b_m^{(2)} P_m + b_l^{(2)} P_l$$

and so forth until after j steps the maximum achievable $\Delta R^2_{(j,j+1)}$ is below a chosen threshold value.

The $\Delta R^2_{(k,k+1)}$ values, $k = 1, \ldots, j-1$, might be used as measures of uncertainty importance for the j uncertain data included in the model although they do not tell the direction of their influence on the uncertainty of the model result. The latter is given by the signs of the corresponding j standardized regression coefficients. (Near) multicollinearity, i.e. its influence on the variance of the estimates b_m^*, is less of a problem in stepwise regression. Extending the matrix $\boldsymbol{P_k}$ (k+1 columns of matrix $\boldsymbol{P}$) by a column that is (nearly) a linear combination of some of the columns already in the matrix will often not provide a sufficiently large $\Delta R^2_{(k,k+1)}$ to also include the corresponding parameter into the least squares linear approximation.

6.3.1.4 The Correlation Ratio as Uncertainty Importance Measure

The ranking derived from standardized regression coefficients will not be satisfactory if the R^2 value of the least squares linear approximation is too low, i.e. $R^2 \leq 0.5$. The reason for the low R^2 value could be that:

– Some uncertain data represent model uncertainties that are major contributors to the uncertainty of the model result. The data values used in the uncertainty analysis are indices, arbitrarily assigned to alternative model formulations. Correlation between the sample index numbers and the sample values of the model result Y will most likely not be suitable to measure the uncertainty importance nor will the estimates of standardized regression coefficients for these uncertain data.

The R^2 value will generally be rather low in this case, thereby indicating that a significant fraction of the sample variance of Y might be due to model uncertainties but is not represented by the linear least squares approximation $\widehat{Y}$.

Among those important model uncertainties may be uncertain input functions over interim model results or over independent variables like time and/or space (see Sect. 3.3.4). A sample of possibly true input functions may have been used in the analysis, and their index values will be used as the values of a substitute uncertain datum in the sampling process and uncertainty importance analysis.

The correlation ratio discussed below could be used for ranking in these cases provided the sample size is sufficiently smaller (for instance, $\leq N^{1/2}$) than the sample size N of the uncertainty analysis, so that several model runs use the same possibly true model formulation or input function.

– Some of the uncertainties may be the output from a model (feeder model) that provides input to the computer model application. This output may consist of a mixture of single numbers (scalar values) and sequences, arrays or tables of values. An uncertainty analysis of the feeder model application by Monte Carlo simulation provides a state of knowledge expression for these results in the form of a multivariate sample, drawn according to their joint subjective probability distribution (see Sect. 4.4.1.13 for the treatment of this sample in the uncertainty analysis of the computer model application). Obviously, standardized regression coefficients are not a suitable importance measure for those uncertainties that are mixtures of sequences, arrays or tables of values. Again, the correlation ratio discussed below could be used for ranking in this case provided the sample size of the feeder model analysis is sufficiently smaller (for instance $\leq N^{1/2}$) than the sample size N of the uncertainty analysis of the computer model result so that several computer model runs use the same model output set from the feeder model as input. However, uncertainty importance measures computed for the input from the feeder model would not permit to differentiate between the uncertainty importance of individual sequences, arrays or tables that are part of the mixture. If the substitute uncertain datum, used for the input from the feeder model, is found to be of sufficient importance, the ranking will have to be redone, this time, however, based on the sample of the underlying uncertain data of the feeder model in order to decide about state of knowledge improvements. The sample values of the uncertain data (input sample) of the feeder model may be combined with those of the uncertainty analysis of the computer model and importance measures may be computed for the combined set of uncertain data.

A low R^2 value may, however, also be due to the fact that an approximation $\widehat{Y}$, linear in the uncertain data, is not capable of explaining the contributions to the sample variance of Y from strongly non-linear (particularly non-monotone) relationships between Y and some of the uncertain data. Hora (2003) suggests a method for the detection of non-monotone relationships.

The measure to be computed in these situations is the so-called correlation ratio (Iman and Hora 1990; McKay et al. 1992; Manteufel 1996; McKay 1997; Kruskal 1958; Kendall and Stuart 1973). It is not meant to substitute the correlation

coefficients or the standardized regression coefficients but to enable a correction of their uncertainty importance ranking particularly in those cases where model uncertainty is represented by uncertain data with indices of alternative model formulations as data values. A scatter plot of the model result versus the uncertain datum should be investigated wherever the correlation ratio assigns a significantly higher rank than the correlation coefficient or the standardized regression coefficient. This may help to explain the difference.

The correlation ratio (Kruskal 1958; Kendall and Stuart 1973; Iman and Hora 1990) makes use of the approximation

$$\widehat{Y} = \mathrm{E}\{Y|P_m\} = \mathrm{h}_\mathrm{Y}(P_m) \tag{6.53}$$

of the computer model result Y. $\mathrm{E}\{Y|P_m\}$ is the conditional mean value of Y, i.e. conditioned on the value of P_m. In the least squares sense, $\mathrm{h}_\mathrm{Y}(P_m)$ is the best approximation of Y by a function of P_m alone, i.e. any value different to $\mathrm{h}_\mathrm{Y}(P_m)$ would give a larger squared error sum (Steiner's Theorem; see Heinhold and Gaede 1968).

The squared value of the correlation ratio δ_m is obtained as

$$\delta_m^2 = \frac{\mathrm{Var}\{\mathrm{E}\{Y|P_m\}\}}{\mathrm{Var}\{Y\}} = 1 - \frac{\mathrm{E}\{\mathrm{Var}\{Y|P_m\}\}}{\mathrm{Var}\{Y\}}. \tag{6.54}$$

from the relationship

$$\mathrm{Var}\{Y\} = \mathrm{E}\{\mathrm{Var}\{Y|P_m\}\} + Var\{\mathrm{E}\{Y|P_m\}\} \tag{6.55}$$

which applies since

$$\begin{aligned}
\mathrm{Var}\{\mathrm{E}\{Y|P_m\}\} &= \mathrm{E}\Big\{(\mathrm{E}\{Y|P_m\})^2\Big\} - \mathrm{E}^2\{\mathrm{E}\{Y|P_m\}\} = \mathrm{E}\big\{\mathrm{E}^2\{Y|P_m\}\big\} - \mathrm{E}^2\{Y\} \\
\mathrm{E}\{\mathrm{Var}\{Y|P_m\}\} &= \mathrm{E}\Big\{\mathrm{E}\big\{(Y|P_m)^2\big\} - \mathrm{E}^2\{Y|P_m\}\Big\} = \mathrm{E}\Big\{\mathrm{E}\big\{(Y|P_m)^2\big\}\Big\} - \mathrm{E}\big\{\mathrm{E}^2\{Y|P_m\}\big\} \\
&= \mathrm{E}\big\{Y^2\big\} - \mathrm{E}\big\{\mathrm{E}^2\{Y|P_m\}\big\}
\end{aligned}$$

and therefore

$$\mathrm{E}\{\mathrm{Var}\{Y|P_m\}\} + \mathrm{Var}\{\mathrm{E}\{Y|P_m\}\} = \mathrm{E}\big\{Y^2\big\} - \mathrm{E}^2\{Y\} = \mathrm{Var}\{Y\}.$$

From (6.54) and (6.55) follows that the smallest value that can be assumed by δ_m is zero and its largest value is 1. Therefore, δ_m cannot carry directional information (i.e. whether Y tends to react in the same or in the opposite direction of a change of P_m). Direction would not make sense anyway in the cases discussed above, namely non-monotone relationships between Y and P_m or indices of model alternatives as values of P_m.

$\mathrm{E}\{\mathrm{Var}\{Y|P_m\}\}$ is a measure of the variability of Y due to uncertain data other than P_m. It is zero if Y is a function of P_m alone.

$\mathrm{Var}\{\mathrm{E}\{Y|P_m\}\}$ is the variance of the mean values of Y conditioned on the values of P_m. It is zero if Y is independent of P_m.

The square of the correlation ratio measures the fraction of the variance of the model result Y that is explained by the variance of the mean values of Y conditioned on P_m. Obviously, the correlation ratio will be influenced by contributions to $\mathrm{Var}\{\mathrm{E}\{Y|P_m\}\}$ that are due to state of knowledge dependence of P_m and any P_l, $l \neq m$. This can be seen from the following simple example:

$$
\begin{aligned}
&Y = a_1P_1 + a_2P_2\\
&\mathrm{Var}\{Y\} = a_1^2\mathrm{Var}\{P_1\} + a_2^2\mathrm{Var}\{P_2\} + 2a_1a_2\mathrm{Cov}\{P_1,P_2\}\\
&\mathrm{E}\{Y|P_1\} = a_1P_1 + a_2\mathrm{E}\{P_2|P_1\}\\
&\mathrm{E}\{\mathrm{E}\{Y|P_1\}\} = a_1\mathrm{E}\{P_1\} + a_2\mathrm{E}\{P_2\} = \mathrm{E}\{Y\}\\
&\mathrm{Var}\{\mathrm{E}\{Y|P_1\}\} = \mathrm{Var}\{a_1P_1 + a_2\mathrm{E}\{P_2|P_1\}\}\\
&= \mathrm{E}\Big\{(a_1P_1 + a_2\mathrm{E}\{P_2|P_1\} - \mathrm{E}\{Y\})^2\Big\}\\
&= \mathrm{E}\Big\{(a_1(P_1 - \mathrm{E}\{P_1\}) + a_2(\mathrm{E}\{P_2|P_1\} - \mathrm{E}\{P_2\}))^2\Big\}\\
&= \mathrm{E}\big\{a_1^2(P_1 - \mathrm{E}\{P_1\})^2 + a_2^2(\mathrm{E}\{P_2|P_1\} - \mathrm{E}\{P_2\})^2\\
&\quad + 2a_1a_2(P_1 - \mathrm{E}\{P_1\})(\mathrm{E}\{P_2|P_1\} - \mathrm{E}\{P_2\})\big\}
\end{aligned}
$$

and since

$$
\begin{aligned}
&\mathrm{E}\{\mathrm{E}\{P_2|P_1\}P_1\} = \mathrm{E}\{P_2P_1\} = \mathrm{E}\{P_2\}\mathrm{E}\{P_1\} + \mathrm{Cov}\{P_1,P_2\}\\
&\mathrm{Var}\{\mathrm{E}\{Y|P_1\}\} = a_1^2\mathrm{Var}\{P_1\} + a_2^2\mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\} + 2a_1a_2\mathrm{Cov}\{P_1,P_2\}
\end{aligned}
\tag{6.56}
$$

$$
\delta_1^2 = \big[a_1^2\mathrm{Var}\{P_1\} + a_2^2\mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\} + 2a_1a_2\mathrm{Cov}\{P_1,P_2\}\big] / \big[a_1^2\mathrm{Var}\{P_1\} + a_2^2\mathrm{Var}\{P_2\} + 2a_1a_2\mathrm{Cov}\{P_1,P_2\}\big]
$$

Y is functionally independent of P_1 if $a_1 = 0$. In this case, the second terms of the denominator and of the numerator remain with $\mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\} \neq 0$ due to state of knowledge dependence. In the case of $a_1 \neq 0$ and no state of knowledge dependence between P_1 and P_2, the first term of the numerator and the first and second terms of the denominator remain.

A sample value of the correlation ratio for Y with respect to P_m is obtained from the approximation

$$
\begin{aligned}
&\mathrm{h_Y}(p_m) = \frac{1}{v}\sum\nolimits_{n=1}^{N} w_n y_n\\
&w_n = 0 \quad \textit{if} \;\; p_{m,n} \neq p_m\\
&w_n = 1 \quad \text{otherwise}
\end{aligned}
\tag{6.57}
$$

and v is the number of indices n with $w_n = 1$.

In other words, $\mathrm{h_Y}(p_m)$ is the average over all y_n to the same value of P_m in the sample (see Fig. 6.2).

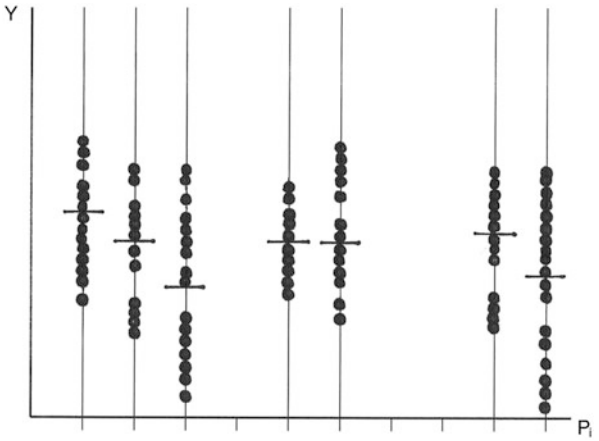


Fig. 6.2 Model results to values of the uncertain datum P_i with discrete subjective probability distribution; the horizontal bars indicate the conditional mean values

In the context of uncertainty analysis by Monte Carlo simulation, all values $p_{m,\,n}$ are random samples. The probability for two or more identical sample values of P_m is zero if the state of knowledge of P_m is expressed not by a discrete subjective probability distribution as in Fig. 6.2 but by a continuous one as in Fig. 6.3. Figure 6.2 would be representative of model uncertainties, quantified by a small set of alternative model formulations, or of multivariate samples used as joint state of knowledge expression (see Sect. 3.5.4). In other words, it illustrates the situation for uncertain data with only a small set of possibly applicable values.

To obtain sample values of the approximate correlation ratio in the continuous case, the set $\{(p_{m,n}),\ n = 1, \ldots, N\}$ is ordered by increasing magnitude and the ordered set is divided into L non-overlapping classes with equally many (K) values in each class $(N = LK)$. The mean values are taken over each class, and $h_Y(p_{m,\,n}) = E\{Y \mid P_m = p_{m,n}\}$ is approximated by

$$h_Y\left(p_{m,n}\right) = \bar{y}_l \tag{6.58}$$

for each sample value $p_{m,n}$ from class l

with $\bar{y}_l = \left(\frac{1}{K}\right)\sum_{class\ l} y_n$.

$s_Y^2 = \left(\frac{1}{N-1}\right)\sum_{l=1}^{L}\sum_{k=1}^{K}\left(y_{l,k} - \bar{y}\right)^2$ corresponds to $[1/(N - 1)]$ SST (estimate of $\mathrm{Var}\{Y\}$)

$\left(\frac{1}{N-1}\right)\sum_{l=1}^{L} K\left(\bar{y}_l - \bar{y}\right)^2$ corresponds to $[1/(N - 1)]$ SSR (estimate of $Var\{\mathrm{E}\{Y \mid P_m\}\}$)

$\left(\frac{1}{L}\right)\sum_{l=1}^{L}\left(\frac{1}{K-1}\right)\sum_{k=1}^{K}\left(y_{l,k} - \bar{y}_l\right)^2$ corresponds to $[1/(N - L)]$ SSE (estimate of $\mathrm{E}\{\mathrm{Var}\{Y \mid P_m\}\}$).

It follows for the square value of the sample value d_m of the correlation ratio δ_m

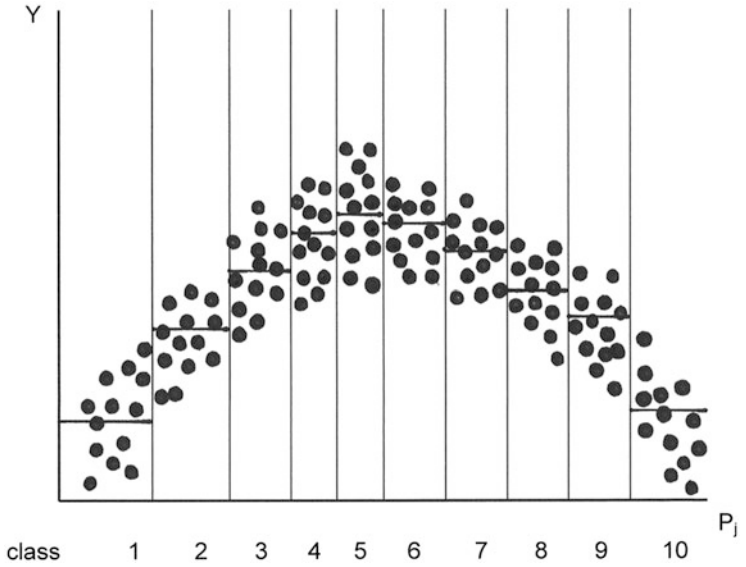


Fig. 6.3 Model results to 10 classes of values of the uncertain datum P_i with continuous subjective probability distribution; the horizontal bars indicate the conditional mean values

$$d_m^2 = \sum_{l=1}^{L} K\left(\bar{y}_l - \bar{y}\right)^2 / \left[\sum_{l=1}^{L} \sum_{k=1}^{K} \left(y_{l.k} - \bar{y}\right)^2\right] \tag{6.59}$$

where L classes of K sample values each are assumed ($N = LK$) and $y_{l,k}$ is the kth sample value of the computer model result Y that is in class l of P_m. As a rule of thumb, the square root of N may be chosen for the number L of classes.

The approximation

$h_Y\left(p_{m,n}\right) = \bar{y}_l$, for each sample value $p_{m,n}$ from class $l, l = 1, \ldots, L$

uses only the class averages of Y. Therefore, it cannot explain the intraclass variability.

Indexing of model alternatives (i.e. their representation by an uncertain datum with the model indices as possibly true values) is and can be rarely done with respect to their influence on the results of the computer model. Regression of Y on the index set would therefore not make sense. Computer models usually provide a set of J model results ($Y_1, \ldots, Y_J$), and the influence of a model alternative may be very different for each of the J different model results and may even change over an independent variable (like time or space) of the model result Y_j. The approximate correlation ratio uses the variability of the class means $\bar{y}_l$, $l = 1, \ldots, L$. Numbering

and sequential order of the classes is therefore irrelevant. This feature makes the correlation ratio the only uncertainty importance measure (of those discussed) suitable for model uncertainties quantified by a set of $L > 2$ alternatives.

Examples, which demonstrate the capabilities of the correlation ratio as uncertainty importance measure, are provided in Sect. 6.3.3. Some limitations and their remediation are presented in Ratto et al. (2009).

6.3.1.5 Variance Decomposition

Variance decomposition provides uncertainty importance measures as fractions of the total variance of the model result $Y = \mathrm{h}(P_1, \ldots, P_M)$. In what follows, $\boldsymbol{P}$ denotes the vector of uncertain data $P_1, \ldots, P_M$ and $\boldsymbol{p}$ a realization thereof. A complete decomposition of $\mathrm{Var}\{Y\}$ (Sobol 1993; Saltelli et al. 1999, 2010; Jansen 1999; Saltelli 2002) is achieved by defining the following terms:

$$
\begin{aligned}
h_0 &:= \mathrm{E}\{Y\} \\
h_i &:= \mathrm{E}\{Y|P_i\} - h_0 && i = 1, \ldots, M \\
h_{i,j} &:= \mathrm{E}\{Y|P_i, P_j\} - h_0 - h_i - h_j && i = 1, \ldots, M; j > i \\
h_{i,j,k} &:= \mathrm{E}\{Y|P_i, P_j, P_k\} - h_0 - h_i - h_j - h_k - h_{i,j} - h_{i,k} - h_{j,k} && i = 1, \ldots, M; j > i; k > j \\
&\vdots \\
h_{1,\ldots,M} &:= \mathrm{E}\{Y|P_1, \ldots, P_M\} - h_0 - \sum_{i=1}^{M} h_i - \sum_{i=1}^{M} \sum_{j>i}^{M} h_{i,j} - \sum_{i=1}^{M} \sum_{j>i}^{M} \sum_{k>j}^{M} h_{i,j,k} - \cdots - \sum_{i=1}^{M} h_{-i}.
\end{aligned}
\tag{6.60}
$$

The index "$-i$" indicates all indices except index i.

$$
\begin{aligned}
\mathrm{h}(P_1, \ldots, P_M) = h_0 &+ \sum_{i=1}^{M} h_i + \sum_{i=1}^{M} \sum_{j>i}^{M} h_{i,j} + \sum_{i=1}^{M} \sum_{j>i}^{M} \sum_{k>j}^{M} h_{i,j,k} + \cdots \\
&+ \sum_{i=1}^{M} h_{-i} + h_{1,\ldots,M}
\end{aligned}
\tag{6.61}
$$

since all that remains from adding up the right-hand sides of the definitions in (6.60) is $\mathrm{E}\{Y | P_1, \ldots, P_M\}$ which equals $\mathrm{h}(\boldsymbol{P})$.

(6.60) defines a decomposition of $\mathrm{h}(\boldsymbol{P})$ into a constant h_0, $\mathrm{C_M}^1$ terms (h_i) quantifying the difference between the local least squares approximation of Y at each value of P_i and the mean value of Y, $\mathrm{C_M}^2$ terms ($h_{i,j}$) quantifying the difference

between the local least squares approximation of Y at each pair of values of P_i and P_j and the mean value of Y minus h_i and h_j and so forth. C_M^k is the number of all combinations of M elements into groups of k different elements. Following the definition of the conditional mean values, the integrals of each of the expressions $h_i, h_{i,j}, \ldots, h_{1,\ldots,M}$ of (6.60), multiplied by the joint density function of the uncertain data in the conditions, equal zero.

Squaring both sides of (6.61), subtracting h_0^2 on both sides, multiplying each term of the squared right-hand side by the joint density function of the uncertain data in the conditions and integrating with respect to these data equals the squared term on the left-hand side multiplied by the joint density function $f(\boldsymbol{p})$ and integrated with respect to $\boldsymbol{p}$. Assuming that the latter integral exists, one arrives at

$$\begin{aligned}&\int h^2(\boldsymbol{p})\mathrm{f}(\boldsymbol{p})\mathrm{d}\boldsymbol{p} - h_0^2 = \mathrm{Var}\{Y\} = \textstyle\sum_{i=1}^{M} \int h_i^2 \mathrm{f}_i(p_i)\mathrm{d}p_i \\ &+ \textstyle\sum_{i=1}^{M} \sum_{j>i}^{M} \int\int h_{i,j}^2 \mathrm{f}_{i,j}(p_i,p_j)\mathrm{d}p_i\mathrm{d}p_j + \ldots + \sum_{i=1}^{M} \int \ldots \int h_{-i}^2 \mathrm{f}_{-i}(\boldsymbol{p}_{-i})\mathrm{d}\boldsymbol{p}_{-i} \\ &+ \int \ldots \int h_{1,\ldots,M}^2 \mathrm{f}(\boldsymbol{p})\mathrm{d}\boldsymbol{p}\end{aligned} \tag{6.62}$$

The right-hand side of (6.62) may be written as

$$\begin{aligned}\mathrm{Var}\{Y\} = &\sum_{i=1}^{M} V_i + \sum_{i=1}^{M} \sum_{j>i}^{M} V_{i,j} + \sum_{i=1}^{M} \sum_{j>i}^{M} \sum_{k>j}^{M} V_{i,j,k} \\ &+ \ldots + \sum_{i=1}^{M} V_{-i} + V_{1,\ldots,M}\end{aligned} \tag{6.63}$$

$$\begin{aligned}
V_i &:= \mathrm{Var}\{\mathrm{E}\{Y|P_i\}\} && i = 1,\ldots,M \\
V_{i,j} &:= \mathrm{Var}\{\mathrm{E}\{Y|P_i,P_j\}\} - V_i - V_j && i = 1,\ldots,M; j > i \\
V_{i,j,k} &:= \mathrm{Var}\{\mathrm{E}\{Y|P_i,P_j,P_k\}\} - V_i - V_j - V_k - V_{i,j} - V_{i,k} - V_{j,k} && i = 1,\ldots,M; j > i;\ k > j \\
&\vdots \\
V_{1,\ldots,M} &:= \mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P}\}\} - \textstyle\sum_{i=1}^{M} V_i - \sum_{i=1}^{M} \sum_{j>i}^{M} V_{i,j} - \sum_{i=1}^{M} \sum_{j>i}^{M} \\
&\quad \textstyle\sum_{k>j}^{M} V_{i,j,k} - \ldots - \sum_{i=1}^{M} V_{-i}
\end{aligned} \tag{6.64}$$

since the integrals of all mixed terms $h_0(\mathrm{E}\{Y|\ldots\} - h_0)$ are zero and all other mixed terms cancel out and $\mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P}\}\} = \mathrm{Var}\{Y\}$.

Some of the terms in (6.64), involving two or more uncertain data in the condition, may be negative in the case of state of knowledge dependence.

$$V_i^T = V_i + \sum_{j>i}^{M} V_{i,j} + \sum_{j>i}^{M} \sum_{k>j}^{M} V_{i,j,k} + \ldots \quad \text{for } i = 1, \ldots, M \tag{6.65}$$

is the individual total contribution of each of the uncertain data to $\mathrm{Var}\{Y\}$.

$V_i^T/\mathrm{Var}\{Y\}$ is called the total uncertainty importance index of the uncertain datum P_i, while $V_i/\mathrm{Var}\{Y\}$ is the principal (also called "first-order") uncertainty importance

index of P_i with respect to the model result Y. $V_i/\mathrm{Var}\{Y\}$ is known as the square value of the correlation ratio of P_i with respect to Y (see Sect. 6.3.1.4).

Each of the terms in (6.64) is the result of a multiple integral, i.e.

$$V_i = \int \left[\int \ldots \int \mathrm{h}(\boldsymbol{p})\mathrm{f}_{-\mathrm{i}}(\boldsymbol{p}_{-\boldsymbol{i}})d\boldsymbol{p}_{-\boldsymbol{i}}\right]^2 \mathrm{f}_{\mathrm{i}}(p_i)\mathrm{d}p_i - (\mathrm{E}\{Y\})^2$$

$$V_{i,j} = \int\int \left[\int \ldots \int \mathrm{h}(\boldsymbol{p})\mathrm{f}_{-\mathrm{i},-\mathrm{j}}(\boldsymbol{p}_{-\boldsymbol{i},-\boldsymbol{j}})\mathrm{d}\boldsymbol{p}_{-\boldsymbol{i},-\boldsymbol{j}}\right]^2 \mathrm{f}_{\mathrm{i,j}}(p_i,p_j)\mathrm{d}p_i\mathrm{d}p_j - (\mathrm{E}\{Y\})^2 - V_i - V_j$$

and so forth. The corresponding conditional densities are required under the integrals in the case of state of knowledge dependence.

The computational cost of evaluating these integrals over M dimensions for the mean value $\mathrm{E}\{\mathrm{h}(\boldsymbol{P})\}$ and for each of the $\sum_{m=1}^{M} C_M^m = 2^M - 1$ terms in (6.63) is out of the question for the number of model results to be analysed and for the number of uncertain data involved in computer models of practical relevance where the number M of uncertain data is usually large and the uncertainty importance measures are required for various model results often evaluated at several points of time or space. Analytic solutions are usually not available, and approximate solutions are to be obtained either numerically or by Monte Carlo simulation. Both ways will require evaluations of the model results for large numbers of data value vectors $\boldsymbol{p}$.

In order to obtain the total variance contribution of the uncertain datum P_i, it is not necessary to evaluate each of the terms in (6.65) since

$$V_i^T = \mathrm{E}\{\mathrm{Var}\{Y|\boldsymbol{P}_{-\boldsymbol{i}}\}\} = \mathrm{Var}\{Y\} - \mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P}_{-\boldsymbol{i}}\}\} \tag{6.66}$$

with

$$Var\{E\{Y|\boldsymbol{P}_{-\boldsymbol{i}}\}\} = \int \ldots \int \left[\int h(\boldsymbol{p})f_{i|-i}(p_i)dp_i\right]^2 f_{-i}(\boldsymbol{p}_{-\boldsymbol{i}})d\boldsymbol{p}_{-\boldsymbol{i}} - (E\{Y\})^2$$

and $\boldsymbol{P}_{-\boldsymbol{i}}$ is the vector of all uncertain data except P_i.

Examples

The following two simple examples serve to illustrate the method and the results obtained:

(a)

$$Y = \mathrm{h}(P_1, P_2, P_3) = P_1 + P_2P_3$$

The three uncertain data are pairwise state of knowledge independent.

$$\begin{aligned}\mathrm{E}\{Y\} &= \mathrm{E}\{P_1\} + \mathrm{E}\{P_2)E\{P_3\}\\ \mathrm{Var}\{Y\} &= \mathrm{Var}\{P_1\} + \mathrm{Var}\{P_2P_3\}\\ &= \mathrm{Var}\{P_1\} + \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\} + \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\} + \mathrm{Var}\{P_2\}\mathrm{Var}\{P_3\}\end{aligned}$$

$$V_1 = \mathrm{Var}\{\mathrm{E}\{Y|P_1\}\} = \int \left[\int\int (p_1 + p_2 p_3)\mathrm{f}_{2,3}(p_2, p_3)\mathrm{d}p_2\mathrm{d}p_3\right]^2 \mathrm{f}_1(p_1)\mathrm{d}p_1 - (\mathrm{E}\{Y\})^2$$
$$= \mathrm{Var}\{P_1\}$$
$$V_2 = \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\}$$
$$V_3 = \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\}$$
$$\sum_{i=1}^{3} V_i < \mathrm{Var}\{Y\}$$
$$V_{1,2} = \mathrm{Var}\{\mathrm{E}\{Y|P_1, P_2\}\} - V_1 - V_2$$
$$= \int\int \left[\int (p_1 + p_2 p_3)\mathrm{f}_3(p_3)\mathrm{d}p_3\right]^2 \mathrm{f}_{1,2}(p_1, p_2)\mathrm{d}p_1\mathrm{d}p_2 - (\mathrm{E}\{Y\})^2 - V_1 - V_2$$
$$\mathrm{Var}\{\mathrm{E}\{Y|P_1, P_2\}\} = \mathrm{Var}\{P_1\} + \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\}$$
$$V_{1,2} = 0$$
$$\mathrm{Var}\{\mathrm{E}\{Y|P_1, P_3\}\} = \mathrm{Var}\{P_1\} + E^2\{P_2\}\mathrm{Var}\{P_3\}$$
$$V_{1,3} = 0$$
$$\mathrm{Var}\{\mathrm{E}\{Y|P_2, P_3\}\} = \mathrm{Var}\{P_2\}\mathrm{Var}\{P_3\} + \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\} + \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\}$$
$$V_{2,3} = \mathrm{Var}\{P_2\}\mathrm{Var}\{P_3\}$$
$$\sum_{i=1}^{3} V_i + \sum_{i=1}^{3}\sum_{j>i}^{3} V_{i,j} = \mathrm{Var}\{Y\}$$
$$\mathrm{Var}\{\mathrm{E}\{Y|P_1, P_2, P_3\}\} = \mathrm{Var}\{Y\}$$
$$V_{1,2,3} = \mathrm{Var}\{\mathrm{E}\{Y|P_1, P_2, P_3\}\} - \sum_{i=1}^{3} V_i - \sum_{i=1}^{3}\sum_{j>i}^{3} V_{i,j} = 0$$
$$V_1^T = V_1 + V_{1,2} + V_{1,3} + V_{1,2,3} = \mathrm{Var}\{P_1\}$$
$$V_2^T = \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\} + \mathrm{Var}\{P_2\}\mathrm{Var}\{P_3\}$$
$$V_3^T = \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\} + \mathrm{Var}\{P_2\}\mathrm{Var}\{P_3\}$$
$$\mathrm{Var}\{Y\} - V_2^T = \mathrm{Var}\{P_1\} + \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\}.$$

If the true value of P_2 is p_2, then

$$Var\{Y|P_2 = p_2\} = Var\{P_1\} + p_2^2 Var\{P_3\} = Var\{Y\} - V_2^T$$

evaluated for $P_2 = p_2$.

(b)

$$Y = \mathrm{h}(P_1, P_2, P_3) = P_1 + P_2 P_3$$

The uncertain data P_1 and P_2 are state of knowledge dependent.

$$\mathrm{E}\{Y\} = \mathrm{E}\{P_1\} + \mathrm{E}\{P_2)E\{P_3\}$$

$$\begin{aligned}
\mathrm{Var}\{Y\} &= \mathrm{E}\{Y^2\} - \mathrm{E}^2\{Y\} = \int\int\int [p_1 + p_2 p_3]^2 \mathrm{f}_{1|2}(p_1|p_2)\mathrm{f}_2(p_2)\mathrm{f}_3(p_3)\mathrm{d}p_1\mathrm{d}p_2\mathrm{d}p_3 \\
&\quad -\left(\mathrm{E}^2\{Y\}\right) = \mathrm{Var}\{P_1\} + \mathrm{Var}\{P_2P_3\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
&= \mathrm{Var}\{P_1\} + \mathrm{Var}\{P_3\}\mathrm{Var}\{P_2\} + \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\} \\
&\quad +\mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
V_1 &= \mathrm{Var}\{\mathrm{E}\{Y|P_1\}\} \\
&= \int\left[\iint (p_1 + p_2 p_3)\mathrm{f}_{2|1}(p_2|p_1)\mathrm{f}_3(p_3)\mathrm{d}p_2\mathrm{d}p_3\right]^2 \mathrm{f}_1(p_1)\mathrm{d}p_1 - (\mathrm{E}\{Y\})^2 \\
&= \mathrm{Var}\{P_1\} + \mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\}\mathrm{E}^2\{P_3\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
V_2 &= \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\} + Var\{\mathrm{E}\{P_1|P_2\}\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
V_3 &= \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\} \\
V_{1,2} &= \mathrm{Var}\{\mathrm{E}\{Y|P_1, P_2\}\} - V_1 - V_2 \\
&= \iint\left[\int (p_1 + p_2 p_3)\mathrm{f}_3(p_3)\mathrm{d}p_3\right]^2 \mathrm{f}_{1|2}(p_1|p_2)\mathrm{f}_2(p_2)\mathrm{d}p_1\mathrm{d}p_2 - (\mathrm{E}\{Y\})^2 - V_1 - V_2 \\
\mathrm{Var}\{\mathrm{E}\{Y|P_1, P_2\}\} &= \mathrm{Var}\{P_1\} + \mathrm{E}\{P_3\}\mathrm{Var}\{P_2\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
V_{1,2} &= -\mathrm{E}^2\{P_3\}\mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\} - \mathrm{Var}\{\mathrm{E}\{P_1|P_2\}\} - 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
\mathrm{Var}\{\mathrm{E}\{Y|P_1, P_3\}\} &= \mathrm{Var}\{P_1\} + \mathrm{E}\{P_3^2\}\mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\} + \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\} \\
&\quad + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
V_{1,3} &= \mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\}\mathrm{Var}\{P_3\} \\
\mathrm{Var}\{\mathrm{E}\{Y|P_2P_3\}\} &= \mathrm{Var}\{\mathrm{E}\{P_1|P_2\}\} + \mathrm{Var}\{P_2P_3\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
V_{2,3} &= \mathrm{Var}\{P_2\}\mathrm{Var}\{P_3\}
\end{aligned}$$

$$\sum_{i=1}^{3} V_i + \sum_{i=1}^{3}\sum_{j>i}^{3} V_{i,j} > \mathrm{Var}\{Y\}$$

$$\begin{aligned}
\mathrm{Var}\{\mathrm{E}\{Y|P_1, P_2, P_3\}\} &= \mathrm{Var}\{Y\} = \mathrm{Var}\{P_1\} + \mathrm{Var}\{P_2P_3\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
V_{1,2,3} &= \mathrm{Var}\{\mathrm{E}\{Y|P_1, P_2, P_3\}\} - \sum_{i=1}^{3} V_i + \sum_{i=1}^{3}\sum_{j>i}^{3} V_{i,j} \\
&= -\mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\}\mathrm{Var}\{P_3\} \\
V_1^T &= V_1 + V_{1,2} + V_{1,3} + V_{1,2,3} = \mathrm{Var}\{P_1\} - \mathrm{Var}\{\mathrm{E}\{P_1|P_2\}\} \\
V_2^T &= \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\} + \mathrm{Var}\{P_2\}\mathrm{Var}\{P_3\} - \mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\}\mathrm{E}\{P_3^2\} \\
&= \mathrm{Var}\{P_2\}\mathrm{E}\{P_3^2\} - \mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\}\mathrm{E}\{P_3^2\} \\
V_3^T &= \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\} + \mathrm{Var}\{P_2\}\mathrm{Var}\{P_3\} = \mathrm{Var}\{P_3\}\mathrm{E}\{P_2^2\} \\
\mathrm{Var}\{\mathrm{Y}\} - V_2^T &= \mathrm{Var}\{P_1\} + \mathrm{E}^2\{P_2\}\mathrm{Var}\{P_3\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\} \\
&\quad - \mathrm{Var}\{\mathrm{E}\{P_2|P_1\}\}\mathrm{E}\{P_3^2\}
\end{aligned}$$

If p_2 turns out to be the true value of P_2, then $\mathrm{Var}\{Y\} - V_2^T$ evaluated at $P_2 = p_2$ equals

$$\begin{aligned}
&\mathrm{Var}\{P_1|P_2 = p_2\} + p_2^2\mathrm{Var}\{P_3\} = \mathrm{Var}\{Y|P_2 = p_2\} \\
&\mathrm{E}\{\mathrm{Var}\{Y|P_3\}\} = \int\left[\int\int (p_1 + p_2 p_3)^2 \mathrm{f}_{1|2}(p_1|p_2)\mathrm{f}_2(p_2)\mathrm{d}p_1\mathrm{d}p_2\right. \\
&\quad \left. -\left(\int\int (p_1 + p_2 p_3)\mathrm{f}_{1|2}(p_1|p_2)\mathrm{f}_2(p_2)\mathrm{d}p_1\mathrm{d}p_2\right)^2\right]\mathrm{f}_3(p_3)\mathrm{d}p_3
\end{aligned}$$

$$\mathrm{E}\{\mathrm{Var}\{Y|P_3\}\} = \mathrm{Var}\{P_1\} + \mathrm{E}\{P_3^2\}Var\{P_2\} + 2\mathrm{E}\{P_3\}\mathrm{Cov}\{P_1, P_2\}$$

and using $V_3 = \mathrm{Var}\{\mathrm{E}\{Y\,|\,P_3\}\}$ as well as

$$\mathrm{E}\{P_3^2\}\mathrm{Var}\{P_2\} = \mathrm{Var}\{P_3\}\mathrm{Var}\{P_2\} + \mathrm{E}^2\{P_3\}\mathrm{Var}\{P_2\}$$

it follows

$$\mathrm{Var}\{\mathrm{E}\{Y|P_3\}\} + \mathrm{E}\{\mathrm{Var}\{Y|P_3\}\} = \mathrm{Var}\{Y\}$$

End of Examples

State of knowledge independence of the uncertain data is not required for the variance decomposition since

$$\mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P_L}\}\} = \int[\int \mathrm{h}(\boldsymbol{p})\mathrm{f}_{-\mathrm{L}|\mathrm{L}}(\boldsymbol{p_{-L}}|\boldsymbol{p_L}) \prod_{l\in -L} \mathrm{d}p_l]^2 \mathrm{f}_L(p_L) \prod_{l\in L} \mathrm{d}p_l$$
$$-\left(\int \mathrm{h}(\boldsymbol{p})\mathrm{f}(\boldsymbol{p})\mathrm{d}\boldsymbol{p}\right)^2 = \mathrm{E}\left\{(\mathrm{E}\{Y|\boldsymbol{P_L}\})^2\right\} - \mathrm{E}^2\{Y\}$$

and

$$\mathrm{E}\{\mathrm{Var}\{Y|\boldsymbol{P_L}\}\} = \int \left(\int (\mathrm{h}(\boldsymbol{p}))^2 \mathrm{f}_{-\mathrm{L}|\mathrm{L}}(\boldsymbol{p_{-L}}|\boldsymbol{p_L}) \prod_{l\in -L} \mathrm{d}p_l \right.$$
$$\left. - \left[\int \mathrm{h}(\boldsymbol{p})\mathrm{f}_{-\mathrm{L}|\mathrm{L}}(\boldsymbol{p_{-L}}|\boldsymbol{p_L}) \prod_{l\in -L} \mathrm{d}p_l\right]^2\right) \mathrm{f}_{\mathrm{L}}(\boldsymbol{p_L}) \prod_{l\in L} \mathrm{d}p_l$$
$$= \mathrm{E}\{Y^2\} - \mathrm{E}\left\{(\mathrm{E}\{Y|\boldsymbol{P_L}\})^2\right\}.$$

$Y = \mathrm{h}(\boldsymbol{P})$, $\boldsymbol{P_L}$ is the vector of any subset L of the set $P_1, P_2, \ldots, P_M$ and $\boldsymbol{P_{-L}}$ is its complement, $\mathrm{f}_{-\mathrm{L}|\mathrm{L}}$ is the conditional subjective probability density function of $\boldsymbol{P_{-L}}$ given $\boldsymbol{P_L}$. However,

$$V_L^T = \mathrm{Var}\{Y\} - \mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P_{-L}}\}\} = \mathrm{E}\{\mathrm{Var}\{Y|\boldsymbol{P_{-L}}\}\} \geq V_L$$

may not hold if $\mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P_{-L}}\}\}$ contains contributions to V_L that are due to uncertain data in $\boldsymbol{P_{-L}}$ that are state of knowledge dependent with uncertain data in $\boldsymbol{P_L}$ (compare V_1 and V_1^T of the example (b) above).

The total uncertainty importance index

$$\frac{V_i^T}{\mathrm{Var}\{Y\}} = \frac{\mathrm{E}\{\mathrm{Var}\{Y|\boldsymbol{P_{-i}}\}\}}{\mathrm{Var}\{Y\}} = 1 - \frac{\mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P_{-i}}\}\}}{\mathrm{Var}\{Y\}} \tag{6.67}$$

is a very powerful yet very expensive uncertainty importance measure. As has been shown above, state of knowledge independence of the uncertain data is not required for (6.67) to hold but, as has been explained above, $V_i^T \geq V_i$ may not hold if uncertain data in $\boldsymbol{P_{-i}}$ are state of knowledge dependent with P_i. Once the true value p_i of P_i is known ($\mathrm{Var}\{Y\} - V_i^T$) evaluated at p_i is what remains of $\mathrm{Var}\{Y\}$, i.e. it equals $\mathrm{Var}\{Y\,|\,P_i = p_i\}$.

Numerical integration or Monte Carlo methods will need to be applied unless the multiple integrals can be evaluated analytically. The corresponding computational effort will simply be unaffordable for the results of most computer models of practical relevance. This is not surprising as $\mathrm{E}\{Y|\boldsymbol{P}_{-i}\}$ is the point-wise least squares approximation of Y over the $(M-1)$-dimensional parameter space and $\mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P}_{-i}\}\}$ is its variance. Consequently, with reference to $\boldsymbol{P}_{-i}$, $\mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P}_{-i}\}\}$ is the equivalent of SSR and $V_i^T = \mathrm{Var}\{Y\} - \mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P}_{-i}\}\}$ is the equivalent of SSE in regression analysis so that $\mathrm{Var}\{\mathrm{E}\{Y|\boldsymbol{P}_{-i}\}\}/\mathrm{Var}\{Y\}$ and $V_i^T/\mathrm{Var}\{Y\}$ are the equivalent of R^2 and $1 - R^2$, respectively.

The first-order uncertainty importance index

$$\frac{V_i}{\mathrm{Var}\{Y\}} = \frac{\mathrm{Var}\{\mathrm{E}\{Y|P_i\}\}}{\mathrm{Var}\{Y\}}$$

is the quotient of the variance of the point-wise least squares approximation of Y over P_i and the total variance of Y. Consequently, with reference to P_i,

$$\mathrm{Var}\{\mathrm{E}\{Y|P_i\}\} = \int [\int \dots \int \mathrm{h}(\boldsymbol{p})\mathrm{f}_{-\mathrm{i}|\mathrm{i}}(\boldsymbol{p}_{-i}|\boldsymbol{p}_i)\mathrm{d}\boldsymbol{p}_{-i}]^2 \mathrm{f}_\mathrm{i}(p_i)\mathrm{d}p_i - (\mathrm{E}\{Y\})^2$$

is the equivalent of SSR in regression analysis, while $V_i/\mathrm{Var}\{Y\}$ is the equivalent of the R^2 value.

In order to somewhat reduce the computational effort involved in the estimation of V_i and of V_i^T, mainly two estimation procedures are suggested in the literature for the case of state of knowledge independent uncertain data. They proceed in the following steps:

1. First they require the evaluation of the computer model for two random samples of size N:

Sample No. 1							Sample No. 2					
Data index	1	2	…	…	…	M	1	2	…	…	…	M
Sample element												
1	$p_{1,1}^{(1)}$	$p_{2,1}^{(1)}$	…	$p_{i,1}^{(1)}$	…	$p_{M,1}^{(1)}$	$p_{1,1}^{(2)}$	$p_{2,1}^{(2)}$	…	$p_{i,1}^{(2)}$	…	$p_{M,1}^{(2)}$
2	$p_{1,2}^{(1)}$	$p_{2,2}^{(1)}$	…	$p_{i,2}^{(1)}$	…	$p_{M,2}^{(1)}$	$p_{1,2}^{(2)}$	$p_{2,2}^{(2)}$	…	$p_{i,2}^{(2)}$	…	$p_{M,2}^{(2)}$
⋮												
N	$p_{1,N}^{(1)}$	$p_{2,N}^{(1)}$	…	$p_{i,N}^{(1)}$	…	$p_{M,N}^{(1)}$	$p_{1,N}^{(2)}$	$p_{2,N}^{(2)}$	…	$p_{i,N}^{(2)}$	…	$p_{M,N}^{(2)}$

Evaluation of the model result for the sample elements

$$\boldsymbol{p}_{\boldsymbol{n}}^{(1)} = \left(p_{1,n}^{(1)}, p_{2,n}^{(1)}, .., p_{i,n}^{(1)}, .., p_{M,n}^{(1)}\right) \quad \text{and}$$
$$\boldsymbol{p}_{\boldsymbol{n}}^{(2)} = \left(p_{1,n}^{(2)}, p_{2,n}^{(2)}, .., p_{i,n}^{(2)}, .., p_{M,n}^{(2)}\right)$$

for $n = 1, 2, \dots, N$ provides the two samples of model results $y^{(k)} = \left(y_1^{(k)}, y_2^{(k)}, \dots, y_N^{(k)}\right)$, $k = 1, 2$ with $y_n^{(k)} = \mathrm{h}\left(\boldsymbol{p}_{\boldsymbol{n}}^{(\boldsymbol{k})}\right)$.

2. Estimation of V_i

To this end, the column for the uncertain datum with index i, in sample no. 2, is replaced by the corresponding column in sample no. 1 and the computer model is evaluated for the modified sample no. 2 to give the sample of model result values $\boldsymbol{y}^{(2,i)} = \left(y_1^{(2,i)}, y_2^{(2,i)}, \ldots, y_N^{(2,i)}\right)$. Among the formulae, suggested in the literature for the estimation of V_i, are (see Saltelli et al. (2010) for a compilation):

(a)

$$\widehat{V}_i = N^{-1} \sum_{n=1}^{N} y_n^{(1)} y_n^{(2,i)} - \left(N^{-1} \sum_{n=1}^{N} y_n^{(1)}\right)\left(N^{-1} \sum_{n=1}^{N} y_n^{(2,i)}\right)$$

(b)

$$\widehat{V}_i = \widehat{\mathrm{Var}\{Y\}} - (2N)^{-1} \sum_{n=1}^{N} \left(y_n^{(1)} - y_n^{(2,i)}\right)^2$$

The second formula avoids summation of products of possibly large absolute values. Division of $\widehat{V}_i$ by $\widehat{\mathrm{Var}\{Y\}}$ gives the estimate of the principal (or "first-order") uncertainty importance index for the uncertain datum P_i. Estimation of all M principal uncertainty importance indices requires $N(M + 2)$ runs of the computer model in order to obtain the sample values $\boldsymbol{y}^{(1)}$, $\boldsymbol{y}^{(2)}$ and $\boldsymbol{y}^{(2,i)}$, $i = 1, \ldots, M$.

3. Estimation of V_i^T

In order to obtain an estimate of V_i^T, the column for the uncertain datum with index i in sample no. 1 is replaced by the corresponding column in sample no. 2 and the computer model is evaluated for the modified sample no. 1 to give the sample of model result values $\boldsymbol{y}^{(1,i)} = \left(y_1^{(1,i)}, y_2^{(1,i)}, \ldots, y_N^{(1,i)}\right)$. Among the formulae, suggested in the literature for the estimation of V_i^T, are (see Saltelli et al. 2010 for a compilation):

(a)

$$\widehat{V}_i^T = \widehat{\mathrm{Var}\{Y\}} - N^{-1} \sum_{n=1}^{N} y_n^{(1)} y_n^{(1,i)} - \left(N^{-1} \sum_{n=1}^{N} y_n^{(1)}\right)\left(N^{-1} \sum_{n=1}^{N} y_n^{(1,i)}\right)$$

(b)

$$\widehat{V}_i^T = (2N)^{-1} \sum_{n=1}^{N} \left(y_n^{(1)} - y_n^{(1,i)}\right)^2$$

Division of $\widehat{V}_i^T$ by $\widehat{\mathrm{Var}\{Y\}}$ gives the estimate of the total uncertainty importance index for the uncertain datum P_i. Estimation of all M total uncertainty importance indices requires NM additional runs of the computer model in order to obtain the sample values $\boldsymbol{y}^{(1,i)}$, $i = 1, \ldots, M$.

The estimation procedure for the case of state of knowledge dependence differs from the one described above only in so far as for $\boldsymbol{y}^{(1,i)}$ and $\boldsymbol{y}^{(2,i)}$, $i = 1, \ldots, M$ new samples of parameter values are to be drawn with the values for $p_{i,n}^{(1,i)}$ to be drawn

under the condition $\boldsymbol{P}_{-i} = \boldsymbol{p}_{-i,n}^{(1)}$ and $\boldsymbol{p}_{-i,n}^{(2,i)}$ to be drawn under the condition of $P_i = p_{i,n}^{(1)}$. The computational cost is the same as in the case of no state of knowledge dependence except for the generation of the conditional sample values of the uncertain data concerned.

It may be possible to stop the process of estimating V_i or V_i^T for P_i, i = 1, . . ., M once an uncertain datum has been found with first order or total contribution to Var{Y} that is sufficiently close to the value of Var{Y} so that one may conclude that this uncertain datum is mainly responsible for Var{Y}. In the uncertainty analysis setting, it will also not be necessary to compute estimates of the indices with up to three or more digits (as is sometimes attempted in the literature on sensitivity analysis) since only the main uncertainty contributors are of interest. Last but not least, the sample sizes needed to obtain the estimates of V_i and/or V_i^T are out of the question for uncertainty analyses of results from most computer models of practical relevance (Helton et al. 2012).

The approximate correlation ratio introduced in Sect. 6.3.1.4 is another attempt to reduce the computational effort. Here, the sample values of P_i are ordered by increasing magnitude and the ordered set is divided into an exhaustive set of $L = N^{1/2}$ disjoint classes C_l, $l = 1, \ldots, L$. The conditional mean values $E\{Y \mid p_i \in C_l\}$ are estimated from the $N^{1/2}$ sample values of Y in each class. This approach does not require model evaluations for a second sample of size N. In fact, the same Monte Carlo sample that was used for the uncertainty quantification of Y can be used for this approximate computation of the square roots of the M first-order uncertainty importance measures.

6.3.1.6 Recommendations and Necessary Checks

Table 6.6 provides some practical recommendations. Four problem classes are distinguished. The classes are defined by the presence or absence of non-negligible sample correlations (spurious and/or due to specified state of knowledge dependence) and of uncertainties represented by a set of three or more alternatives using their index values as values of a substitute uncertain datum (i.e. alternative model formulations, multivariate output of a feeder model, etc.).

ΔR^2 measures the growth of R^2 from stepwise regression with every parameter that is included in the least squares linear approximation. The intermediate R^2 to the top ranked parameters is thus immediately available to decide whether to include more parameters into the approximation or to stop the stepwise regression procedure. Clearly, the rankings based on correlation coefficients and on standardized regression coefficients may be misleading in problem classes 2 and 4 since correlation with and regression on arbitrarily assigned index values will generally make no sense. The correlation ratio is computed and compared to the rankings obtained from correlation coefficients and standardized regression coefficients. This comparison may help to explain any differences at the top of the ranking. Those differences might be due either to uncertain data expressed by a set of more than two alternatives

Table 6.6 Recommended measures for uncertainty importance ranking

Problem class	Non-negligible sample correlations	>2 alternatives and/or strong non-linearities[a]	Measure	Purpose
1	No	No	SRC[b]	Ranking, direction, R^2
			ΔR^2	Intermediate R^2
2	No	Yes	SRC[b]	Ranking, direction, R^2
			CR	Ranking[c]
			ΔR^2	Intermediate R^2
3	Yes	No	CC	Ranking, direction
			SRC	Ranking, direction, R^2 [d]
			ΔR^2	Intermediate R^2
4	Yes	Yes	CC	Ranking, direction
			SRC	Ranking, direction, R^2 [d]
			CR	Ranking[c]
			ΔR^2	Intermediate R^2

SRC standardized regression coefficient, *CC* correlation coefficient, *CR* correlation ratio, ΔR^2 growth of the sample coefficient of determination R^2 due to the inclusion of an uncertain datum into the least squares linear approximation

[a]And/or non-monotone relationships between the model result and some of the uncertain data

[b]CC and SRC will provide the same ranking for this class (see (6.48))

[c]Compare to the ranking from SRC (and CC, in case of class 4) to check the influence of uncertainties expressed by more than two alternatives or of non-monotone and/or strongly non-linear relationship with the model result. If, for instance, an important model uncertainty is represented by more than 2 model alternatives, then its ranking, obtained from SRCs and CCs, may be wrong due to arbitrary indexing of the alternatives

[d]Compare the ranks to those obtained with CC in order to check for the influence of spurious correlations on the latter (see (6.48))

or to strong non-linearity and particularly non-monotonicity of the model result Y with respect to some uncertain data. It is recommended to seek confirmation of the explanation by viewing scatter plots of the sample values of the model result versus the sample values of the uncertain data involved. In classes 3 and 4, the correlation coefficients and standardized regression coefficients are computed to compare the two rankings and to explain differences among the top ranks on the basis of the sample correlations $r_{m,\,l}$ that are either spurious or due to specified state of knowledge dependences. This also helps to single out the influence of spurious correlations on the parameter ranking derived from correlation coefficients.

– What if R^2 is small?
  A small R^2 value means that the ranking derived from standardized regression coefficients refers to only a small fraction of the sample variance s_Y^2 of the model result. It may be due to strong non-linearities of $Y = \mathrm{f}(P_1, \ldots, P_M)$. These may lead to extreme sample values y_n (outliers). Non-linearities may also lead to changes of direction of influence, indicated by a change of sign of the standardized regression coefficient over some independent variable of the model result (time or space), for one or more of the main contributors to model output

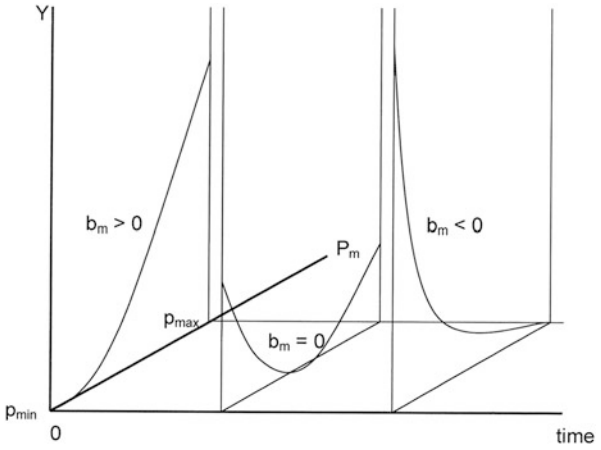


Fig. 6.4 Example of a regression coefficient estimate b_m changing sign over the independent variable "time". A linear relationship between Y and P_m explains much of the variability of Y before and after the change of sign, while little or none at all will be explained at and around the time of the change of sign

uncertainty. As a consequence, R^2 may be small in a neighbourhood of that value of the independent variable where the change in sign takes place (Fig. 6.4).

A small value of R^2 means that a least squares linear approximation to Y in the uncertain data or in a subset thereof is not capable of explaining a major fraction of the sample variance s_Y^2 of the model result. It follows that neither standardized regression coefficients nor correlation coefficients may be suitable for ranking. What about the correlation ratio? It will be suitable since it is not restricted to a linear model. However, the correlation ratio gives no indication of the direction of the influence of the uncertain datum on the model result. The correlation ratio may be large due to few outliers in the sample. They often have a distorting influence on the sample variance of the model result and therefore on the estimates of uncertainty importance measures. To avoid this, a transformation of the sample values is often performed before the uncertainty importance measures are computed.

6.3.2 Uncertainty Importance Measures Computed from Rank Transformed Data

Outliers are often characteristic of random samples from strongly skewed probability distributions. The subjective probability distribution of the model result Y may be strongly skewed due to a non-linear relationship between Y and one or more of the uncertain data. In this case, fractions of the sample variance s_Y^2 explained by a linear least squares regression model may not be a suitable measure of the contribution of an uncertain datum to the uncertainty of the model result since s_Y^2 will be strongly affected by outliers. In this situation, ranges containing given amounts of subjective probability may be better quantitative expressions of the uncertainty of the model result and therefore better suited as reference for uncertainty importance measures.

Therefore, measures derived from the cumulative probabilities of the model result Y and of the uncertain data, instead of their actual values, may be more suitable for ranking. They indicate whether selection of an upper quantile value of an uncertain datum P_m will, in tendency, lead to an upper (or lower) quantile value of Y, irrespective of the absolute values of Y and P_m involved. Only differences in cumulative probability count, i.e. contributions to uncertainty represented by the bulk of the population of values of Y and not so much by extreme values. Uncertainty importance of P_m is now ranked the higher, the better the cumulative probabilities (or complementary cumulative probabilities) of Y agree with the cumulative probabilities of P_m. Thus, the bulk of the population of values of Y has a much higher weight in determining uncertainty importance than the extreme values that are relatively few in probability content. This leads to the use of probability integral transformed data for the purpose of computing uncertainty importance measures.

The probability integral transformed value of an uncertain datum P_m at $P_m = p$ is the value $\mathrm{F_m}(p)$, i.e. the cumulative subjective probability at p.

The probability integral transformed value of the model result Y at $Y = y$ is the value $\mathrm{F_Y}(y)$, i.e. the cumulative subjective probability at y.

The focus of uncertainty importance ranking is not so much on generating an approximate functional expression of the encoded relationship between the model result and the uncertain data but rather on obtaining information about whether there are some uncertain data with a stronger tendency to lead to an upper or lower quantile value of Y, for an upper (lower) quantile value of their subjective probability distribution, than others (Fig. 6.5).

With F_Y the probability integral transformation of the model result $Y = \mathrm{h}(P_1, \ldots, P_M)$ and F_m the probability integral transformation of the uncertain datum P_m, least squares linear regression works with the representation of $\mathrm{F}_Y(y)$ by a linear function in the $F_m(p_m)$, $m = 1, \ldots, M$ including an error term E:

$$\mathrm{F_Y}(y) = \alpha_0 + \sum_{m=1}^{M} \alpha_m \mathrm{F_m}(p_m) + E. \tag{6.68}$$

Unfortunately, the subjective probability distribution of Y and thus its probability integral transformation $\mathrm{F_Y}$ are generally unknown (Fig. 6.6). However, an empirical

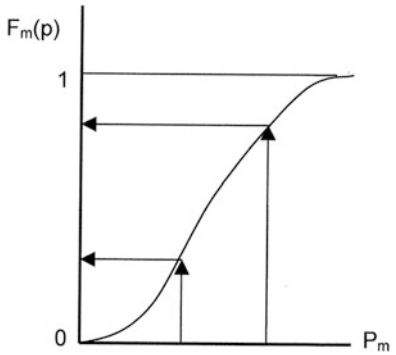


Fig. 6.5 Probability integral transformation of P_m

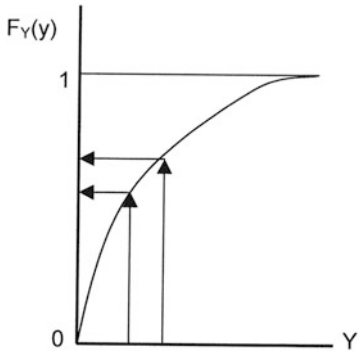


Fig. 6.6 Probability integral transformation of Y

approximation is available from the sample values y_n through rank transformation of the y_n, $n = 1, \ldots, N$. This transformation arranges the y_n in ascending order and assigns an ordinal number to each. The same transformation is performed with the sample values $p_{m,\,n}$, $n = 1, \ldots, N$ for $m = 1, \ldots, M$ (Table 6.7).

From the rank transformed values, estimates of the rank regression coefficients, rank correlation coefficients and rank correlation ratios may be computed, using the ranks $\mathrm{rk}(y_n)$ and $\mathrm{rk}(p_{m,\,n})$ instead of the raw data in the relationships and procedures described in Sect. 6.3.1. See Sect. 3.5.2.5 for a computationally easier way of computing rank correlation coefficients.

Estimates $a_0, a_1, \ldots, a_M$ of the rank regression coefficients are obtained from fitting the expression

$$w_n = a_0 + \sum_{m=1}^{M} a_m \mathrm{rk}\left(p_{m,n}\right) \tag{6.69}$$

to $\mathrm{rk}(y_n)$ for n = 1, …, N such that the following condition is satisfied:

$$\sum_{n=1}^{N} \left(\mathrm{rk}(y_n) - w_n\right)^2 = \sum_{n=1}^{N} \left[\mathrm{rk}(y_n) - a_0 - \sum_{m=1}^{M} a_m \mathrm{rk}\left(p_{m,n}\right)\right]^2 = \textit{minimal}! \tag{6.70}$$

Table 6.7 Example of a rank transformation of the sample values of a model result Y and of the corresponding sample values of the uncertain datum P_m

y	$\mathrm{rk}(y)$	p_m	$\mathrm{rk}(p_m)$
125.36	9	0.25	1
−17.25	1	0.96	9
83.12	6	0.63	5
2.37	3	0.77	7
95.87	7	0.47	2
13.56	4	0.65	6
101.18	8	0.48	3
75.39	5	0.58	4
−1.29	2	0.83	8

The least squares estimates of the rank regression coefficients are identical to the estimates of the standardized rank regression coefficients since mean and variance of the sample elements rk(y_n) equal those of the sample elements rk($p_{m,n}$), $m = 1, \ldots, M$; n = 1, …, N. This follows from (6.36).

The (standardized) rank regression coefficient says by how many rank units the linear least squares approximation W of the rank transformed model result changes if rk(P_m) is changed by one rank unit, all other rk(P_l), $l \neq m$, remaining unchanged. The $R^2_{rk(Y),W}$ value is now the fraction of the sample variance of rk(Y) that is represented by the sample variance of the approximation W. Just as is the case for raw data (see (6.48)), the sample rank correlation coefficient $r_{rk(Y),rk(P_m)}$ is not only an estimate of the amount of rank units by which the approximation W of rk(Y) changes due to the change by one rank unit of rk(P_m), all rk(P_l), $l \neq m$, remaining constant. It also includes estimates of the amounts of rank unit changes that are due to the fractions of rank unit changes of the rank transformed P_l, $l \neq m$, which are state of knowledge dependent on P_m (and/or spuriously correlated with P_m). Practical examples are discussed in Sect. 6.3.3.

If the sample value of the empirical coefficient of determination R^2, computed from rank transformed data, is larger than that computed from raw data, the uncertainty importance ranking from rank transformed data is usually adopted. The latter is with respect to contributions to uncertainty as expressed by the bulk of the sample values of the model result and not so much by the extreme sample values since it is only differences in rank orders that count and not differences in value.

What if also the sample value of R^2, obtained with rank transformed data, does not exceed 0.5? This leaves the uncertainty importance analysis with the rankings from the correlation ratio on raw data or rank transformed data and with viewing scatter plots. Studying scatter plots will be very cumbersome if the analysis has to deal with many uncertain data and with many model results and especially in the case of model results that are a function of time or space. In the presence of spurious correlations, scatter plots may be just as misleading as the correlation coefficients and correlation ratios from raw or rank transformed data.

Approximations (or so-called meta-models), which may be used for uncertainty importance analysis, are discussed in Gatelli et al. (2009), Storlie et al. (2009) and Blatman and Sudret (2010). The statistical analysis of scatter plots is the subject of Kleijnen and Helton (1999).

6.3.3 Practical Examples

Example 1

Figure 1a of Hofer (1999) shows sample correlation coefficients and Fig. 1b shows estimates of standardized regression coefficients as uncertainty importance measures. Spurious correlations among the sample values of the 37 uncertain data are the main cause for the differences between both figures. The few specified state of

knowledge dependences involve only data with very small standardized regression coefficients and therefore hardly change the picture. A good example for the misleading uncertainty importance information from correlation coefficients is the uncertain datum P_{19}. The estimate of the standardized regression coefficient is negligible, while that of the correlation coefficient ranks P_{19} as one of the main uncertainty contributors. The contributions to the estimate of the correlation coefficient are mainly due to positive spurious correlations of P_{19} with P_5, P_9 and P_{11} and negative spurious correlation with P_{16}. The standardized regression coefficients b_5^*, b_9^* and b_{11}^* are negative, while b_{16}^* is positive. Their effect on $r_{Y,\,19}$ can be seen from the relationship (6.48) which reads in this case as

$$r_{Y,19} \approx b_{19}^* + b_5^* r_{19,5} + b_9^* r_{19,9} + b_{11}^* r_{19,11} + b_{16}^* r_{19,16}.$$

The uncertainty importance information for P_{31} and P_{33} is also affected by spurious correlations with P_5, P_9, P_{11} and P_{16}. In this case, the resulting estimates of the correlation coefficients for P_{31} and P_{33} rank these uncertain data as unimportant contributors to the uncertainty of Y, while the standardized regression coefficients rank these uncertain data as important.

Example 2

Comparison of sample rank correlation coefficients and estimates of rank regression coefficients, from an uncertainty analysis of another application of the same computer model as in Example 1, illustrates the difference between rank regression coefficients and rank correlation coefficients as uncertainty importance measures. In this case, a specified correlation among the multivariate samples of values of uncertain data is responsible for the difference. Strong state of knowledge dependence was identified between the uncertain data P_2 and P_3. It was quantified by the rank correlation coefficient $r_{rk(2),\,rk(3)} = -0.95$. The estimates of the rank regression coefficients indicate that P_3 strongly contributes to the uncertainty of the model result Y and that large values of P_3 tend to lead to large values of Y (positive sign of the rank regression coefficient) over the time range up to about 450 s, while the estimate of the rank regression coefficient for P_2 indicates almost negligible uncertainty contribution to Y. This suggests that Y is only weakly (or not at all) a function of P_2. However, due to the relationship (6.48) P_2 receives a high ranking if rank correlation coefficients are used as importance measures. This is due to the uncertainty contributor that P_2 and P_3 have in common. It was the reason for the specification of the strong negative rank correlation $r_{rk(2),\,rk(3)}$. The rank correlation coefficients indicate about equally strong contributions of P_2 and P_3 to uncertainty, however in opposite directions. Clearly, it does not suffice to look at correlation coefficients only, if specified and/or spurious correlations among uncertain data are large. Comparison of correlation coefficients and regression coefficients helps to avoid pitfalls in the identification and interpretation of main contributions to model result uncertainty. This observation is independent of whether raw or rank transformed data are used.

Example 3
Comparison of the estimates of the rank correlation coefficients and of the rank regression coefficients with the uncertainty importance measures of Example 1 (see Figs. 7 and 1a in Hofer 1999), for the same analysis and for the same model result, shows the most striking difference in the importance measure for the uncertain datum P_{16}. The uncertainty importance of P_{16} could not be seen from rank correlation coefficients or from rank regression coefficients. High values of the model result are associated with values of P_{16} in the upper half of the uncertainty range. Rank transformation turns the large absolute differences between these values of the model result and those in the bulk of the sample into unit differences. This way the measure of uncertainty importance of P_{16} is largely reduced. The scatter plot helps to explain this difference. After rank transformation, it is mainly the variability in the bulk of data points, i.e. the cloud at the bottom of the scatter plot, that determines the estimates of the rank correlation and regression coefficient. On raw data, the large gap between the two subsets of model result values is reflected in the estimates of the correlation and standardized regression coefficient. In fact, rank transformation is intended to reduce the effect of outliers (Iman and Conover 1979). But in this example, the focus is exactly on those parameters that give rise to the large model result values in the sample. Since the sample value of R^2 from rank transformed data is here larger than that of R^2 obtained with raw data, one would be inclined to prefer the uncertainty importance information from rank transformed data and would thus miss the important contribution of P_{16} to model result uncertainty as measured by the sample variance s_Y^2 and thus by the importance measures obtained from raw data. The message of this example is that uncertainty importance measures computed from rank transformed data should always be compared to those computed from raw data. Striking differences need to be clarified by viewing scatter plots.

Example 4
Glaeser et al. (1994) show in Figs. 11 and 12 the scatter plots of a model result versus the uncertain data P_{46} and P_{48}, each at a different point in time. The data represent two different model uncertainties. While the values of P_{46} are those of a correction factor to the output from the preferred model formulation, those of P_{48} are the indices of three different model options. A small set of values at the lower end of the uncertainty range of P_{46} leads to very high values of the model result, while all other sample values produce model results from a significantly lower value range. This is due to strong non-linearity of the model result with respect to P_{46}. Accordingly, the correlation ratio ranks P_{46} among the most important uncertainties with a value of 0.86 for the importance measure, while the sample correlation coefficient and rank correlation coefficient assign only a medium to low ranking to P_{46}. Their values are -0.49 and -0.38, respectively. At least they indicate the right direction of the influence of the uncertainty of P_{46}.

In the case of P_{48}, model formulations 1 and 3 lead to a wide range of model result values at a specific point in time, while the results obtained using formulation 2 are mainly concentrated in the lower value range. The correlation ratio ranks P_{48} among the important data uncertainties. Its value is 0.56. On the other hand, the sample

correlation coefficient and rank correlation coefficient rank P_{48} among the unimportant data uncertainties. Their values are 0.1 and 0.03 respectively.

The importance of both model uncertainties could only be detected with the correlation ratio.

Example 5
The application of a small model, with only four uncertain data, provided the model result Y as a function over the time span of 1 s. The sample correlation coefficients ranked P_2, P_3 and P_4 as unimportant uncertainties, while P_1 was ranked as most important with the correlation coefficient decreasing from 1.0 at time t = 0 to 0.65 at t = 0.97. The sample R^2 value from least squares linear regression decreased from 1.0 at t = 0 to 0.5 at t = 0.97. If one had been satisfied with this importance information, one would have missed out on an important uncertainty contributor. The correlation ratio showed the importance of P_3 as increasing from 0.3 at t = 0 to 0.7 at t = 0.97 and that of P_1 as decreasing from 1.0 to about 0.68. The value of the correlation ratio for P_3 increased as fast as the sample R^2 value from linear least squares regression decreased over time. The scatter plot of Y versus P_3 at time t = 0.97 confirmed the ranking of P_3. It is due to non-monotone behaviour of Y with respect to P_3.

Example 6
This is a counter example from the uncertainty analysis of results of a computationally demanding model application. The scatter plot of the model result Y versus the data uncertainty P_{27} shows that the sample values of Y come from two distinct value ranges. Few sample values are from the upper range. They are loosely spread over the range of possibly true values of P_{27} with three sample values concentrated in a narrow value range of P_{27}. All other sample values of Y are from the lower value range of Y and are more or less evenly spread over the value range of P_{27}. The estimate of the correlation ratio assumes a value as high as 0.42 due to the few very high values of the model result in a narrow sub-range of the value range of P_{27}. The sample size is $N = 92$, and it may be expected that a larger sample size would show the values in the upper range of Y to be more evenly spread over the uncertainty range of P_{27} and would thus yield a smaller estimate of the correlation ratio.

Example 7
Returning to Fig. 6.1 and using the explanations of the various importance measures given in this chapter, it will now be possible to identify the main contributors to uncertainty as a function of the distance from the hazardous operation. The estimate of the standardized regression coefficient of the uncertain datum P_4 is -0.6 at distance 400 m and slowly approaches -0.35 at distance 600 m. Since the values of the coefficients for the other three uncertain data move in a band between zero and -0.35 over all distances, P_4 is identified as the main source of uncertainty of the model result at small distance from the operation. The negative sign of the estimate of the standardized regression coefficient says that small values of P_4 tend to lead to large values of the model result and vice versa. The sample R^2 value, however, is

rapidly decreasing over the independent variable "distance". Yet, in the range of largest model result uncertainty it lies well above 0.5. Consequently, the state of knowledge of P_4 needs to be improved if the uncertainty of the hazard index at small distance is to be reduced most effectively.

6.4 Explaining the Outliers

The following is a stepwise description of an algorithm for the application of a statistical test that may be used to determine the uncertain data that are mostly responsible for the top $u\%$ of the population of possibly true values of a computer model result.[2]

Given:

- A computer model that provides the result Y as a function of M uncertain data $P_1, \ldots, P_M$.
- A simple random sample[3] of N (see the remarks below for minimum sample sizes N required) times M values of the uncertain data drawn according to the joint subjective probability distribution of $P_1, \ldots, P_M$, together with the corresponding N values of Y.

One would like to know:

- Which of the uncertain data are mostly responsible for the top $u\%$ of the N sample values of Y.

The steps of the suggested algorithm are:

1. Sort the N sample values of Y, beginning with the smallest value, and store them along with their model run numbers from the Monte Carlo simulation.
   Example:

Rank	Y	Run no.
1	1.3567E-1	124
2	1.4683E-1	15
⋮	⋮	⋮
N	2.5687E+2	73

2. Take the run numbers to the top $u\%$ values (i.e. the largest values of Y). These are $N*u/100$ numbers.

[2]Instead of this algorithm, it may be sufficient to use graphical means like mappings as in Fig. 6.8 for small sample sizes N.

[3]The algorithm should also serve its purpose for a Latin Hypercube sample although the N × M values of the uncertain data are in this case not independently sampled nor are the N values of the model result.

3. Draw at random $N^*u/100$ numbers from a uniform distribution over the interval $[0.5, (100 - u)^*N/100 + 0.5]$ and round these numbers to their nearest integer value.

   Consider them as ranks of values of Y.

4. Find the run numbers that correspond to these ranks in the table of Step 1.
5. There are now two sets of N*u/100 run numbers each. Call the set from Step 2 the "Top Set" and the set from Step 4 the "Bottom Set". The following steps are performed for each of the uncertain data $P_1, \ldots, P_M$ separately.

   For P_m, $m = 1, \ldots, M$ do:

6. Rank the $N^*u/100$ sample values of P_m, which belong to the "Top Set", by increasing magnitude.[4]
7. Rank the $N^*u/100$ sample values of P_m, which belong to the "Bottom Set", by increasing magnitude.
8. Start with the set that contains the smallest sample value of P_m.[5] Call this set of sample values the "Start Set" and

   Assign the value 1 to a variable V
   The value 0 to a variable W
   The absolute value $|V - W|$ to a variable Z.

9. Successively climb up the ranks of the "Start Set" and with each rank

   - Set a variable T equal to the sample value of P_m that corresponds to this rank in the "Start Set"
   - Increase V by 1
   - Increase W to the number of values in the other set that are below or equal to T
   - If $|V - W|$ is larger than Z, replace Z by the new absolute value of the difference.

10. Stop for P_m, as soon as either V or W has reached the value N*u/100.
11. Divide Z by N*u/100 and store the result as d_m. Increase the index m by 1 and proceed to the next uncertain datum. Start with Step 6.

[4] Should a sample value appear several times, assign the corresponding rank numbers accordingly. Example:

Rank	Value
1	1.3567E-2
2	1.4683E-2
3	1.4683E-2
4	1.5925E-2

[5] Should the smallest values of both sets be equal, select one set as "Start Set" and assign the value 1 to both V and W.

12. Once all M uncertain data have been dealt with, compare the values $d_1, \ldots, d_M$ to $g(\alpha)/(N^*u/100)^{1/2}$. α is the significance level of the test in percent divided by 100 and $g(0.01) = 2.30$, $g(0.05) = 1.92$, $g(0.10) = 1.73$ (see Table A20 in Conover 1980). Consider only those parameters P_m with d_m exceeding the value $g(\alpha)/(N^*u/100)^{1/2}$ as possibly important for the upper $u\%$ of values of Y.

    If none of the d_m, $m = 1, ..., M$ exceeds $g(\alpha)/(N^*u/100)^{1/2}$, then it was not possible, at the significance level $\alpha\%$, to identify an uncertain datum as mainly responsible for the magnitude of the top $u\%$ of values of Y in the sample of size N. A larger sample size N may possibly lead to the identification of such a parameter or increasing the significance level α may do.
13. The larger the value of d_m in the significant subset of values from Step 12, the more likely it is that P_m is responsible for the top $u\%$ of sample values of Y.

Remarks

This algorithm makes use of the two-sided Kolmogorov–Smirnov Test (Conover 1980) of the null hypothesis that both samples of $N^*u/100$ values of the uncertain datum (see steps 5, 6 and 7) are sampled according to the same distribution against the alternative hypothesis that they are sampled according to different distributions.

The sample size N needs to be at least $40/(u/100)$ so that $N^*u/100$ is at least 40. For smaller N, the number to be compared to the test statistic d_m may be taken from Table A20 in Conover (1980).

Instead of the statistical test graphical means could be used as shown in Fig. 6.7 (Kurowicka and Cooke 2006).

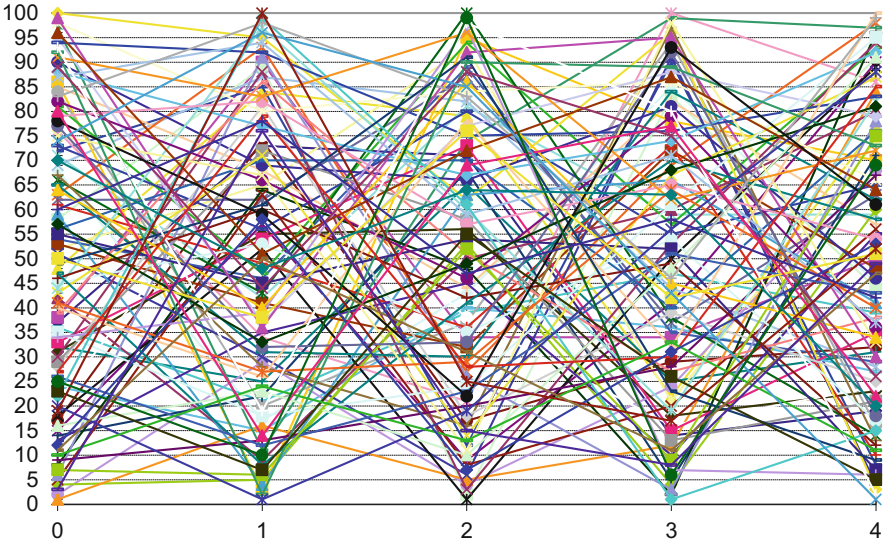


Fig. 6.7 Mapping of the rank transformed values of the four uncertain data of a simple model into the rank transformed values of the model result. The latter are shown on the left-hand side for a sample of size 100

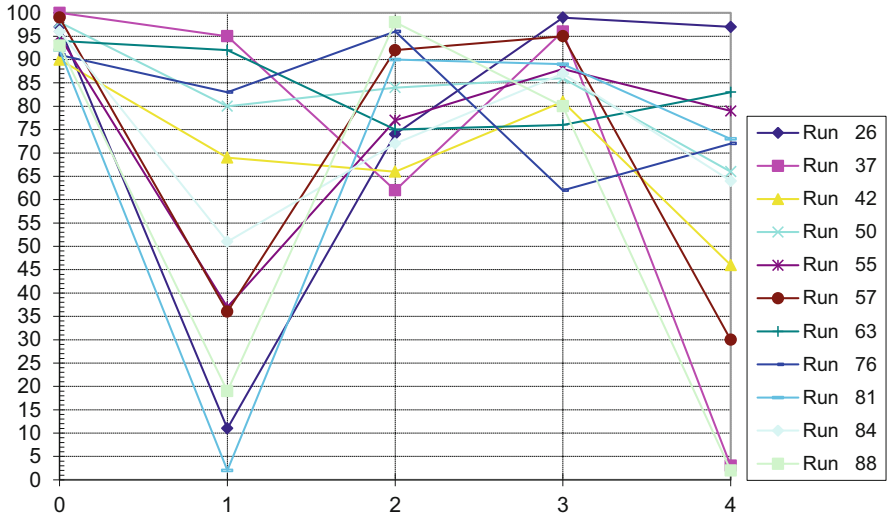


Fig. 6.8 Mapping of the rank transformed values of the four uncertain data into the top 10 ranks of the model result. Clearly, the values of P_2 and P_3 from their upper uncertainty range are mainly responsible for the top 10% of the model result values in the sample of size 100

In Fig. 6.8, the graph has been restricted to the top $u\%$ model result values in order to better identify the uncertain data that are mainly responsible for this part of the subjective probability distribution.

6.5 Contributions to Quality Assurance

There are several ways of how uncertainty analysis and the associated ranking of the uncertain data can contribute to quality assurance:

- Model runs end abnormally for some sets of data values in the input sample of the uncertainty analysis and the uncertainty importance measures for the two-valued model output "run ends normally = 0", "run ends abnormally = 1" points out where to look for causes.
- Due to an error in the model, in the coding, in the input specifications or in the input file, uncertainty importance measures show associations that should not exist or do not show associations that should exist. For instance, in Example 1, the uncertainty importance measure to P_1 ranked this uncertainty as unimportant while it should actually rank among the most influential uncertainties with respect to a specific model result. It turned out to be of negligible uncertainty importance only because the sampled parameter values were entered in the wrong position of the input deck, while the correct position carried the "best estimate" value in all model runs of the Monte Carlo simulation.
- Uncertainty importance measures computed for the "model output" "processor-time per run" may point out model inefficiencies.

– Measured experimental values (including state of knowledge expressions of their measurement errors) are not at all (or only in parts) contained within the uncertainty range of the corresponding values provided by the computer model. The uncertainty importance measures may suggest improvements to the computer model or state of knowledge improvements to those uncertain data that were identified by the uncertainty analysis as having the potential of effectively reducing the discrepancy between measured and computed values.

The uncertainty analysis presented in Kolev and Hofer (1996) serves as an illustration for the last point. Figure 4a of this publication shows the alternative time histories obtained from the uncertainty analysis of a model result (sample size N = 96) for the problem time span of 3 s. The two-sided (90%, 95%) statistical tolerance limit is presented as a function of time in Fig. 4b (lines ending at 2.7 s). The measured values are also shown in Fig. 4b (up to 3.0 s). Figure 4b says that at least 90% of the population of model result values, which follow from the combined effect of all uncertain data accounted for by the uncertainty analysis, lie within the two-sided statistical tolerance limit lines (ending at t = 2.7 s) at a confidence level of at least 95%. The measurement result first runs along and then closely below the upper limit line for times up to about 1.3 s and stays well above the upper limit line for all times above 1.5 s.

The conclusion is that the model result tends to be lower than the measured value. The rank regression coefficients in Fig. 9a identify two uncertain data (P_4 and P_6) as the main contributors to model result uncertainty. They also tell the direction in which the model result tends to change with a given change in the value of these two uncertain data. Positive sign (for P_6) means same direction, and negative sign (for P_4) stands for opposite direction. The sample R^2 value from the rank regression model decreases steadily from a high value of almost 1.0 to 0.54 at time t = 2.7 s. Investigations of how P_4 and P_6 are handled in the model revealed that P_4 was taken as the maximum value (twice the average) where the average value is actually required ("maximum" and "average" not with respect to uncertainty range but in a modelling context). The uncertainty importance ranking, derived from correlation ratios (Fig. 9b), also identifies P_4 and P_6 as the main contributors to model result uncertainty. The correlation ratio does, however, not carry directional information.

The 96 model runs were performed in parallel on an Intel Paragon XP/S MP system with 96 compute nodes.

6.6 Graphical Presentation of the Uncertainty Importance Measures

The most common ways of presenting uncertainty importance measures graphically are the following:

– Bar charts
– Line charts

– Scatter plots
– Mappings

6.6.1 The Bar Chart

The bar chart is used for uncertainty importance measures to single value model results or to continuous value model results at chosen values of the independent variable (i.e. time and/or space) or for their minima or maxima over the full time (space) range or over a sub-range thereof. For small to moderate numbers M of uncertain data it will make sense to present the bars for correlation coefficients, standardized regression coefficients and/or correlation ratios to each of the M uncertain data as vertical bars. For large numbers M stepwise regression will be used and the chart will contain only bars for standardized regression coefficients to those uncertain data that are included in the least squares linear approximation. The bar chart of standardized regression coefficients also needs to show the corresponding sample R^2 value in the legend. Sometimes the chart is drawn vertically (horizontal bars) if M is large and printing space requires.

6.6.2 The Line Chart

The line chart is used for uncertainty importance measures to model results that are a function of some independent variable like time and/or space. The computer model will provide the result only at a finite number of points on the axis of the independent variable. The uncertainty importance analysis will be performed at each of these points or at a subset thereof. For each individual uncertain datum, the line chart then connects the values of the measure along the sequence of points by segments of straight lines. For small to moderate numbers M of uncertain data this can be done for all data although it will make sense to show the uncertainty importance measure for at most 10 uncertain data per line chart. Consequently, a sequence of charts will often be needed to present the lines for all M data. For large numbers M, stepwise regression will be used. The line chart will then need to show the standardized regression coefficients for all uncertain data that are included in the least squares linear approximation at anyone of the set of points in time and/or space. The line connecting the sample R^2 values computed at each point is usually shown in a separate chart.

6.6.3 The Scatter Plot

Scatter plots for the chosen pair of uncertain datum P_m and model result Y_j (or Y_j at a discrete value of an independent variable like time and/or space), show the set of

pairs of values $\{(p_{m,n}, y_{j,n}) \mid n = 1, \ldots, N\}$ from the N runs of the computer model as points in a diagram with the coordinates (P_m, Y_j). They are useful in the search for and detection of relationships that the chosen uncertainty importance measures may not have been able to capture and for the comparisons mentioned in Table 6.6. Sometimes it is helpful to use the number n of the model run as marker of the position in the plot.

6.6.4 The Mapping

Mappings are suitable for single value model results or for continuous value model results at chosen values of the independent variable or for their minima or maxima over the full range of the independent variable or over a sub-range thereof. Their advantage is the capability of visualizing value combinations from a subspace of the M-dimensional space of uncertain data that tend to lead to values from a specific sub-range (i.e. upper or lower range) of the total uncertainty range of the model result (see Fig. 6.2 or Kurowicka and Cooke (2006) for more examples). Obviously, the number of uncertain data shown in one Figure is limited so that a sequence of Figures (using the same colour scheme for the model runs) may be needed to accommodate all uncertain data.

6.7 Conclusions

Uncertainty analysis focuses on the combined influence of all potentially important uncertainties on the model result. To this end, the state of knowledge of model formulations, parameters, application-specific input data and parameters of numerical solution algorithms is quantified and expressed by subjective probability distributions in Step 3 of the analysis. As a consequence of these quantifications and of the logic encoded in the computer model, a subjective probability distribution follows for each of the model results in Step 4. Two aspects of this distribution are of interest depending on the assessment question that is to be answered with the help of the model result:

(a) How large is the subjective probability content of a specified value range of the model result?
(b) How large is the value range that contains subjective probability u/100 with the value of u commonly specified as 90 or higher?

Methods like importance sampling or subset sampling (see Sects. 4.4.3 and 4.4.4) are in use if the subjective probability content of the specified value range is expected to be very small. The main contributors to the subjective probability content are determined in the course of its estimation.

Monte Carlo simulation, using Latin Hypercube sampling or simple random sampling (SRS), is commonly in use otherwise. The simulation provides an estimate of the probability content of a specified value range as well as a one- or two-sided value range with the specified subjective probability content u/100, depending on the question to be answered. If SRS is used, one- or two-sided statistical tolerance limits are available that contain at least u/100 subjective probability at a confidence level of at least v%. This is important if only small sample sizes [in the order of 100 or several 100 executions of the model application (for short called "model runs")] are affordable. The values of u and v are commonly chosen as 90 or higher. If the value range with specified probability content is found to be so large that the model result does not permit meaningful decision-making, then it is required to identify the main contributors to this large uncertainty range. The size of this range is determined by the variance of the subjective probability distribution. Uncertainty importance measures are therefore chosen such that they explain a maximum fraction of the variance. The affordable sample size of the Monte Carlo simulation, i.e. the number of model runs, is usually small for computationally demanding models. For these models, uncertainty importance analysis cannot afford a separate set of model runs performed specifically for the purpose of uncertainty importance analysis. For consistency and efficiency reasons, uncertainty importance analysis has to use those model runs executed for the purpose of uncertainty analysis. Consequently, correlation coefficients, correlation ratios (the square root of the first order importance measures from variance decomposition, available only in their approximate form due to sample size restrictions) and standardized regression coefficients (the latter often obtained from stepwise regression) are with or without transformation into ranks, a reasonable choice of importance measures. The input sample for the model runs is a multivariate random sample of sets of values, each set consisting of one value for each of the numerous uncertain data. For computationally demanding models, the number of uncertain data will be large and the size of the input sample (i.e. the affordable number of model runs) will be comparatively small. It will therefore show the effects of spurious correlations in addition to those of the specified state of knowledge dependences. Due to the presence of correlations within the multivariate sample, it will not be sufficient to look at the uncertainty importance ranking derived from correlation coefficients alone. Differences between this ranking and one obtained from standardized regression coefficients need to be understood if significant. Additional insights can be gained and pitfalls avoided if the estimates of correlation coefficients and standardized regression coefficients are compared. The effect of spurious correlations on uncertainty importance measures may often be singled out by this comparison.

The correlation ratio is an indispensable measure whenever model uncertainty is expressed by more than two alternatives with their indices used as values of a substitute uncertain datum (see Sect. 3.4.1) or when measures, quantifying the extent of linear or monotone relationship between an uncertain datum and a model result, are not adequate.

Differences in uncertainty importance rankings, obtained before and after rank transformation, are due to different ways of viewing the data. Rank transformation is

performed if the uncertainty importance measures should focus more on the bulk of the data and less on outliers or if strongly non-linear relationships between the model result and its main uncertainty contributors are suspected. The differences between importance measures from raw and from rank transformed data provide additional insights and need to be understood.

Uncertainty importance analysis tells where to improve the state of knowledge in order to reduce the uncertainty of the model result most effectively. In other words, it provides guidance as to whether further model development is primarily needed or improved knowledge of parameter and input data values of the model application. This is a choice between two very different directions of further activities with differing costs and chances for success. For instance, the importance analysis may suggest that additional tests, experiments and theoretical investigations, leading to improved model development, are required or that further data collection has priority.

Table 6.6 offers recommendations for the choice of uncertainty importance measures in 4 problem classes. An excellent compilation of additional uncertainty importance measures is to be found in Helton et al. (2006).

Finally, the contribution of uncertainty analysis and of its uncertainty importance ranking to the quality assurance of the model results cannot be appreciated highly enough. It must receive the strong emphasis that it obviously deserves.

References

Bischof, C. H., Hovland, P., & Norris, B. (2008). On the implementation of automatic differentiation tools. *Higher-Order and Symbolic Computation, 21*(3), 311–331.

Blatman, G., & Sudret, B. (2010). Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliability Engineering and System Safety, 95*, 1216–1229.

Bowerman, B. L., O'Connell, R. T., & Koehler, A. B. (2005). *Forecasting, time series and regression* (4th ed.). Belmont, CA: Brooks/Cole.

Cacuci, D. G. (2003). *Sensitivity and uncertainty analysis, Volume I: Theory*. Boca Raton: Chapman & Hall/CRC Press.

Conover, W. J. (1980). *Practical nonparametric statistics* (2nd ed.). New York: Wiley.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York: Wiley.

Freund, R. J., & Minton, P. D. (1979). *Regression methods – A tool for data analysis*. New York: Marcel Dekker.

Gatelli, D., Kucherenko, S., Ratto, M., & Tarantola, S. (2009). Calculating first-order sensitivity measures: A benchmark of some recent methodologies. *Reliability Engineering and System Safety, 94*, 1212–1219.

Glaeser, H., Hofer, E., Kloos, M., & Skorek, T. (1994). Uncertainty and sensitivity analysis of a post-experiment calculation in thermal hydraulics. *Reliability Engineering and System Safety, 45*, 19–33.

Heinhold, J., & Gaede, K.-W. (1968). *Ingenieur-Statistik*. Wien: R. Oldenbourg München.

Helton, J. C., Johnson, J. D., Sallaberry, C. J., & Storlie, C. B. (2006). Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering and System Safety*, *91*, 1175–1209.

Helton, J. C., Hansen, C. W., & Sallaberry, C. J. (2012). Uncertainty and sensitivity analysis in performance assessment for the proposed high-level radioactive waste repository at Yucca Mountain, Nevada. *Reliability Engineering and System Safety, 107*, 44–63.

Hofer, E. (1999). Sensitivity analysis in the context of uncertainty analysis for computationally intensive models. *Computer Physics Communications, 117*, 21–34.

Hora, S. C. (2003). A distribution-free test for the relationship between model input and output when using Latin Hypercube sampling. *Reliability Engineering and System Safety, 79*, 333–339.

Iman, R. L., & Conover, W. J. (1979). The use of the rank transform in regression. *Technometrics, 21*, 499.

Iman, R. L., & Hora, S. C. (1990). A robust measure of uncertainty importance for use in fault tree system analysis. *Risk Analysis, 10*, 401.

Jansen, M. J. W. (1999). Analysis of variance designs for model output. *Computer Physics Communications, 117*, 35–43.

Janssen, P. H. M. (1994). Assessing sensitivities and uncertainties in models: A critical evaluation. In J. Grasman & G. van Straten (Eds.), *Predictability and nonlinear modelling in natural sciences and economics* (p. 344). Dordrecht: Kluwer Academic.

Kendall, M., & Stuart, A. (1973). *The advanced theory of statistics* (Vol. 2, 3rd ed.). New York: Mac Millan.

Kleijnen, J. P. C., & Helton, J. C. (1999). Statistical analysis of scatterplots to identify important factors in large-scale simulations, 1: Review and comparison of techniques. *Reliability Engineering and Systems Safety, 65*, 147–185.

Kolev, N. I., & Hofer, E. (1996). Uncertainty and sensitivity analysis of a postexperiment simulation of non-explosive melt-water interaction. *Experimental Thermal and Fluid Science, 13*, 98–116.

Kruskal, W. H. (1958). Ordinal measures of association. *Journal of the American Statistical Association, 53*, 814–861.

Kucherenko, S., Rodriguez-Fernandez, M., Pantelides, C., & Shah, N. (2009). Monte Carlo evaluation of derivative-based global sensitivity measures. *Reliability Engineering and System Safety, 94*, 1135–1148.

Kurowicka, D., & Cooke, R. M. (2006). *Uncertainty analysis with high dimensional dependence modelling*. New York: Wiley.

Manteufel, R. D. (1996). Variance-based importance analysis applied to a complex probabilistic performance assessment. *Risk Analysis, 16*, 587.

McKay, M. D. (1997). Nonparametric variance-based methods of assessing uncertainty importance. *Reliability Engineering and System Safety, 57*, 267.

McKay, M.D., Beckman, R.J., Moore, L. M., & Picard, R. R. (1992). An alternative view of sensitivity in the analysis of computer codes. In *Proceedings of the Section on Physical and Engineering Sciences of the American Statistical Association*.

Ratto, M., Pagano, A., & Young, P. C. (2009). Non-parametric estimation of conditional moments for sensitivity analysis. *Reliability Engineering and System Safety, 94*, 237–243.

Rosner, B. (1995). *Fundamentals of biostatistics* (4th ed.). New York: Duxbury Press.

Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications, 145*, 280–297.

Saltelli, A., Tarantola, S., & Chan, K. (1999). A quantitative, model independent method for global sensitivity analysis of model output. *Technometrics, 41*(1), 39–56.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications, 181*, 259–270.

Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computational Experiment, 1*, 407–414.

Storlie, C. B., Swiler, L. P., Helton, J. C., & Sallaberry, C. J. (2009). Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering and System Safety, 94*, 1735–1763.

Chapter 7
Step 6: Present the Analysis and Interpret Its Results

7.1 Presentation of the Analysis

A document or publication presenting the results from an uncertainty analysis of a computer model application should have the following structure:

1. The assessment question
   This section presents the questions that are to be answered by the application of the model. In this context, one needs to clarify whether there is a single true answer to each of the questions or whether some of the questions have a population of true answers that is to be summarized by a probability distribution. The latter situation requires the quantification of Type A (aleatoric) uncertainty. This task may either already be covered by the computer model or may be performed in the course of the uncertainty analysis, namely in a second (inner) Monte Carlo loop with the first (outer) loop handling the Type B (epistemic) uncertainties. The approach with two nested Monte Carlo simulations is also sometimes called a 2DMC approach (see Chap. 9).
2. The assessment model
   This section presents the main features of the computer model that was applied to answer the assessment questions.
3. The analysis tool
   Commonly, an uncertainty analysis does not start from scratch but applies a commercially available or custom-made software package that covers at least Steps 3, 4 and 5 of the uncertainty analysis. This section needs to present the main features of the package and to quote the corresponding documentation.
4. The elicitation process
   Experts are not only required for state of knowledge quantifications but even one step earlier (Step 1) namely in the search for and identification of the potentially important uncertainties. This section needs to present the details of the selection

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_7

of the panel of experts together with the structured approach used to identify uncertainties and to elicit the experts' state of knowledge quantifications.

5. The potentially important uncertainties
   The main body of this section is a table presenting the uncertainties identified by the experts as potentially important together with a short explanation of their role within the computer model. This section also explains the methodologies chosen for representing model uncertainties by uncertain parameters.[1]

6. State of knowledge quantifications
   The main body of this section is a table that shows for each uncertain datum the type of subjective probability distribution and its specifying details chosen to quantitatively express the corresponding state of knowledge.

7. State of knowledge dependences
   The main body of this section is a table that contains for all pairs of uncertain data, identified as significantly state of knowledge dependent, the details of their dependence quantification.

8. Propagation of the states of knowledge through the assessment model
   A short description of the Monte Carlo procedure, particularly of the sampling technique chosen, is given in this section. It explains how random sampling at the level of uncertain data (input sample) and the evaluation of the computer model for each sample element result in a random sample of model result values (output sample). The latter is drawn according to the unknown joint subjective probability distribution of the model results that follows logically from the state of knowledge quantifications at the level of the uncertain data and from the encoded instructions of the computer model. The sample of model result values is used to derive uncertainty statements as well as other useful information contained in the sample (like correlations among model results) and, together with the input sample, a ranking of the uncertain data.

9. Uncertainty statements for selected model results
   Of particular interest are those model results that are to be used in answering the assessment questions. The uncertainty statements will be in the form of one- or two-sided ($u\%$, $v\%$) statistical tolerance limits. These limits define intervals of model result values with one endpoint left open in the case of one-sided limits. The intervals say that one can be at least $v\%$ confident that at least $u\%$ subjective probability of the model result are between the endpoints of the interval. This statement also accounts for the sampling error that might be due to the fact that only a sample of finite size was drawn according to the unknown joint subjective probability distribution of the model results.

[1] Model uncertainties are represented by uncertain parameters, and uncertain parameters are considered as uncertain input data. Consequently, only the term uncertain data is used.

The results obtained with best estimate or reference values of the uncertain data may also be of interest particularly if shown in relation to the statistical tolerance limits.

If the model result is a function of time (also called a time history) or any other independent variable, the (u%, v%) statistical tolerance limits will be shown at selected values of this variable. Sometimes, the upper and lower endpoints of the two-sided limits are continuously connected for illustration purposes. This does, however, not necessarily mean that at least u% of the sample of time histories computed for a particular model result run within the continuous connections of the lower and the upper endpoints at a confidence level of at least v%. Since the (u%, v%) statistical tolerance limits are determined locally, at selected points in time, etc., they do indeed contain at least u% of the population of model result values at each individual point at a confidence level of at least v%; however, these values may not belong to the same set of time histories at every selected point in time. Any particular one of the sample of time histories may have its value within the endpoints of the two-sided statistical tolerance limit at some values of the independent variable but not necessarily at all of them. Therefore, it will also be of interest to show the complete sample of time histories (or sequences of model result values over space, etc.).

In addition to the statistical tolerance limits, it will be required to also show the local (at a selected value of the independent variable) empirical subjective probability distribution summarizing the local variability of the output sample values of the model result. This may be of interest at several selected values of the independent variable. In this empirical distribution, the position of specific values like the endpoints of a two-sided statistical tolerance limit and the result obtained with best estimate or reference values of the uncertain data should be indicated. It may also be required to show the empirical subjective probability distribution of the sample of global (over all values of the independent variable) maximum, minimum, mean and median model result values.

In many situations, sample correlations that might exist between any pair of model results will also be important. They are to be presented by scatter plots of the respective sample of pairs of model results at selected values of the independent variable. The plots illustrate the extent of a relationship between the states of knowledge of the two model results in the pair.

10. Uncertainty importance rankings for selected model results

    The purpose of this chapter is to present the main contributors to the uncertainty of the model results. In other words, to show the model users, model developers and the decision-makers primarily where the state of knowledge would need to be improved in order to reduce the uncertainty of the model results most effectively. The uncertainty importance rankings will also show whether these improvements should be performed more at the data side (suggesting further efforts in data gathering) or at the side of model formulations (suggesting further model development).

The information from several uncertainty importance measures will need to be evaluated and presented. This is particularly so if state of knowledge dependence has been specified and/or if spurious correlations in the sample drawn for the uncertain data (input sample) cannot be excluded. Standardized regression coefficients (probably from stepwise regression without or with rank transformation) are a good starting point together with the sample value of R^2 that indicates how much of the uncertainty of the model result is explained by the measures. Low sample values of R^2 as well as uncertainties represented by a set of alternatives, with their indices being used as the values of an uncertain datum, will make it necessary to also look at the rankings derived from the correlation ratio. In the presence of state of knowledge dependences and/or spurious correlations, the ranking derived from correlation coefficients always needs to be compared to the ranking derived from standardized regression coefficients. Additional important recommendations are given in Table 6.6 and in the conclusions of Chap. 6. If the computer model result is a function of some independent variable all measures will need to be presented locally, i.e. at selected values of the independent variable. For illustration purposes, these local values of the uncertainty importance measures will be connected along the axis of the independent variable by straight lines. Uncertainty importance measures may also need to be computed for the global (over all values of the independent variable) maximum, minimum, mean and median model result values.

Uncertainty importance information obtained with respect to the processor time required by each of the model runs helps to identify improvements that aim at a more economical (with respect to run time requirements) computer model.

Last but not least, the uncertainty importance ranking will provide some additional insights into the model internals and will most likely have suggested improvements to the computer model and its application. Some of these insights might be worthwhile to be communicated to the parties concerned.

11. Conclusions

This should be a summary of the main analysis results together with a description of what was learnt about the computer model and its application. It will also contain some recommendations aiming at the reduction of the uncertainty of the model results. Running the computer model for many sets of data values and model formulations, selected from the combined multidimensional uncertainty range, as is the case in uncertainty analysis, is certainly a thorough robustness check of the model. This chapter may, therefore, also include remarks about any model improvements initiated by the analysis or may point out that the computer model successfully passed this robustness test.

7.2 Interpretation of the Uncertainty Estimate

The uncertainty information is contained in the joint subjective probability distribution that follows for the computer model results by propagation of the states of knowledge at the level of parameters, input data and model formulations (for short: uncertain data) through the encoded instructions of the computer model. Uncertainty estimates are derived from a random sample of values of the model results drawn according to this joint distribution. The random selection is done indirectly by drawing a random sample according to a joint subjective probability distribution for all uncertain data with state of knowledge quantifications provided in Step 2 (Chap. 3) and by evaluating the computer model for this sample.

This sequence of analysis steps makes it clear that the correct interpretation of the uncertainty information is as follows:

– A (u%, v%) statistical tolerance interval for the model result Y contains, at a confidence level of at least v%, at least u% of the population of model result values that follow from the combined effect of the state of knowledge quantifications for all uncertain data accounted for by the uncertainty analysis.

Or:

– The model result Y lies, at a confidence level of at least v%, with subjective probability of at least u/100 within the given (u%, v%) statistical tolerance interval. The subjective probability is the analyst's degree of belief that follows from the state of knowledge quantifications at the level of the uncertain data.

The population of model result values (that follow from the combined effect of the state of knowledge quantifications for all uncertain data accounted for by the uncertainty analysis) is a population of possibly true answers to the assessment question under the condition that the following assumptions hold:

– The computer model is not seriously flawed.
– The most important contributors to the uncertainty of its results have been accounted for in the analysis (the combined effect of all those not accounted for may be neglected).
– The states of knowledge at the level of parameters, models and input data have been appropriately expressed by a joint subjective probability distribution.
– Any changes to the interpretation of "subjective probability" (see Sect. 3.4.1) are of only minor consequence.

Some caveats may be contained in the state of knowledge documentation sheets (Sect. 3.6.1.5). The interpretation of the uncertainty information will be the appropriate place to mention these caveats and to discuss the limitations (if any) they may place on the usability of the uncertainty information.

If the above assumptions hold, according to the judgment of the analysis team, then the (u%, v%) statistical tolerance limits for the model result may be interpreted as follows:

- The two-sided ($u\%$, $v\%$) statistical tolerance limit contains at least $u\%$ of the population of possibly true answers to the assessment question at a confidence level of at least $v\%$.
- The upper ($u\%$, $v\%$) statistical tolerance limit says that at least $u\%$ of the population of possibly true answers to the assessment question do not exceed this limit value at a confidence level of at least $v\%$.
- At least u% of the population of possibly true answers to the assessment question do not exceed the given upper safety limit (i.e. comply with the given upper limit value), at a confidence level of at least $v\%$, if the ($u\%$, $v\%$) upper statistical tolerance limit coincides with (or is just below) the given safety limit.

Or equivalently,

- The two-sided ($u\%$, $v\%$) statistical tolerance limit contains the true answer to the assessment question with subjective probability of at least $u/100$ at a confidence level of at least $v\%$.
- The upper ($u\%$, $v\%$) statistical tolerance limit says that the true answer to the assessment question does not exceed this limit value with subjective probability of at least $u/100$ at a confidence level of at least $v\%$.
- The subjective probability, for the true answer to the assessment question not to exceed the given upper safety limit (i.e. to comply with the given upper limit value), is at least $u/100$, at a confidence level of at least $v\%$, if the ($u\%$, $v\%$) upper statistical tolerance limit coincides with (or is just below) the given safety limit.

The subjective probability is the degree of belief of the analysis team.

7.3 Interpretation of the Uncertainty Importance Ranking

The uncertainty importance information is in the form of a ranking of the main contributors to the uncertainty of a model result Y. It is not automatically justified to conclude that Y is most sensitive to the top ranked uncertain datum P_m. The ranking considers not only the sensitivity of Y to P_m but also the uncertainty of P_m. Therefore, if a model result Y is most sensitive to changes in the value of P_m but there is little uncertainty about the true value of P_m, then its contribution to the uncertainty of Y, which is made up of the sensitivity of Y to changes in the value of P_m and the uncertainty of P_m, will be moderate to small. The ranking of P_m will consequently be medium to low. On the other hand, Y may be less sensitive to changes in the value of P_l, but the uncertainty about the true value of P_l is large and therefore the ranking of P_l is high in the uncertainty importance information for the model result Y.

It is important to keep this distinction in mind when interpreting uncertainty importance rankings obtained from an uncertainty analysis. Much too often is the information misinterpreted as being purely a measure of sensitivity while in fact it is

a measure of the combination of the sensitivity of the model result to changes in the value of the uncertain datum and of the uncertainty about the true value of that datum.

Uncertainty importance rankings need to be supplemented by either the R^2 value (if the importance measures are standardized regression coefficients) or by some other quantitative measure (see the example in Sect. 10.2) to show how much of the uncertainty of the model result is actually explained by the top ranked uncertain data.

Chapter 8
Practical Execution of the Analysis

8.1 Stepwise Support by Analysis Software

8.1.1 STEP 1: Search

The analyst conducts, together with the client (representative of the team performing the computer model application) and the modeller (representative of the team that developed the model), a preliminary search for potentially important uncertainties. This search is performed along the lines of the discussion of sources of uncertainty in Chaps. 1 and 2. There may be simple enough situations allowing the search to be conducted by only one person.

The analyst and the client go through the scenario description and note down the potentially important uncertainties that come to mind together with names of experts who might be able to provide state of knowledge quantifications. Then they go through the conceptual model together with the modeller, again associating names of experts with the identified uncertainties. Next, the analyst and the modeller search the mathematical model for uncertain model formulations and parameters and try to identify suitable experts. The input data file is checked next for possibly important uncertainty contributors. They also go through the numerical model and through the list of model options that are available to the client and note down any potentially important uncertainties in the choice of discretization schemes, parameters that control the numerical solution process, and in the choice of combinations of model options that might be used for the intended application.

Ideally, the analysis software is installed on a laptop computer and supports the documentation of the findings of this step online, i.e. during the meetings of analyst, client and modeller. The documentation should be organized such that the analyst can easily trace back the origin of an uncertainty and follow up where and how an uncertainty has been taken into account or why it was excluded from further consideration.

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_8

8.1.2 STEP 2: Quantify

The analyst will now, one by one, go through the list of uncertainties set up in Step 1 and contact the identified experts. In the case of an uncertain model formulation, the analyst and the expert will first arrive at a parameterization using the most appropriate method from those discussed in Sect. 3.4. For each uncertain datum, the analyst fills in the state of knowledge document discussed in Sect. 3.6.1.5. It consists of the cover sheet, the state of knowledge quantification and probabilistic modelling sheet and the state of knowledge dependence information sheet. The analysis software should offer a page for each sheet with pre-assigned places for the various entries. This way, the state of knowledge quantification and probabilistic modelling sheet can be completed in cooperation with the expert during the elicitation session (see Sect. 3.6), and graphs of the density and cumulative probability function of the subjective probability distribution fitted to the expert's state of knowledge quantifications can be immediately shown in order to obtain the expert's approval. Ideally, the software allows for the presentation of several fitted distributions in one figure so that they may be easily compared. The software should support all the distribution types discussed in Sect. 3.6.1.3 and possibly additional ones.

The software must number the uncertain data consecutively, and the analyst should be able to change this numbering, according to his preferences, with the software taking care of the needed adjustments. The software must also set up a list of the uncertain data showing number and short name and provide a link to the details of the corresponding state of knowledge modelling like type and parameters of the subjective probability distribution as well as state of knowledge dependences and their modelling.

The page for the state of knowledge dependence information sheet should provide for at least those modelling options that are discussed in Sect. 3.5. Some of them will require the entry of encoded instructions needed to compute the values of dependent uncertain data, and consequently a link to a compiler is required so that these instructions can be executed in Step 3 of the analysis. The software must also offer the option to show a scatterplot for any two or three state of knowledge dependent uncertain data so that the expert may immediately be able to judge the adequacy of his state of knowledge dependence quantification and modelling.

The analyst needs to indicate, on the cover sheet, whether the uncertain datum is already part of the model and whether the input file of the computer model application needs to be extended in order to allow the entry of values for this uncertain datum. Note, that all those data that are not yet part of the model require changes so that they may be read from the input file and processed by the computer model as intended.

Upon request, the software should print out the state of knowledge documents for all uncertain data or for a selection in a format suitable for documentation purposes. It should also be possible to restrict the print out to either the cover sheets or to the state of knowledge quantification and modelling sheets or to the state of knowledge dependence information sheets. Examples of the three types of sheets are given in Sect. 3.6.1.5.

Finally, after all uncertainties have been dealt with and the corresponding state of knowledge documents have been completed, the analyst presents the complete set of state of knowledge documents to each of the consulted experts. The purpose is to ask each expert whether any of the uncertain data with state of knowledge quantified by him might be state of knowledge dependent to an uncertain datum with its state of knowledge quantified by anyone of the other experts. Any state of knowledge dependence identified this way and judged to be potentially important needs to be quantified and probabilistically represented, e.g. by one of the ways discussed in Sect. 3.5. This has to be done by the analyst in cooperation with the two experts involved. The corresponding information sheets dealing with state of knowledge dependence (see Sect. 3.6.1.5) are then updated accordingly.

8.1.3 STEP 3: Propagate

All state of knowledge quantification and probabilistic modelling has been done in the previous step. Step 2 is the most laborious task of the whole analysis. Now it is the task of the software to collect the distribution and state of knowledge dependence information for all uncertain data from the respective sheets and to draw a random sample of N sets of M data values each. N is the sample size chosen by the analyst, and M is the total number of uncertain data (model uncertainties are represented by uncertain parameters and uncertain parameters are categorized as uncertain data—see Sects. 3.4 and 3.6.2). The analyst determines whether a simple random sample or a Latin Hypercube sample is to be drawn (see Sect. 4.4 for the difference). The software stores the sampled data values in an $N \times M$ array (N rows and M columns) for the performance of the corresponding evaluations (for short called "runs") of the computer model and for their further use in Step 5 of the analysis.

The software may provide a skeleton program offering the possibility to either enter the program statements of the encoded computer model directly by "copy and paste" or to call the model program which may be on another file on the laptop computer. The skeleton program will also read the sampled values from their array, row by row, and assign the values to the short names given on the cover sheets of the state of knowledge documents. The analyst only has to make sure that these names are those that are being used in the model code and that the required compiler is installed on the laptop computer. Using the skeleton would be the option to choose if the computer model is not too large and the run-time is not too long on the laptop computer. Otherwise the model would need to be run on an external machine. In this case, the client takes the array of sampled parameter values and provides a program that inserts the values in the right places of the input file of the computer model application.

There may be the possibility of running the model on a machine with many processors in parallel, each capable of executing the complete model or it may be possible to make use of a cluster of machines so that all N model runs can be performed in parallel. This approach requires the preparation of N copies of the input file.

The output from the N model runs is needed in Steps 4 and 5 of the uncertainty analysis in order to obtain the actual analysis results. Usually not every one of the many model results needs to be analysed. Some of those that do may be sequences of values of discretized functions over one or more independent variables (like time and/or space), and others may be single values like minima or maxima (over time and/or space) or fields of values (vectors or matrices, for instance). Whatever is supposed to be analysed will have to be stored by the client in a fashion suitable for further processing by the analysis software on the laptop. The analysis software should be capable of processing model output that is a mixture of all of the output types mentioned above. The output file is generated automatically by the software for further processing in Steps 4 and 5 if the model runs are executed under the program skeleton offered by the analysis software on the laptop.

The output file provided by the client has to contain the following information for every model result that is to be analysed:

- Number of the model result (from among the set of model results that are to be analysed).
- Short name of the model result. The name will be used in the graphical presentations of the analysis results.
- Type of result, i.e. single value, array of values or sequence of values of a discretized function.
- Independent variable of the sequence (if any).
- Measuring unit of the result.
- Measuring unit(s) of the independent variable(s).
- Value of the result (in the case of single values) or the sequence of pairs (or triplets) of values of the discretized function [value of the result, value(s) of the independent variable(s)].
- Sequence of specified values of the independent variable(s) at which the result should be analysed and graphically presented.

Some of the N runs of the computer model may fail. For this reason, it will be of advantage to include a model output that receives the value 1 if the run fails and 0 otherwise. An uncertainty importance analysis (Step 5) using this output may help to pinpoint the reason for the failure. After the computer model has been improved accordingly or the state of knowledge quantifications have been revised, a new sample of parameter values needs to be selected and the N computer model runs will need to be redone. The analysis software should, however, offer the option to perform Steps 4 and 5 of the analysis with selected model runs, for instance with only those computer model runs that could be completed successfully.

8.1.4 STEP 4: Estimate Uncertainty

This step can be fully executed by the analysis software. The analyst only needs to select from among the available options like those described in Chap. 5.

8.1.5 STEP 5: Uncertainty Ranking

This step too can be fully performed by the analysis software with the analyst only required to choose from the available options. His choice will most likely include those options described in Chap. 6.

8.1.6 STEP 6: Present and Interpret

In this step, the analyst discusses the analysis results together with the client and the modeller. The analyst will adopt the ways of interpretation presented in Chap. 7, and the client will need to judge whether the quantified uncertainty of the model results is acceptable for the purpose of the computer model application. If it is not acceptable, the analyst will consult the uncertainty importance measures together with the modeller in order to identify candidates among the uncertain data that promise sufficient potential for state of knowledge improvement in order to effectively reduce the uncertainty of the model result in question. The modeller on the other hand will look at the uncertainty importance measures from the perspective of their reasonableness. Any measure that seems suspiciously small or large needs to be scrutinized whether it may be due to errors or insufficiencies of the computer model or of the scenario description or may be due to an erroneous transfer of the data values into the input file of the computer model application. It could, however, also be due to the inadequacy of the uncertainty importance measure chosen.

The modeller may also be interested in the run-time differences between the N model runs and may for this reason include the runtime as an additional model output that is to be subjected to the uncertainty importance analysis. The uncertainty importance measures may provide an indication of where the computer model could be improved in order to make it more economical to run.

Once satisfied with the analysis results, it has to be decided by the analyst and the client which information is to be included in their documentation. They also need to decide about those analysis results that should be given to decision-makers together with the model results. The documentation of the analysis results may adopt the structure suggested in Sect. 7.1.

8.2 Comparison of Four Software Packages

SUSA (Kloos 2015) is a menu-driven software for uncertainty analysis that satisfies most of the requirements mentioned above. It has been used, among numerous other occasions, in the practical examples of Chap. 10. SUSA 3.6 runs on personal computers under the operating system MS-WINDOWS. It makes use of MS-EXCEL only for the purpose of generating graphics. A FORTRAN compiler is required if the analyst expresses some state of knowledge dependences by FORTRAN encoded functional relationships between uncertain data and if the (FORTRAN encoded) computer model is to be evaluated under the control of SUSA using the provided program skeleton that is encoded in FORTRAN. Otherwise the computer model can be written in any programming language as only the output file with the model results, that are to be analysed, needs to be provided for further processing by SUSA. SUSA 4.0 (Kloos 2015) does not make use of MS-EXCEL but has a graphical user interface (GUI) written in Visual Basic.NET (vb.net). The GUI is menu-driven and guides through the steps of the analysis. SUSA 4.0 is available for 32 bit and 64 bit computer systems. The free GNU FORTRAN compiler is automatically installed with SUSA 4.0.

It is important to note that the computer model does not need to run on the personal computer since SUSA provides for all necessary data exchanges between an external machine and the personal computer.

The only limitations of SUSA 3.6 are those of memory space available for the operation of SUSA and the limitations of MS-EXCEL. The latter concern mainly the graphical presentation of data series in the case of index-dependent model results (model output that is a sequence of values of a discretized function over some independent variable). Specifically:

- An index-dependent model result can only be presented for the maximum number of 255 model runs in one chart.
- The maximum number of index-dependent values per model run to be presented in a chart is 32,000.
- The maximum number of data points of all model runs to be presented in a chart is 256,000.

Although the output from computer models may be in the order of MBs or GBs, the limitations listed above are usually of no concern since the results actually required to answer the assessment question, and therefore to be analysed, are generally rather limited in number.

None of the limitations listed above apply to SUSA 4.0.

Other well-known analysis software systems are @RISK, Cristal Ball and ModelRisk. Descriptions of their capabilities can be obtained from their websites. Neither SUSA nor any of the other systems offer all the options discussed in Chaps. 2–7. There are differences in the number of distribution types supported as well as in the kind of input accepted for state of knowledge quantification. Also, none of the systems provides all options for state of knowledge dependence

modelling. Only ModelRisk offers several copula functions. With respect to sampling, there is basically no difference between SUSA, @RISK and Cristal Ball as they all support Latin Hypercube (LHS) (with the exception of ModelRisk) and simple random sampling (SRS). None of them supports importance sampling or subset sampling. Only SUSA computes statistical tolerance limits and intervals in Step 4 of the analysis. SUSA is also the only system that computes the correlation ratio (square root of the approximate first-order uncertainty importance measure from variance decomposition). It does so without requiring model evaluations additional to those used for the uncertainty analysis of the model result.

While @RISK, Cristal Ball and ModelRisk offer more or less the same support for the analysis steps as SUSA, there is, however, one important difference. Crystal Ball, @RISK and ModelRisk are MS-EXCEL based and are therefore limited to models that can be formulated using spreadsheets. They are therefore not suited for the analysis of computer model applications where the model is written in FORTRAN or any other programming language and comprises thousands or tens of thousands of lines of code. SUSA on the other hand expects the model to be entered in FORTRAN into the offered template either directly or via the call function of FORTRAN. For models running not under the control of SUSA, comfortable data exchange options are provided. In this case, the model may run on any other computer (even in parallel on a system of compute nodes) and may be encoded in a programming language other than FORTRAN. This way there is no limitation to the size of the computer model application that is to be analysed with SUSA.

Reference

Kloos, M. (2015). *Main features of the tool SUSA 4.0 for uncertainty and sensitivity analyses*. ESREL 2015, European Safety and Reliability Conference, Zürich.

Chapter 9
Uncertainty Analysis When Separation of Uncertainties Is Required

9.1 Introduction

Some assessment questions are not fully specific and therefore have not only one true answer but a population of true answers (see Chap. 1). Their wording does not include a specific but only a generic reference unit. Examples of generic reference units are for instance:

– . . . Per throw of the die.
– . . . Per year of operation of the technical facility X.
– . . . Per repetition of experiment XYZ.
– . . . Per individual of the exposed population.
– . . . Per N individuals of the exposed population.

Specific reference units would be for instance:

– . . . The next throw of the die.
– . . . The next 20 years of operation of the technical facility X.
– . . . The next repetition of experiment XYZ.
– . . . The individual z_i from the exposed population.
– . . . The individuals $z_1,\ldots,z_N$ from the exposed population.

While the specific reference unit permits only one true answer, variables that change their value with every realization of the generic reference unit lead to a population of true answers. As long as these realizations are not specified, the variables are stochastic variables and their values are aleatoric uncertainties of the assessment. Due to these aleatoric uncertainties, it cannot be known which answer to give for an unspecified (i.e. any) realization of the reference unit. In what follows, the analysis that quantifies aleatoric uncertainties of the assessment will be called "analysis of Type A uncertainty". It arrives at a probability distribution that summarizes the population of true answers. This probability distribution is then the answer to the assessment question with generic reference unit. It is a probabilistic

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_9

answer. This probabilistic answer is the result of the application of a computer model that quantifies the aleatoric uncertainty, i.e. performs the analysis of Type A uncertainty.

Analyses of Type A uncertainty make use of random laws that model the stochastic variability and thus express quantitatively the aleatoric uncertainties. The applicable models and the appropriate values of their parameters are fixed yet unknown or imprecisely known. They are therefore epistemic uncertainties of the Type A analysis result. Together with other epistemic uncertainties, they are the subject of an epistemic uncertainty analysis of the probabilistic answer to the assessment question. This epistemic uncertainty analysis will subsequently be referred to as "analysis of Type B uncertainty".

Many computer models are designed to specifically quantify uncertainty due to stochastic variability, i.e. they include the analysis of Type A uncertainty. Among them are all probabilistic risk or safety assessments as well as reliability analyses of any kind of system. Their results are estimates of the probability distributions that summarize the population of true answers as well as estimates of their characteristics like mean value, standard deviation, quantile values, cumulative and complementary cumulative probabilities such as the probability (now in its frequentistic interpretation) that a safety limit is exceeded by a true answer to the assessment question.

The analysis of the epistemic uncertainty of these results runs as described in Chaps. 2–7 except for some differences within each of the six analysis steps. Before the differences are discussed, some remarks are in place. They concern notation and the two kinds of analysis output that are connected to the two types of uncertainty.

The following notation is being used in this chapter:

$X_1,\ldots,X_J$ the J results produced by an application of the model, each having a population of true values; they are the J components of the vector $\boldsymbol{X}$.

$g_1,\ldots,g_J$ the corresponding J functions; they are the J components of the function vector $\mathbf{g}$.

$Y_1,\ldots,Y_I$ the I results produced by an application of the model, each having only one true value; they are the I components of the vector $\boldsymbol{Y}$.

$h_1,\ldots,h_I$ the corresponding I functions; they are the I components of the function vector $\mathbf{h}$.

$A_1,\ldots,A_K$ the K aleatoric uncertainties (stochastic variables) used by some or all of the J components of the function vector $\mathbf{g}$; they are the K components of the vector $\boldsymbol{A}$.

$P_1,\ldots,P_M$ the M epistemic uncertainties used by some or all of the $J + I$ functions. They are the M components of the vector $\boldsymbol{P}$ and their state of knowledge is expressed by a joint subjective probability distribution.

For a specific choice $\boldsymbol{P} = \boldsymbol{p}$ of values for the epistemic uncertainties, one arrives at the J-dimensional random variable

$$\boldsymbol{X}_{\boldsymbol{P}=\boldsymbol{p}} = g(\boldsymbol{A}|\boldsymbol{P} = \boldsymbol{p}).$$

A possibly true summary of its population of values is given by the possibly true joint probability distribution $\mathbf{F}(\boldsymbol{X}|\boldsymbol{P} = \boldsymbol{p})$. This distribution quantifies the aleatoric uncertainty that is due to the stochastic variables $\boldsymbol{A}$ given $\boldsymbol{P} = \boldsymbol{p}$, while

$$\boldsymbol{X} = \mathbf{g}(\boldsymbol{A}, \boldsymbol{P})$$

defines a population of possibly true J-dimensional random variables and therefore a population of possibly true joint probability distributions for $X_1, \ldots, X_J$ that logically results from the state of knowledge quantifications for the epistemic uncertainties $\boldsymbol{P}$. Figure 9.1 shows a small sample from the population of possibly true probability distributions for the one-dimensional (i.e. $J = 1$) random variable X. For every quantile percentage value q, there is a population of possibly true quantile values $x_{q\%}$ that can be summarized by a subjective probability distribution for the quantile $x_{q\%}$. This is illustrated in Fig. 9.2 where the sample of $q(=50)\%$ quantile values is taken from Fig. 9.1.

There is also a subjective probability distribution for the mean value $\mathrm{E}_{\mathbf{A}}\{X\}$ ($\mathrm{E}_{\mathbf{A}}\{.\}$ denotes the mean taken over the stochastic variability) as well as for the probability $\mathrm{w}(X > x^*)$ that X exceeds a given limit value x^* and for the probability w $(X \leq x^*)$ that X does not exceed x^*.

The model results $X_{1,\,\boldsymbol{P}=\boldsymbol{p}}, \ldots, X_{J,\,\boldsymbol{P}=\boldsymbol{p}}$ will share some or all of the influences of the stochastic variables $\boldsymbol{A}$. Consequently, there will be stochastic dependence among

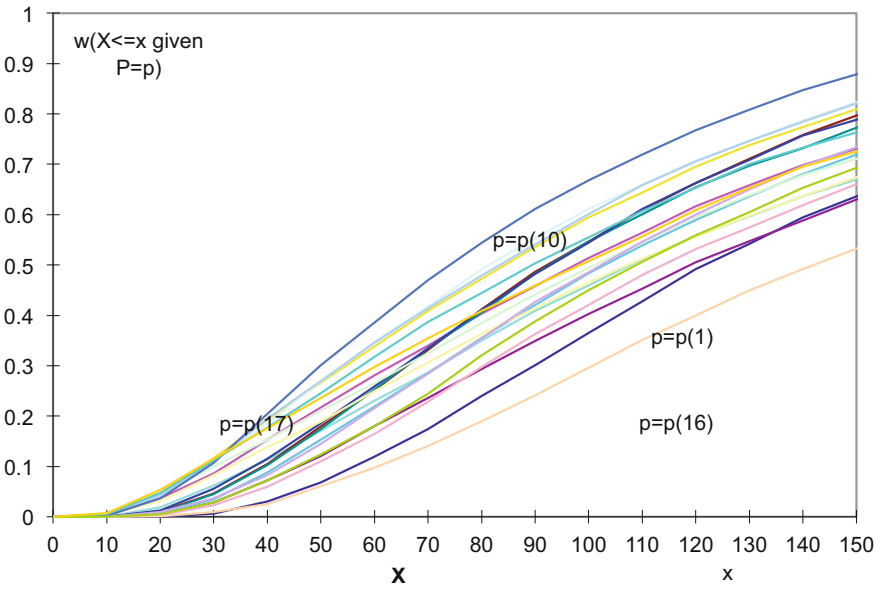


Fig. 9.1 Random sample of 20 conditional cumulative distribution functions (cut off at $X = 150$) each quantifying Type A uncertainty under the condition that $\boldsymbol{P} = \boldsymbol{p}_n$, $n = 1,\ldots,20$. The sample is output from the analysis of Type B uncertainty of the results from the analysis of Type A uncertainty

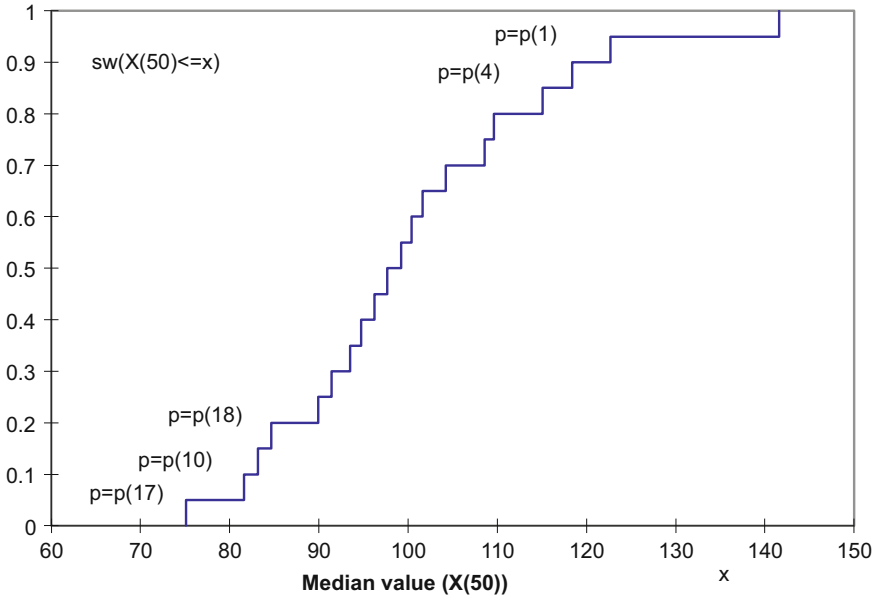


Fig. 9.2 Empirical subjective probability distribution of the sample of twenty 50% quantile values in Fig. 9.1; it is obtained by the analysis of Type B uncertainty of the results of the analysis of Type A uncertainty. The random sample of cumulative distribution functions, each a possibly true quantification of Type A uncertainty, is shown in Fig. 9.1

the components of $\boldsymbol{X}$. This dependence leads to a matrix of pairwise correlation coefficients. It tells whether there are any strong concurrent or counter-current pairwise linear or monotone (if rank correlation coefficients are chosen) relationships between the model results $X_1, \ldots, X_J$ due to the common aleatoric uncertainties**.** Obviously, the entries of this correlation matrix are subject to epistemic uncertainty, i.e. their state of knowledge is expressed by a joint subjective probability distribution that is one of the results of the analysis of Type B uncertainty.

On the other hand, since some or all of the uncertain data $\boldsymbol{P}$ enter the computations of the quantiles $x_{j,\,q\%}, j = 1, \ldots, J$, there will be state of knowledge dependence between these quantiles. The same applies to the mean values $\mathrm{E}_{\mathbf{A}}\{X_j\}$, as well as to the probabilities $\mathrm{w}(X_j \leq x_j^*)$ and $\mathrm{w}(X_j > x_j^*)$ for $j = 1, \ldots, J$.

It follows now the discussion of the differences between an analysis of Type B uncertainty of the results from the application of a model that does not require separation of uncertainties as compared to one that does require separation. The discussion goes through Steps 2–7 of the Type B uncertainty analysis. It points out differences that are due to the fact that the model application now quantifies aleatoric (Type A) uncertainty in order to answer the assessment question.

9.2 STEP 1: Search

Anything that contributes to the stochastic variability of $\boldsymbol{X}$ needs to be expressed by random variables $A_1, \ldots, A_K$ (Type A uncertainties), and their corresponding random laws are needed in the analysis of Type A uncertainty. They are therefore part of the computer model. The random laws are sometimes only assumptions or are fitted to empirical distributions summarizing observations. While their functional forms are model uncertainties, their parameter values are parameter or data uncertainties. Both are to be handled within the analysis of Type B uncertainty. In addition, there may be stochastic dependences between the random variables $A_1, \ldots, A_K$. They need to be expressed by suitable relationships (will be model uncertainties) or measures of association (will be data uncertainties) as part of the computer model.

9.3 STEP 2: Quantify

The same procedures that were discussed in Chap. 3 for model uncertainties in general apply to the quantification of the state of knowledge of the uncertain random laws expressing aleatoric uncertainty. The random laws model stochastic variability and therefore the quantification of the state of knowledge of their parameter values can be achieved by using the likelihood function of available observations and by applying the Bayesian method.

The Bayesian method plays a prominent role in the specification of subjective probability distributions for the uncertain parameters of random laws. It uses the likelihood of realizations of the respective random variables given specific values for the uncertain parameters. In his essay "Towards solving a problem in the doctrine of chances", (Bayes 1958) Thomas Bayes explored more than 250 years ago how to rationally use observation for an update of the state of knowledge of the probability of an event if nothing is known about this probability a priori. His method requires a relation between the observation and the uncertain parameter in question, i.e. the likelihood of the observation, given specific parameter values.

For instance, the method provides for the parameter P of the binomial distribution the subjective probability density function (state of knowledge quantification)

$$\mathrm{f}_1(p|m;n) = \pi(m;n|p)\mathrm{f}_0(p) \Big/ \int_0^1 \pi(m;n|p)\mathrm{f}_0(p)\mathrm{d}p \tag{9.1}$$

with

P	the uncertain parameter;
$\mathrm{f}_0(p)$	the subjective probability density function that expresses the a priori (before the observations became available) state of knowledge of the true value of P;

$\pi(m; n\|p)$	the probability of the observation (m;n) given the true value of P is p, i.e. the probability to observe m failures in n trials if p is the true failure probability per trial. It is the value of the likelihood of $P = p$ given (m;n);
$\int_0^1 \pi(m; n\|p) f_0(p) dp$	the subjective unconditional (not conditioned on any specific value of P) probability to observe m failures in n trials. It serves as normalizing constant;
$f_1(p\|m; n)$	the subjective probability density function that expresses the a posteriori (after the observation became available) state of knowledge of the true value of P.

Formula (9.1) is also known as Bayes' theorem in classical statistics. There, however, it is not applied to an uncertain fixed parameter value (epistemic uncertainty) but to conditional probabilities of aleatoric uncertainties. Bayes did explicitly consider the case of a fixed but unknown parameter value, and he expressed his state of knowledge by probability. The wording of the problem formulation in his essay[1] uses the term "chance" (later in the essay it is called "the chance to be in the right"). It makes clear that a specific interpretation of the term "probability" is required if Eq. (9.1) deals with a fixed but unknown parameter value.

Because of his pioneering use of the subjectivistic interpretation of probability, it has become customary to use the term "Bayesian" whenever subjective probabilities come into play.

In his essay, Bayes chose the density of the uniform distribution over the interval (0, 1) as $f_0(p)$, i.e. as expression of the a priori state of knowledge (nothing known prior to the observation). What if somebody decides to work with the parameter Q, where $Q = P^{1/2}$? If this person uses the uniform distribution for Q over (0, 1) as the expression of the prior state of knowledge (nothing known a priori), then a non-uniform distribution follows for P through transformation. This is a contradiction to the assumption "nothing known about P a priori". Its root cause was most likely the reason why the essay was not published by Bayes himself but was submitted to the Royal Society by a friend after Bayes' death. The contradiction can be avoided by the use of non-informative a priori subjective probability distributions. Approximate non-informative distributions are obtained using Jeffrey's rule (Box and Tiao 1973). To this end, one determines a metric (transformation of P) in which the shape of the likelihood function of the observation does not depend on the actual observation; however, the position of its maximum value on the parameter axis in this metric does. This achieves a separation of the information contained in the likelihood function into two parts. One part contains what is known about the underlying chance mechanism (i.e. Bernoulli law, etc.) and is independent of the actual observation, and the other part contains the information from what was actually observed. The first part

[1] "Given the number of times in which an unknown (nothing known a priori) event has happened and failed: Required the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named".

In definition no. 6 of his essay, it says: "By chance I mean the same as probability".

can serve as a priori information. This is achieved by choosing a uniform prior for the transformed parameter, just as Bayes did in the original metric. Back transformation of the uniform prior into the original metric (P) then provides the non-informative a priori subjective probability density function $f_0(p)$.

The following examples refer to technical components. The underlying random laws, however, apply to Type A uncertainties in a wide range of computer models.

Examples:

– P is the failure probability per demand of a technical component

The stochastic variable (aleatoric uncertainty) is the random variable X that assumes only the values 1 (if the component fails on demand) or 0 (if it does not fail). The corresponding random law is assumed to be the Bernoulli distribution with the uncertain parameter P. The non-informative a priori density function $f_0(p)$ is proportional to $p^{-1/2}(1-p)^{-1/2}$ (Box and Tiao 1973), and the likelihood is the Binomial probability $\pi(m; n|p) = [n!/m!(n-m)!]\ p^m(1-p)^{n-m}$ if m failures are observed among n demands. There are several good reasons for choosing a distribution of the Beta type to express the state of knowledge for P. The density function[2] of a Beta distribution over the interval (0, 1) is $f(p) = \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] p^{\alpha-1}(1-p)^{\beta-1}$. From Bayes' method, one obtains the a posteriori Beta density function $f_1(p)$ with parameters $\alpha = m+1/2$ and $\beta = (n-m) + 1/2$. The mean value is $\alpha/(\alpha+\beta)$, and Beta distribution quantile values can be obtained from the relationship between the Beta and the F-distribution. For instance, the 90%-quantile $p_{90\%}$ of the Beta distribution for P over (0, 1) with parameters (α, β) is $\alpha w_{90}/(\beta + \alpha w_{90})$. w_{90} is the 90%-quantile of the F-distribution with degree of freedom $(2\alpha, 2\beta)$ (See Table IV in Winkler and Hays (1975)).

– P *i*s the failure rate per hour of operation of a technical component

The stochastic variable is the random variable X that assumes any value larger than 0. It is the operation time of the component until it fails. The corresponding random law is often assumed to be the exponential distribution with the uncertain parameter $P = \Lambda$, the failure rate per unit time of operation. The non-informative a priori density function $f_0(p)$ is proportional to $\lambda^{-1/2}$ (Box and Tiao 1973) and the likelihood is the Poisson probability $\pi(s;\ t|\lambda) = (\lambda t)^s e^{-\lambda t}/s!$. s is the number of failures observed in the total operation time t. s and t are usually accumulated over a pool of components that are assumed to have the same failure rate. There are several good reasons for choosing a distribution of the Gamma type to express the state of knowledge about the uncertain parameter "failure rate". The density function[3] of Gamma distributions is $f(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)}$. From Bayes' method, one obtains the a posteriori Gamma density function $f_1(\lambda)$ with parameters $\alpha = s + 1/2$ and $\beta = t$.

[2]Caution: Some write the density function of a Beta distribution with parameters α and $\alpha + \beta$.

[3]Caution: Some write the density function of a Gamma distribution with parameters α and $1/\beta$.

The mean value is α/β, and Gamma distribution quantile values can be obtained from the relationship between the Gamma and the χ^2 distribution. For instance, the 90%-quantile $\lambda_{90\%}$ of the Gamma distribution for Λ with parameters (α, β) is $w_{90}/2\beta$. w_{90} is the 90%-quantile of the χ^2 distribution with degree of freedom 2α (See Table III in Winkler and Hays (1975)).

Remarks to the examples:
The chosen type of the a priori distribution is conjugate to the likelihood function in both cases, meaning that the a posteriori distribution is of the same type as the a priori distribution, and its parameters are easily obtained from those of the a priori and from the observation.

The observation of an individual technical component will most likely not provide sufficient information to obtain reasonable results through Bayes' theorem. Therefore, the components are often grouped such that those in a group show sufficiently close failure behaviour under the conditions considered by the model application. The observations collected for the components in a group are then aggregated or pooled and used in Bayes' theorem to obtain a subjective probability distribution for P, the failure probability per demand, or for Λ, the failure rate per unit time of operation, common to all components in the group. As a logical consequence of this, the model application has to use the same value of P (resp. Λ) for all components in the group (Apostolakis and Kaplan 1981). This is sometimes called "failure probability coupling" or "failure rate coupling". It is a case of complete state of knowledge dependence (see Sect. 3.5) between the failure probabilities (or rates) of the components in the group. The uncertainty of how to define the group boundaries may be seen as a model uncertainty.

The a priori state of knowledge of almost complete ignorance, represented by the non-informative distribution, is usually rare. There are always relevant bits of information about the fixed true but unknown parameter value, before the observation became available. However, it will often be hard to cast them in a defendable manner into an a priori density $f_0(p)$. For this reason, non-informative a priori densities are preferred. The influence of the a priori distribution on the a posteriori distribution decreases with the increase of the amount of available observation.

For reliable components, it is not uncommon to have none or at most one failure ($m = 0$ or 1) in n demands. If, in addition, n is small despite the grouping of components (small compared to the true inverse value of the failure probability per demand), then $f_1(p|m; n)$ may only be adequate if prior knowledge is taken into account.

Observations are also subject to epistemic uncertainties. Firstly, they are sometimes difficult to interpret. Was it really a failure or only close to a failure? Secondly, it may be uncertain whether the failure observed in one group could also occur in the group and under the conditions the model application is concerned with. Last but not least, the question remains whether all failures within the accumulated number of demands or operation times have been reported.

The simplest form of prior knowledge is based on plausibility considerations. Such considerations provide, for instance, justifiable limits beyond which the fixed

true yet unknown parameter value cannot lie. Such limits will be specifiable in all instances where logical upper and/or lower limits apply and where the true values of uncertain data could be found through measurement, albeit subject to measurement error. Additional information may justify not to use a uniform distribution over the interval defined by the limits but to choose a subjective probability distribution that exhibits some characteristic behaviour towards the endpoints of the interval. Example would be a triangular or truncated normal or a lognormal distribution among others (see Sect. 3.6.1.3). The choice of a distribution type will need to observe the maximum entropy principle (Buckley 1985).

The specification of limit values (truncations of the a priori distribution) means that the a priori density $f_0(p)$ is chosen to be zero over those values of P that lie outside the limit values. The a posteriori density $f_1(p|m; n)$ obtained from Eq. (9.1) can then never be non-zero over those parts no matter what the observation $(m;n)$ says that became available after the specification of the a priori distribution. Obviously, such limit values must be well justified in order to be defendable.

The same procedures that are discussed in Sect. 3.5 are available for the specification of any state of knowledge dependences between the uncertain distribution functions and their parameters used to quantify Type A uncertainty as well as for any state of knowledge dependence to any of the other Type B uncertainties of the model application.

All remarks made on expert judgement elicitation in Sects. 3.6 and 3.7 are equally valid for the quantification of the state of knowledge of all the uncertain model formulations and parameters needed to quantify Type A uncertainty.

9.4 STEP 3: Propagate

In some situations, it may be possible to obtain the probability distributions of the model results $X_1, \ldots, X_J$ analytically. These distributions summarize the Type A or aleatoric uncertainty that is due to stochastic variability. The mean values, standard deviations, quantile values and cumulative or complementary cumulative probabilities of interest are then also available as single valued results of the analysis of Type A uncertainty. The analysis of Type B or epistemic uncertainties of these expressions of Type A uncertainty is then performed as described in Chaps. 2–7.

9.4.1 Two Nested Monte Carlo Simulation Loops

Generally, the computer model will be too complex to use an analytical approach for the analysis of Type A uncertainty. Monte Carlo simulation may, therefore, already be part of the model or the model may be supplemented by a Monte Carlo simulation to obtain an estimate of a possibly true joint probability distribution that summarizes the stochastic variability of $X_1, \ldots, X_J$ and thus quantifies the Type A uncertainty.

Together with the Monte Carlo simulation that is to propagate the Type B uncertainty, two nested Monte Carlo simulation loops are therefore to be executed. This is sometimes called a double randomization or two stage sampling approach or a two-loop (2LMC, see Fig. 9.3) or two-dimensional (2DMC) Monte Carlo simulation. In the outer loop, one value is sampled for each of the Type B or epistemic uncertainties $P_1, \ldots, P_M$ according to their joint subjective probability distribution that quantifies their state of knowledge. Using these values, the inner loop (it may already be part of the computer model) samples one value each of the Type A or aleatoric uncertainties $A_1, \ldots, A_K$. The approximate possibly true joint probability

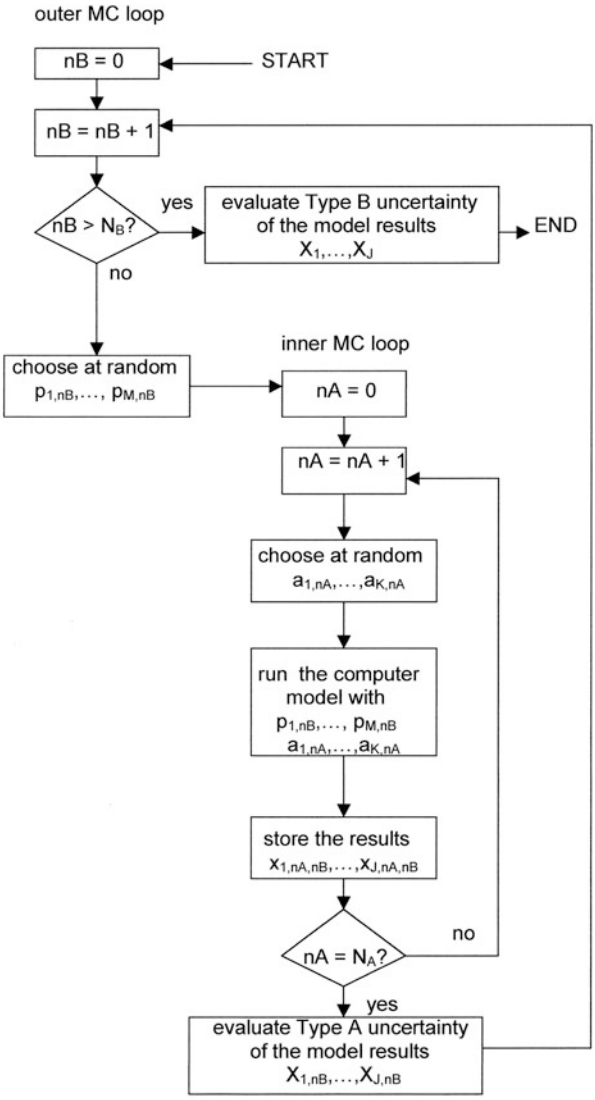


Fig. 9.3 The two nested Monte Carlo simulation loops required for the analysis of Type A and Type B uncertainty

distribution of the model results $X_1, \ldots, X_J$ is then computed. It is an empirical summary of the population of true values that is due to the stochastic variability of $A_1, \ldots, A_K$, given the values $\boldsymbol{P} = \boldsymbol{p}$ chosen for the epistemic uncertainties in the outer loop. The inner loop also computes estimates of the distribution characteristics of interest like mean values, standard deviations, quantile values and cumulative or complementary cumulative probabilities and all other values that are output of the computer model. The outer loop runs over the sample size N_B chosen for the analysis of Type B uncertainty while the number of runs performed by the inner loop corresponds to the sample size N_A chosen for the analysis of Type A uncertainty. See Sect. 4.4 for sampling techniques that might be used in these Monte Carlo simulations.

9.4.2 Low Probability Extreme Value Answers

The population of true answers to risk-related assessment questions with generic reference unit often contains a small subset of low probability and extreme values for some indices $j \in [1, J]$. This subset is usually of particular concern to decision-makers. The probability distribution summarizing the variability within the population of true answers is therefore required to represent this subset adequately. To this end, the Monte Carlo simulation of the inner loop in Fig. 9.3 would require a very large sample size N_A. Even a system of many compute nodes, each capable of processing the full set of functions **g** and **h** of the computer model, would often require a prohibitively long time to perform the N_A model runs needed. In addition, N_B repetitions of the N_A runs are to be performed to quantify the influence of the epistemic uncertainties on the estimation of the probability distribution summarizing the population of true answers. For this reason, analysts frequently resort to a so-called scenario analysis in conjunction with subset sampling. In the context of risk assessment for an industrial plant, the first set (C_{Y_0}) would contain the population of all accident initiating events. This population is divided into subpopulations, and representative members are selected from each subpopulation (often those judged to have the potential of leading to the worst outcome from among those in the subpopulation). The first subset $(C_{Y_1} \subset C_{Y_0})$ would contain all those members of C_{Y_0} that are followed by some safety system failure. Again, division into subpopulations and choice of representative members from each leads to the next subset $(C_{Y_2} \subset C_{Y_1} \subset C_{Y_0})$. C_{Y_2} would contain all those members of C_{Y_1} followed by plant damage (i.e. all sequences "accident initiating event → safety system failure → plant damage"). Again, division into subpopulations and choice of representative members from each leads to subset $(C_{Y_3} \subset C_{Y_2} \subset C_{Y_1} \subset C_{Y_0})$ where C_{Y_3} would contain all those members of C_{Y_2} followed by contamination of the plant environment. Since the populations of the subsets are not known, the subpopulations are indirectly defined by choosing first a set of representatives such that the corresponding model outcomes may be considered a satisfactory

representation of the model outcomes of the subpopulations, and the union of these model outcomes may be thought of as an adequate representation of the (unknown) population of model outcomes for this subset. Following this scenario approach through all subsets requires only the run of those parts of the computer model that are needed by the chosen sequences of representatives, i.e. the scenarios. Additionally, the conditional probabilities of the subpopulations in each subset, each represented by chosen representative members, are to be determined. This is usually achieved by using a mixture of operating experience, expert judgment, analytic modelling and Monte Carlo simulation. The model outcomes obtained for the scenarios together with the conditional probabilities assigned to the subpopulations of the subsets are then summarized in the form of a probability distribution that is assumed to satisfactorily quantify the Type A uncertainty, given the set of values $\boldsymbol{P} = \boldsymbol{p}$ chosen for the epistemic uncertainties in the outer Monte Carlo loop.

Considering the need of the scenario approach to choose representatives from the subpopulations of the various subsets, it is not surprising that a number of efforts have been and are still being undertaken to improve the situation (Siu 1994; Cojazzi 1996; Devooght and Smidts 1996; Hsueh and Mosleh 1996; Chang and Mosleh 1998; Labeau et al. 2000; Hofer et al. 2002; Kloos and Peschke 2006), i.e. to replace the assumptions inherent in the choice of representatives by as much realism as is possible with the presently available computing power. These efforts are concentrated on the analysis of the Type A uncertainty as it is to provide the probabilistic answer to the assessment question. The associated Type B uncertainty analysis by Monte Carlo simulation with sample size N_B multiplies the required computing power by N_B. Therefore, approximate approaches to the Type B uncertainty analysis have been suggested.

Changing the generic reference unit of the assessment question into a specific one (for instance, replacing "per repetition of experiment XYZ" by "the next repetition of experiment XYZ") turns aleatoric uncertainties into epistemic uncertainties. Then, only the analysis of Type B uncertainty is required. The problem of "low probability and extreme value answers" does, however, not go away. It will become one of "low subjective probability possibly true answers of extreme value". The importance sampling and the subset sampling strategies discussed in Sects. 4.4.3 and 4.4.4 are designed to reduce the sample size required for these problems.

9.5 STEP 4: Estimate Uncertainty

The joint probability distribution $\mathrm{F}(\boldsymbol{X}|\boldsymbol{P} = \boldsymbol{p})$ for

$$\boldsymbol{X}_{\boldsymbol{P}=\boldsymbol{p}} = \mathbf{g}(\boldsymbol{A}|\boldsymbol{P} = \boldsymbol{p}),$$

is one possibly true joint probability distribution for $\boldsymbol{X} = X_1, \ldots, X_J$. It quantifies the aleatoric uncertainty of $\boldsymbol{X}$ that is due to the aleatoric uncertainties $\boldsymbol{A}$, given $\boldsymbol{P} = \boldsymbol{p}$.

The marginal distributions $F_j(X_j|\boldsymbol{P} = \boldsymbol{p})$, $j = 1, \ldots, J$, and their distribution characteristics are among the results of the computer model application.

The analysis of the epistemic uncertainties on the other hand is to quantify the state of knowledge of the joint probability distribution and thus of the marginal distributions and of their characteristics like mean value, standard deviation, cumulative probability and complementary cumulative probability as well as pairwise correlations. Their subjective probability distributions result from the propagation of the quantified state of knowledge at the level of uncertain model formulations, parameters and input data through the computer model application that quantifies the Type A uncertainty. Consequently, the relationships

$$\boldsymbol{X} = \mathbf{g}(\boldsymbol{A}, \boldsymbol{P})$$
$$\boldsymbol{Y} = \mathbf{h}(\boldsymbol{P})$$

define a population of possibly true joint probability distributions for $X_1, \ldots, X_J$ as well as a population of possibly true values for all other model results $\boldsymbol{Y}$. These populations are generated by the state of knowledge expressions for the epistemic uncertainties $\boldsymbol{P}$ propagated through the encoded computer model including the analysis of Type A uncertainty.

As an approximate expression of the state of knowledge of the true joint probability distribution of $\boldsymbol{X}$ and of the true vector $\boldsymbol{Y}$, the Monte Carlo simulation (outer loop) provides a sample of joint probability distributions and of vectors $\boldsymbol{y}$, drawn from the population of possibly true distributions and vectors. From this sample, one derives Type B uncertainty statements for all model results. These results include the marginal distributions of $X_1, \ldots, X_J$ and their characteristics of interest as well as all other model results that are components of $\boldsymbol{Y}$. The Type B uncertainty statements are obtained in the same way as was discussed in Chap. 5. Some useful graphical presentations of results from the analysis of Type A (inner simulation loop) and of Type B (outer simulation loop) uncertainty are shown below for a one-dimensional (i.e. $J = 1$) model output X.

Figure 9.4 shows the empirical (obtained from N_A sample values) cumulative distribution function of X as obtained by the inner simulation loop with the reference values of the epistemic uncertainties $\boldsymbol{P}$. Often, the complementary cumulative distribution function is of interest. Figure 9.5 shows the probabilities

$$\mathrm{w}\left(X > x|\boldsymbol{P} = \boldsymbol{p}_{\boldsymbol{Ref}}\right) = 1 - \mathrm{w}\left(X \leq x|\boldsymbol{P} = \boldsymbol{p}_{\boldsymbol{Ref}}\right),$$

to be read from the complementary cumulative distribution function. They are the probabilities that X, as computed with the reference values of the epistemic uncertainties $\boldsymbol{P}$, exceeds given limit values x.

The range of probabilities and values of X_j will extend over orders of magnitude if there is a subset of low probability and extreme value answers. Consequently, the tail end of the ccdf stretches along the abscissa over a wide range of values x while the corresponding complementary cumulative probabilities are hardly distinguishable

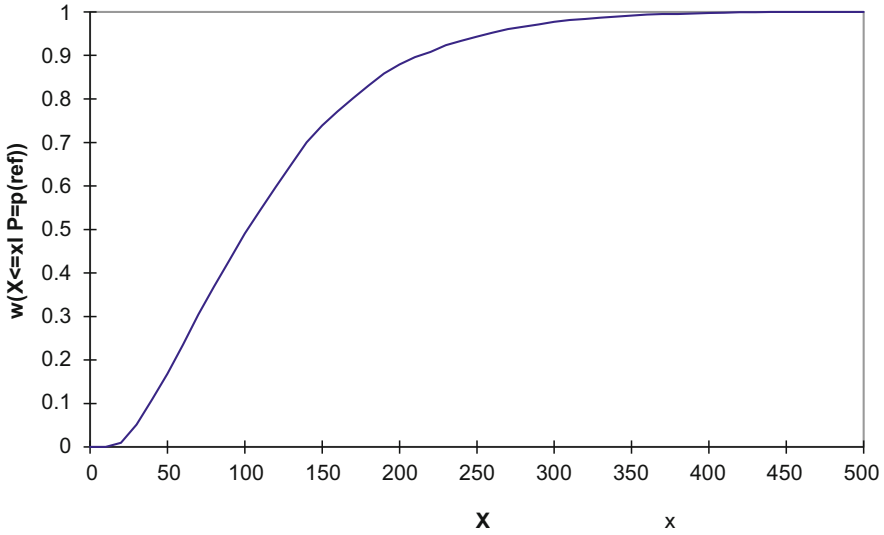


Fig. 9.4 Empirical cumulative distribution function (cdf) of X obtained in the inner simulation loop and using the reference values of $\boldsymbol{P}$

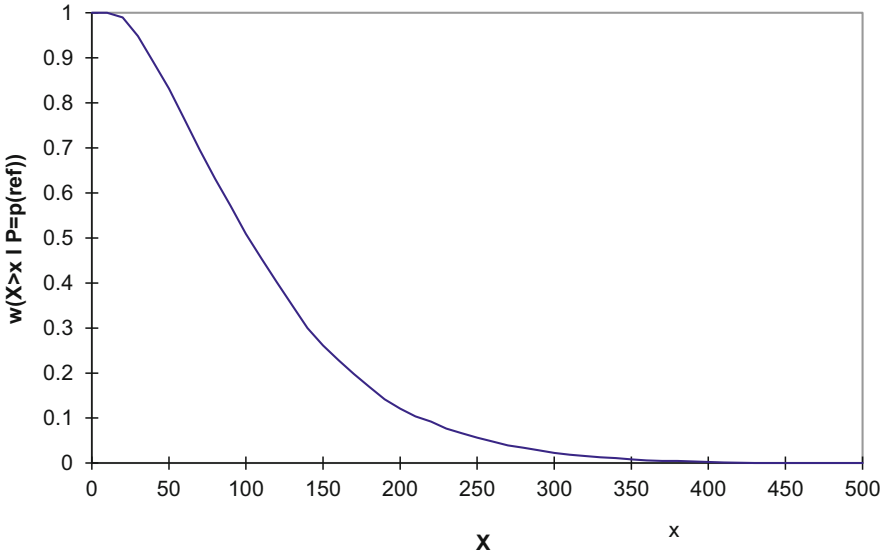


Fig. 9.5 Empirical complementary cumulative distribution function (ccdf) of X, obtained with the reference values of $\boldsymbol{P}$

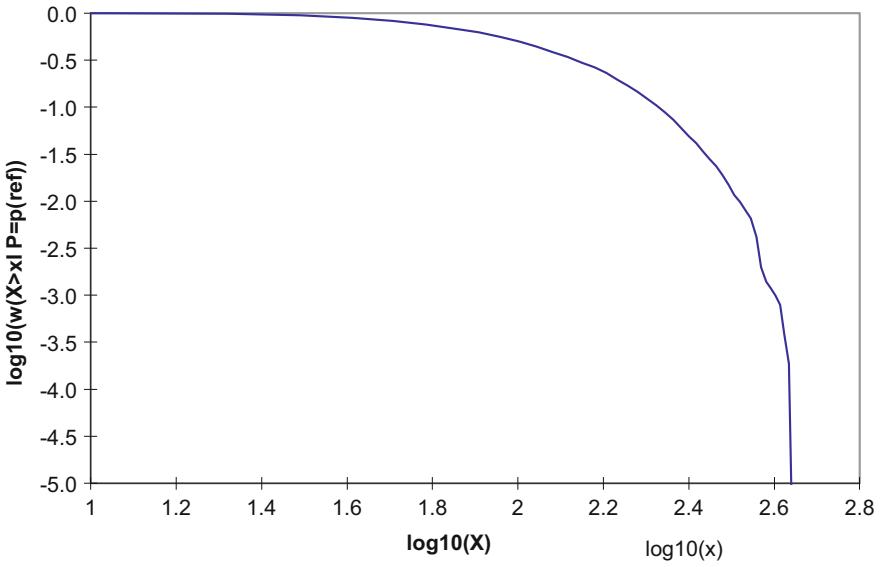


Fig. 9.6 Empirical complementary cumulative distribution function (ccdf) of X, obtained with the reference values of $\boldsymbol{P} = \boldsymbol{p}_{\boldsymbol{ref}}$ and using logarithmic scales

on the ordinate. This suggests the presentation of the ccdf using logarithmic scales as shown in Fig. 9.6.

The following example of an analysis of Type B uncertainty of the result from a computer model quantifying Type A uncertainty is subsequently used for illustration purposes. The model determines the spillage in kg per year of a toxic substance at an industrial site. The example is purely artificial. Since the annual spillage varies stochastically from year to year of operation and the reference unit of the assessment question is generic (. . ..per year of operation), there is a population of true answers summarized by a probability distribution. Due to the epistemic uncertainties involved in the computation of this distribution, there is a population of possibly true distributions. Figure 9.7 shows a sample from the population obtained by the analysis of Type B uncertainty using sample size N_B in the outer loop of the double randomization approach of Fig. 9.3 and sample size N_A in the inner Monte Carlo simulation loop. The corresponding sample of empirical ccdfs is shown in Fig. 9.8.

From the sample of empirical distributions computed in the outer loop of Fig. 9.3, one obtains estimates of the $q\%$ quantiles $w_{q\%}$ of the subjective probability distributions for $w(X \leq x)$ and for $w(X > x)$, i.e. $sw(w(X \leq x) \leq w_{q\%}) = q/100$ and $sw(w(X > x) \leq w_{q\%}) = q/100$ at every value x of X. These estimates of quantile values are continuously connected to arrive at the curve of $q\%$ quantile estimates as shown for $q = 5$, 10, 50, 90 and 95 in Fig. 9.9 for the cumulative distribution function and in Fig. 9.10 for the complementary cumulative distribution function. The estimates of the 50% quantiles (the medians) and of the mean values together with the sample minima and maxima of the empirical probabilities $w(X > x)$ are shown in Fig. 9.11

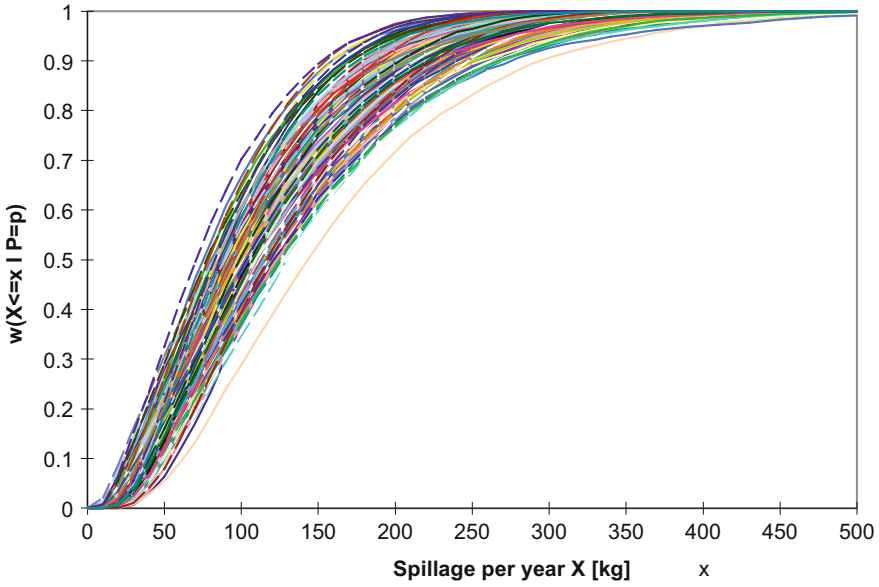


Fig. 9.7 Random sample of empirical cumulative distribution functions. Obtained with $N_B = 100$ sets of values of the epistemic uncertainties $\boldsymbol{P}$. The empirical probability (estimate obtained with sample size N_A) for not more than x kg of spillage per year is indicated at the ordinate

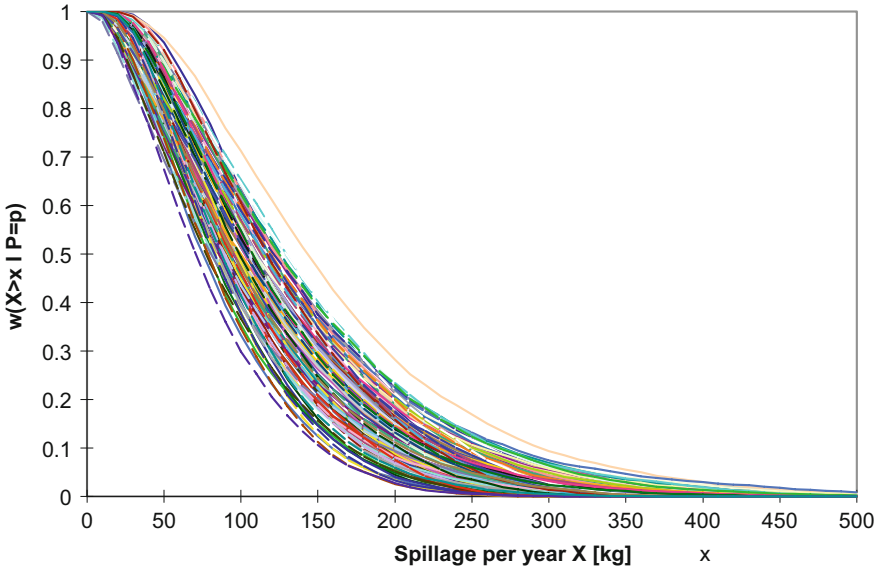


Fig. 9.8 Random sample of empirical complementary cumulative distribution functions obtained with $N_B = 100$ sets of values of the epistemic uncertainties $\boldsymbol{P}$. The empirical probability for more than x kg of spillage per year is indicated at the ordinate

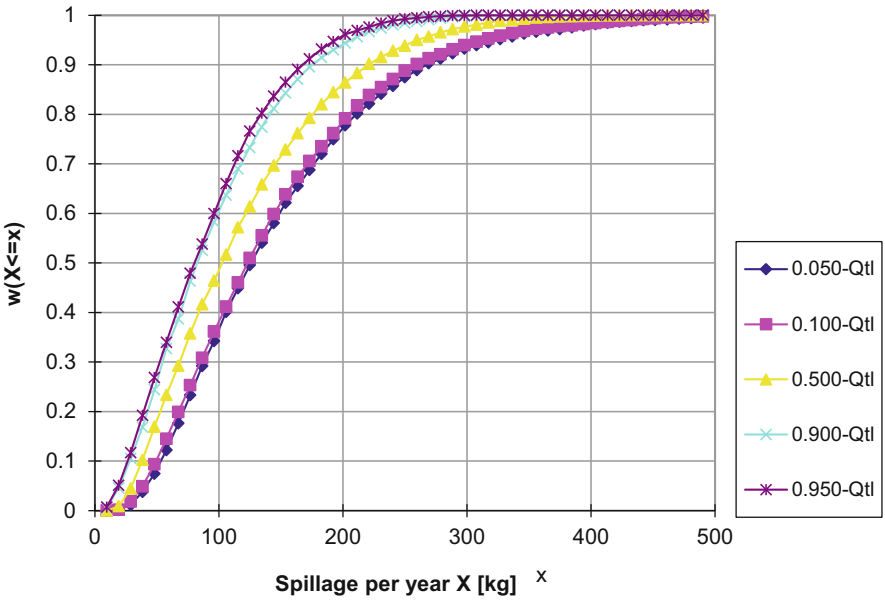


Fig. 9.9 The continuous connections of five quantiles of the empirical subjective probability distributions for the empirical cumulative probabilities at selected values of *X*. This Figure could also include the cdf obtained with "best estimate" or reference values for the epistemic uncertainties ***P***

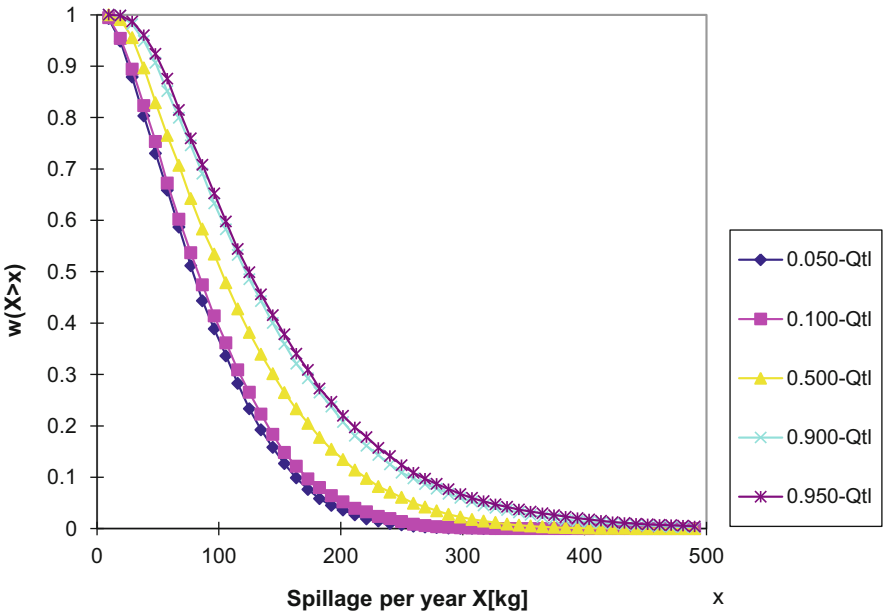


Fig. 9.10 The continuous connections of five quantiles of the empirical subjective probability distributions for the empirical complementary cumulative probabilities at selected values of *X*. This Figure could also include the ccdf obtained with "best estimate" or reference values for the epistemic uncertainties ***P***

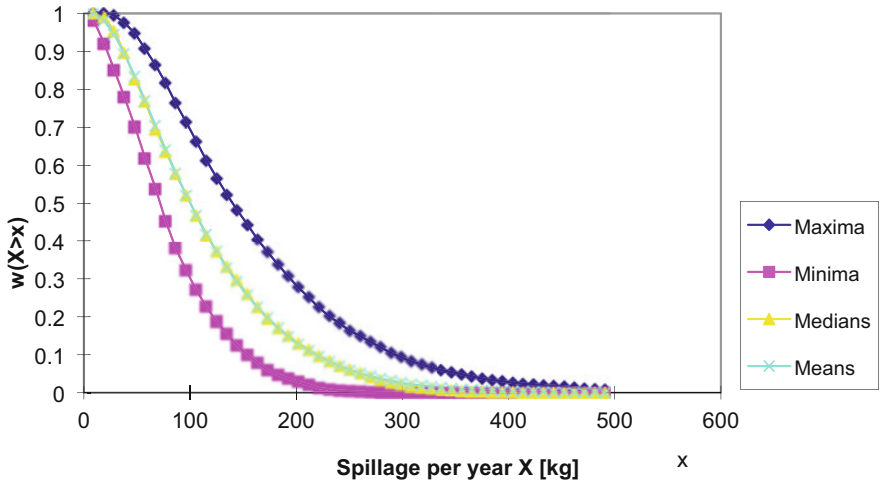


Fig. 9.11 The continuous connections of the mean and median values and smallest and largest values of the sample of N_B empirical complementary cumulative probabilities at selected values of X

for the complementary cumulative distribution function. The equivalent of Fig. 9.10 is shown in Fig. 9.12 for the empirical complementary cumulative distribution functions using the logarithmic scale.

Figure 9.9 provides, for instance, the following type of information: From the $N_B = 100$ model runs, it is estimated that 90% of the population of possibly true

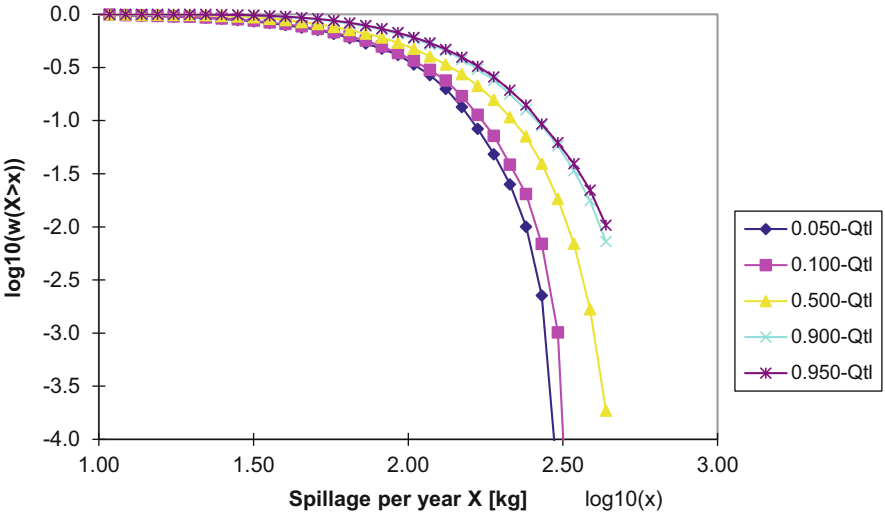


Fig. 9.12 The uncertainty information of Fig. 9.10 presented after $\log_{10}$ transformation in order to permit a more detailed illustration of the range of low probability and high model result values

values for the empirical probability that the spillage per year does not exceed 200 kg are between 0.78 and 0.96.

Figure 9.10 provides, for instance, the following type of information: From the $N_B = 100$ model runs, it is estimated that 90% of the population of possibly true values for the empirical probability that the spillage per year does exceed 200 kg are between 0.04 and 0.22.

Figure 9.13 provides, for instance, the following type of information: From the $N_B = 100$ model runs, it is estimated that at least 90% of the population of possibly true values for the empirical probability that the spillage per year does not exceed 200 kg are between 0.77 and 0.97 at a confidence level of at least 95%. These probability values are given by the intersections of the vertical line at 200 kg with the two limit lines. The values read from the intersections with the line segments, connecting the tolerance limit endpoints, will be subject to interpolation error.

Statistical tolerance limits can also be obtained for the spillage per year at selected values of the empirical cumulative probability (instead of for the empirical cumulative probability at selected values of the spillage per year). The corresponding Figure says that at least 90% of the population of possibly true values of the spillage per year that are not exceeded with empirical probability 0.5 lie between 77 kg and 135 kg at a confidence level of at least 95%. These values are given by the intersections of the horizontal line at cumulative probability 0.5 with the two limit lines.

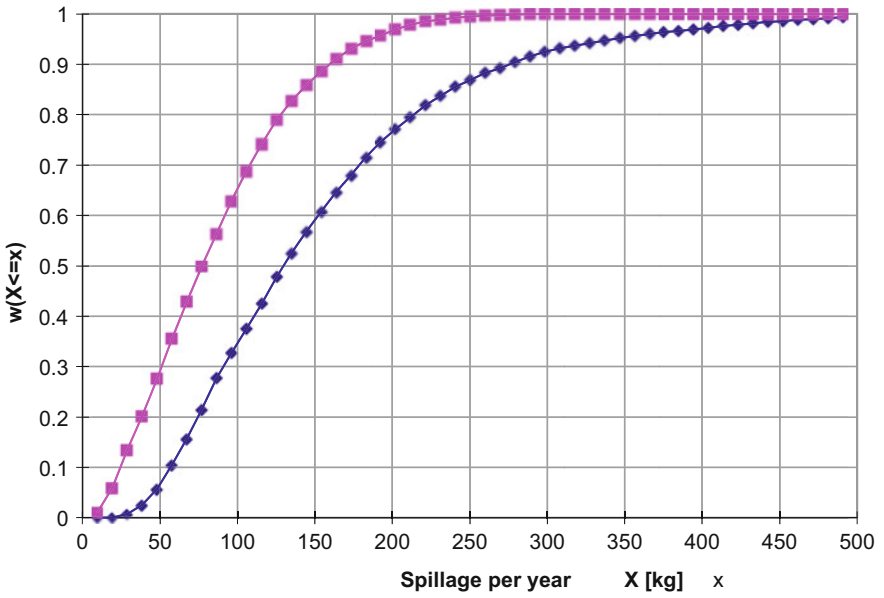


Fig. 9.13 The continuous connections of the endpoints of (90%, 95%) two-sided statistical tolerance limits for the empirical cumulative probability of the spillage X per year, shown at selected values of X

The confidence statement accounts for the possible influence of the estimation error that may be due to the fact that only a sample of $N_B = 100$ was drawn in the outer simulation loop. In addition, the inner loop only provides distribution estimates from a sample of size N_A in practice. The statistical tolerance limits also account for the uncertainty that is due to the variability of empirical probabilities obtained from samples of size N_A in the inner loop.

Analogous information as in Fig. 9.13 is available from Fig. 9.14 for the complementary cumulative probability.

Figure 9.14 provides, for instance, the following type of information: From the $N_B = 100$ model runs, it is estimated that at least 90% of the population of possibly true values for the empirical probability that the spillage per year exceeds 200 kg are between 0.03 and 0.23 at a confidence level of at least 95%.

Of particular interest will be the probability for X to exceed a given limit value x^*. This probability may be read directly from the probability distribution in its complementary form (ccdf). Since there is a population of possibly true ccdfs, due to the epistemic uncertainties involved in its computation, the state of knowledge of the probability for X to exceed x^* is quantified by a subjective probability distribution obtained from the analysis of Type B uncertainty. This distribution tells how well the empirical probability for X to exceed x* can only be known given the epistemic uncertainties and the variability of empirical probabilities obtained from samples of size N_A in the inner loop. Figure 9.15 shows the empirical subjective probability distribution of the empirical probability for the annual spillage to exceed the limit

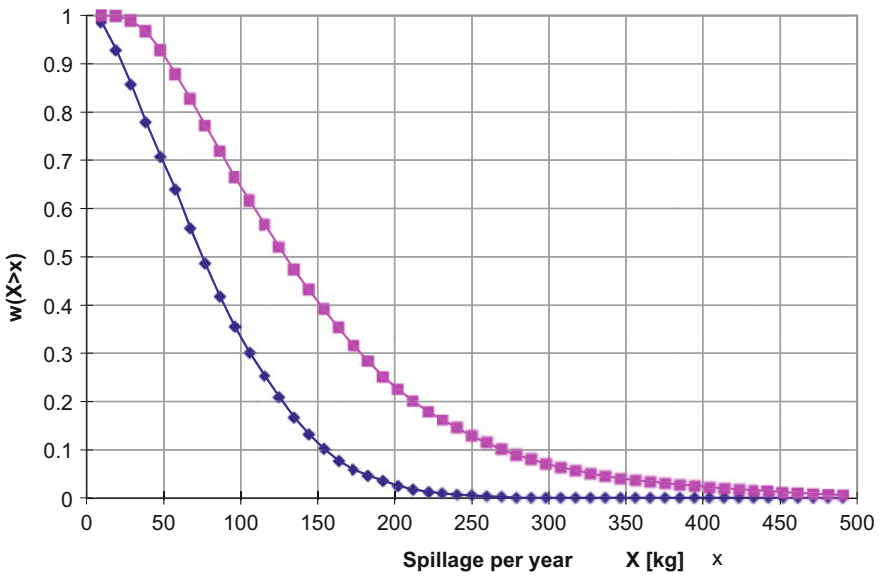


Fig. 9.14 The continuous connections of the endpoints of (90%, 95%) two-sided statistical tolerance limits for the empirical complementary cumulative probability of the spillage per year, shown at selected values of X

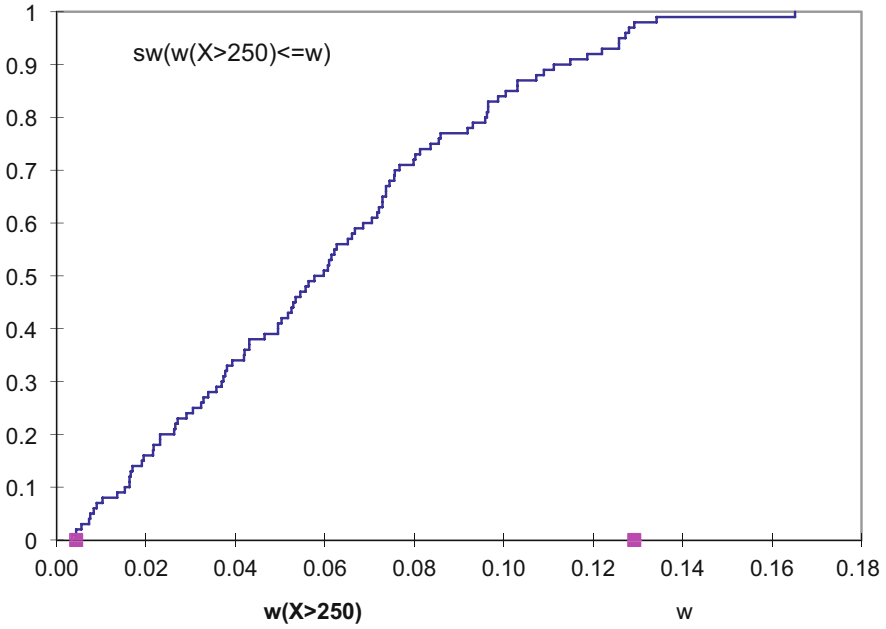


Fig. 9.15 The empirical subjective probability distribution for the empirical probability that the annual spillage X exceeds the limit value of 250 kg. A two-sided (90%, 95%) statistical tolerance limit is indicated on the abscissa. It says that, at a confidence level of at least 95%, at least 90% of the population of possibly true values of this empirical probability lie between 0.004 and 0.129

value of 250 kg. A two-sided (90%, 95%) statistical tolerance limit for this empirical probability is indicated on the abscissa.

The uncertainty information to all single value model results like the complementary cumulative probability $\mathrm{w}(X > x)$ or the cumulative probability $\mathrm{w}(X \leq x)$ or the quantile value $x_{\mathrm{q}\%}$ to a given cumulative (or complementary cumulative) probability $q/100$ or the mean or median value of X may be graphically presented as is shown in Sect. 5.3.

9.6 STEP 5: Rank Uncertainties

Uncertainty importance analysis has to answer different questions at each of the two sampling stages in Fig. 9.3.

After each completion of the inner simulation loop, the question is:

Which of the stochastic variables (aleatoric uncertainties) A_k, $k = 1, \ldots, K$, contribute most to the variance of the empirical distribution summarizing the sample of values of X_j, $j = 1, \ldots, J$, given $\boldsymbol{P} = \boldsymbol{p}$.

After the completion of the outer simulation loop, the question is:

Which of the epistemic uncertainties P_m, $m = 1, \ldots, M$ contribute most to the variance of:

- The sample of possibly true empirical distributions for X_j, $j = 1, \ldots, J$,
- The sample of possibly true values for Y_i, $i = 1, \ldots, I$ and
- The sample of estimates of the importance measures for X_j with respect to each of the A_k, $k = 1, \ldots, K$.

At both Monte Carlo simulation loops, the same types of uncertainty importance measures may be computed that were discussed in Chap. 6 for the epistemic uncertainties. However, their messages are quite different:

The measures derived after each completion of the inner simulation loop say where the assessment question would need to be changed in order to reduce the variances of the empirical marginal probability distributions of the X_j, $j = 1, \ldots, J$ most effectively or to shift the locations of these distributions to less critical ranges of values. This change of the assessment question may be achieved either through

- The use of a less generic reference unit, thereby restricting the variability of some of the A_k. For example: The reference unit might be changed from "... per repetition of experiment XYZ" to "... per repetition of experiment XYZ with molten mass between M_l and M_u" if the molten mass is one of the stochastic variables A_k and is free to vary between technically possible limits describing a larger range than (M_l, M_u).

or through

- a system modification, i.e. a technical change to the system under investigation. For example: If the time to failure of the heating system is one of the stochastic variables A_k that are identified as important contributors to the variance of the probability distribution of X_j, then the heating system of the experimental facility may be changed to one with improved reliability.

The uncertainty importance measures derived after completion of the outer simulation loop carry the same message as discussed in Chap. 6. They rank the epistemic uncertainties P_m, $m = 1, \ldots, M$ with respect to their contribution to the epistemic uncertainty of the empirical distributions of the X_j, $j = 1, \ldots, J$, the single output values for Y_i, $i = 1, \ldots, I$ and even the values of the uncertainty importance measures of the X_j with respect to each of the A_k, $k = 1, \ldots, K$. The importance measures from the outer simulation loop do not directly indicate changes to the reference unit nor to the system under investigation but they indicate where the state of knowledge of the uncertain data P_m, $m = 1, \ldots, M$, should be improved primarily in order to reduce the epistemic uncertainty of the results most effectively. Indirectly, this ranking may also lead to system changes. If, for instance, the epistemic uncertainty of a complementary cumulative probability is large due to lack of knowledge of the failure rate of a technical component, instead of improving its state of knowledge for the component presently in use, it may be decided to exchange the component for one with less uncertainty about its failure rate.

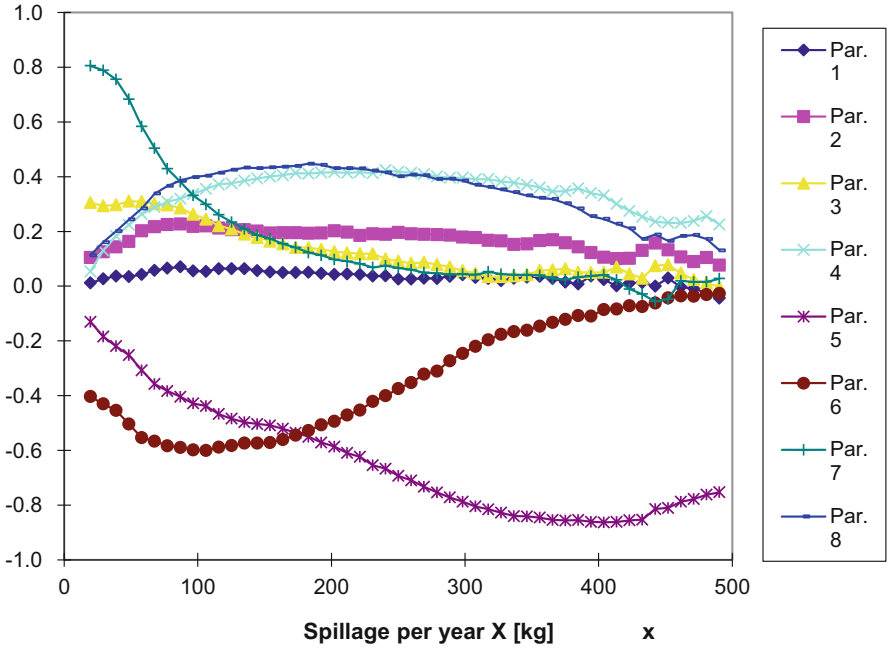


Fig. 9.16 The continuous connections of regression coefficients as uncertainty importance measures derived from the empirical complementary cumulative distribution functions in Fig. 9.8 after rank transformation of the sample values of the epistemic uncertainties $\boldsymbol{P}$ and of the sample ccdf values at the selected equidistant values of spillage. The epistemic uncertainties nos. 5 and 6 are identified as main contributors to the uncertainty of the empirical probability for more than x kg of spillage per year for most of the value range of X. Large values for parameter nos. 5 and 6 tend to lead to small values for the empirical probability of the spillage to exceed x

Improving the state of knowledge for any of the parameters of the random laws expressing the aleatoric uncertainty of A_k, $k = 1, \ldots, K$, does not change the true stochastic variability (or aleatoric uncertainty) of the model results X_j, $j = 1, \ldots, J$, but it changes the state of knowledge of the empirical probability distributions summarizing their stochastic variability.

Figure 9.16 shows the continuous connections of regression coefficients as uncertainty importance measures for the empirical complementary cumulative distribution function (ccdf) after rank transformation of the sample values of the epistemic uncertainties $\boldsymbol{P}$ and of the sample ccdf values at the selected equidistant values of spillage. The corresponding R^2 values are shown in Fig. 9.17. Figure 9.18 gives rank regression coefficients as uncertainty importance measures for the empirical probability that the spillage per year exceeds the limit value of 250 kg. The corresponding R^2 value is given in the Figure caption.

Figure 9.18 shows that most of the uncertainty of the empirical probability for the spillage to exceed the limit value of 250 kg per year is due to the uncertainty of

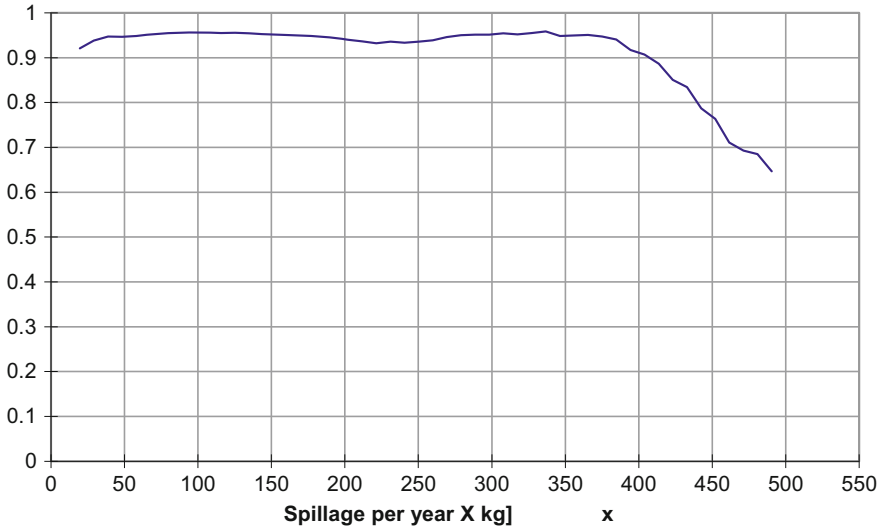


Fig. 9.17 The R^2 values for the set of rank regression coefficients shown in Fig. 9.16. The uncertainty importance measures in Fig. 9.16 explain more than 90% of the uncertainty of the empirical probability for more than x kg of spillage per year for x up to 400

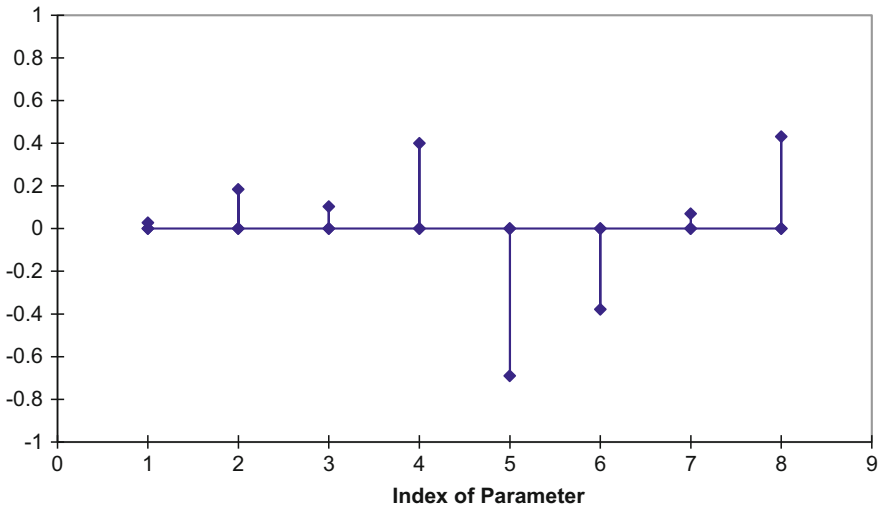


Fig. 9.18 Rank regression coefficients as uncertainty importance measures of the empirical probability that the spillage per year exceeds the limit value of $x = 250$ kg. The corresponding R^2 value is just below 0.95 indicating that the measures explain almost 95% of the uncertainty of this probability quantified by the empirical subjective probability distribution shown in Fig. 9.15

parameter no. 5 with large parameter values tending to lead to low probability values.

The inner simulation loop may be replaced by an importance sampling or subset sampling procedure (see Sects. 4.4.3 and 4.4.4) if the low probability of a subset of extreme values is to be estimated. The Type B uncertainty importance analysis is performed in the outer simulation loop as usual.

9.7 STEP 6: Present the Analysis and Interpret Its Results

The presentation of the analysis of Type B uncertainty of results from an analysis of Type A uncertainty has the same structure as described in Chap. 7. The only exceptions are:

– The assessment question

  It has a generic reference unit and therefore a population of true answers.

– The assessment model

It is a computer model that includes the analysis of Type A uncertainty in order to arrive at a summary of the population of answers to the assessment question in the form of a probability distribution. Particularly, it includes the identification of the aleatoric uncertainties, leading to the population of answers, and the modelling of their stochastic variability by probability distributions. It also includes the inner Monte Carlo loop of Fig. 9.3 or any other approach (i.e. scenario analysis, etc.) suitable to arrive at an estimate of or approximation to the probability distribution summarizing the population of answers to the assessment question given the values of the epistemic uncertainties sampled in the outer simulation loop. Finally, the presentation of the distribution estimate as well as of its characteristics of interest, like mean value, standard deviation, cumulative and complementary cumulative probabilities at limit values, etc., and their interpretation are all part of the assessment model. The results of the analysis of Type B uncertainty of the output from this assessment model are then determined, presented and interpreted as for any other computer model (see Chap. 7 and the specifics discussed in Sects. 9.5 and 9.6).

References

Apostolakis, G., & Kaplan, S. (1981). Pitfalls in risk calculations. *Reliability Engineering, 2*, 135–145.

Bayes, T. (1958). Thomas Bayes‘essay towards solving a problem in the doctrine of chances. *Studies in the History of Probability and Statistics, Biometrika, 45*(3/4), 296–315.

Box, G. E. P., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Reading, MA: Addison-Wesley.

Buckley, J. J. (1985). Entropy principles in decision making under risk. *Risk Analysis, 5*(4), 303–313.

Chang, Y.-H., & Mosleh, A. (1998). *Dynamic PRA using ADS with RELAP5 Code as its thermal hydraulic module*. In A. Mosleh & R. A. Bari (Eds.), *Proceedings of PSAM-4*. Berlin: Springer.

Cojazzi, G. (1996). The DYLAM approach for the dynamic reliability analysis of systems. *Reliability Engineering and System Safety, 52*, 279–296.

Devooght, J., & Smidts, C. (1996). Probabilistic dynamics as a tool for dynamic PSA. *Reliability Engineering and System Safety, 52*, 185–196.

Hofer, E., Kloos, M., Krzykacz-Hausmann, B., Peschke, J., & Sonnenkalb, M. (2002). *Dynamic event trees for probabilistic safety analysis*. Eurosafe Conference, Berlin.

Hsueh, K. S., & Mosleh, A. (1996). The development and application of the accident dynamic simulator (ADS) for dynamic probabilistic risk assessment of nuclear power plants. *Reliability Engineering and System Safety, 52*, 297–314.

Kloos, M., & Peschke, J. (2006). MCDET: A probabilistic dynamics method combining Monte Carlo simulation with the discrete dynamic event tree approach. *Nuclear Science and Engineering, 153*, 137–156.

Labeau, P. E., Smidts, C., & Swaminathan, S. (2000). Dynamic reliability: Towards an integrated platform for probabilistic risk assessment. *Reliability Engineering and System Safety, 68*, 219–254.

Siu, N. (1994). Risk assessment for dynamic systems: An overview. *Reliability Engineering and System Safety, 43*, 43–73.

Winkler, R. L., & Hays, W. L. (1975). *Statistics: Probability, inference, and decision* (2nd ed.). New York: Holt, Rinehart and Winston.

Chapter 10
Practical Examples

10.1 Introduction

The model applied in Sect. 10.2 deals with the population dynamics of the Peruvian anchovy. It was developed in an undergraduate modelling contest (Caulkins et al. 1985) and was modified to some extent in order to serve the intended illustration purpose of this chapter. State-of-the-art population dynamics models for animal populations, in particular for various species of fish, are very sophisticated. Documentation of the uncertainty analysis of their application would fill a report and would therefore not be suitable to serve here as a practical example.

The uncertainty of results from the application of a dose reconstruction model, again developed for illustration purposes only, is analysed in Sect. 10.3. The assessment question asks for the dose value of each of N specific individuals. No Type A uncertainty is therefore involved. The model was specifically designed to illustrate the difference in the treatment of the two kinds of measurement error of Sect. 3.3.1 as well as the effect of uncertain data shared by the reconstructed dose values of several or all individuals.

The results from the population dynamics model are functions of the independent variable "time". A specific sequence of 20 years of fishing activity is modelled. Therefore, there is no Type A uncertainty involved. During this time, an event called "El Nino" may occur. It is unknown how often this will be the case, when the event will occur and if it occurs how severe it will be. The El Nino uncertainty is epistemic since the model application deals with a specific sequence of 20 years. Together with various uncertain data, the El Nino uncertainty is the subject of the Type B uncertainty analysis.

The results from the dose reconstruction model, i.e. the dose value of each of the N specific individuals, can be arranged in a probability distribution summarizing the variability of dose over this population. Due to the uncertain data of the dose reconstruction model, there are infinitely many possibly true distributions. The uncertainty analysis deals with the percentage of individuals (in the given

E. Hofer, *The Uncertainty Analysis of Model Results*,
https://doi.org/10.1007/978-3-319-76297-5_10

population) with dose value above a given limit as well as with the dose received by selected individuals. It uses the two-loop Monte Carlo simulation approach discussed in Chap. 9 in order to quantify the epistemic (Type B) uncertainty of a population value, i.e. the percentage of individuals in the population with dose above the given limit value, and to quantify the epistemic uncertainty of individual dose values while properly accounting for their state of knowledge dependences that are due to shared data uncertainties.

10.2 Uncertainty Analysis of Results from the Application of a Population Dynamics Model for the Peruvian Anchovy

The following short summary of the background information to the assessment questions is taken from Caulkins et al. (1985):

"The anchovy is important for Peru's large fishing industry. The Peruvian anchovy inhabits a narrow strip of water that runs 1500 miles along the coast of Peru. Strong prevailing winds cause a constant northern flow of the surface water. The displaced water is replaced by nutrient-enriched upwelling that supports heavy concentrations of plankton. This plankton is essential for the survival of the anchovy larvae. An unpredictable ocean counter-current, known as El Nino, stops the upwelling and thus has devastating effects on the anchovy larvae. El Nino together with heavy fishing has in the past severely depleted anchovy stocks. The ultimate goal is to find an optimal fishing management policy.

The anchovies have three enemies: The guano-birds, other fish, and man. Since 80% to 95% of the guano-bird's diet consists of adult anchovies, the bird population is highly sensitive to fluctuations in the anchovy population."

10.2.1 The Assessment Questions

The model is applied to answer the following questions for the specified time period of 20 years:

- How large is the harvestable anchovy biomass at discrete times?
- How large is the population of guano-birds at discrete times?

  If an annual fishing quota of 5 Mt is adhered to?

10.2.2 The Model

The anchovy population dynamics model from the literature (Caulkins et al. 1985) was modified and implemented as a computer model. The model incorporates the peculiar characteristics of the anchovy ecosystem. In particular, it

- Divides the population into 6 age cohorts:

  Cohort 1: up to 4 months
  Cohort 2: over 4 and up to 8 months
  Cohort 3: over 8 and up to 12 months
  Cohort 4: over 12 and up to 16 months
  Cohort 5: over 16 and up to 20 months
  Cohort 6: over 20 months (adult anchovies)

- Proceeds in time periods of 4 months:

  Period 1: January to April
  Period 2: May to August
  Period 3: September to December

- Considers the spawning behaviour of the anchovy:

  75% of the total spawning in September
  25% in January

- Relates the number of eggs to body length (age)
- Accounts for the effect of overpopulation on spawning
- Considers infant mortality and predation by guano-birds and by non-bird enemies
- Stops harvest if population is driven below minimum
- Includes the effects of El Nino on:

  Spawning behaviour
  Fecundity
  Infant mortality
  Predation by non-bird enemies
  The breaking up of anchovy schools (scatter)
  Minimum population

- Contains the dynamics of the guano-bird population:
  - Growth rate
  - Predation on population of anchovies
  - Survival if there is a shortage of anchovies

- Accounts for El Nino over the time period considered with respect to

  Years of occurrence
  Severity

10.2.3 The Analysis Tool

The uncertainty and sensitivity analysis was performed with version 3.6 of the analysis software SUSA (Kloos 2015).

SUSA supports the probabilistic modelling of the state of knowledge of uncertain data. It offers a large selection of distribution types, takes distribution truncations into account, derives distribution parameters from quantile values and plots the selected distributions. For the quantification of state of knowledge dependences, it accepts population measures of association, sample measures, conditional distributions, restrictions, complete dependence and functional relationships between uncertain data and provides scatter plots to illustrate the selected dependence quantification.

SUSA offers simple random sampling and Latin Hypercube sampling, derives distribution-free quantile estimates as well as statistical tolerance limits and performs statistical tests of distribution hypotheses and fitting of distributions to data. A choice of uncertainty importance measures is available for model output in the form of single values as well as of discretized functions of an independent variable like time.

10.2.4 The Elicitation Process

There was no elicitation process for this demonstration of an uncertainty analysis. Provisional quantitative expressions of the state of knowledge are provided by the analyst and are used for illustration purposes only.

10.2.5 The Potentially Important Uncertainties

A total of 37 uncertain data were identified in addition to the uncertainty of when El Nino will occur.

10.2.6 Provisional State of Knowledge Quantifications

The state of knowledge to all uncertain data is expressed by subjective probability distributions. Provisional subjective probability distributions are used in this analysis for demonstration purposes only. They are presented in Fig. 10.1. The information in this Figure is part of the automatic documentation support offered by SUSA.

Additionally, a subjective probability of 0.14 was specified for the occurrence of an El Nino for every year; while the subjective probability for the El Nino to be a severe one was set at 0.4.

Figure 10.1 (a–e) presents the SUSA output of the state of knowledge quantifications.

Par. No.	Short Par. Name	Full Parameter Name	Symbol	Unit	Best Estimate	Notes	Distribution Type	Distrib. Par. p1	Distrib. Par. p2	Min.	Max.	Proportion Class
1	guano gro	growth rate of guano birds per period	guanogro	per 4 months	0.05		Uniform Distr.	0.03	0.08	0.03	0.08	
2	catch6 pr	probability that guano bird can catch fish from cohort 6	catch6pr	-	0.9		Uniform Distr.	0.8	0.9	0.8	0.9	
3	catch5 pr	probability that guano bird can catch fish from cohort 5	catch5pr	-	0.5		Uniform Distr.	0.4	0.6	0.4	0.6	
4	loss1	loss from cohort 1 due to predation by non-bird enemies (proportion of total loss)	loss1	-	0.8		Uniform Distr.	0.7	0.9	0.7	0.9	1
5	loss2	loss from cohort 2 due to predation by non-bird enemies (proportion of total loss)	loss2	-	0.125	Distribution for conditional proportion	Uniform Distr.	0.6	0.8	0.6	0.8	1
6	loss3	loss from cohort 3 due to predation by non-bird enemies (proportion of total loss)	loss3	-	0.05	Distribution for conditional proportion	Uniform Distr.	0.5	0.7	0.5	0.7	1

Par. No.	Short Par. Name	Full Parameter Name	Symbol	Unit	Best Estimate	Notes	Distribution Type	Distrib. Par. p1	Distrib. Par. p2	Min.	Max.	Proportion Class
7	loss4	loss from cohort 4 due to predation by non-bird enemies (proportion of total loss)	loss4	-	0.0125	Distribution for conditional proportion	Uniform Distr.	0.5	0.7	0.5	0.7	1
8	loss5	loss from cohort 5 due to predation by non-bird enemies (proportion of total loss)	loss5	-	7.50E-03	Distribution for conditional proportion	Uniform Distr.	0.5	0.7	0.5	0.7	1
9	loss6	loss from cohort 6 due to predation by non-bird enemies (proportion of total loss)	loss6	-	5.00E-03		Uniform Distr.	0.0	1.0	0.0	1.0	
10	loss pred	proportion lost due to non-bird predation (all cohorts)	losspred	-	0.3		Uniform Distr.	0.2	0.4	0.2	0.4	
11	alive 50	percentage of birds kept alive if 50% of total desired food are available	alive50	%	75		Uniform Distr.	40	80	40	80	
12	max bioma	anchovy biomass at which fecundity starts to decline	maxbioma	Mt	20		Uniform Distr.	10	30	10	30	
13	mort nino	factor to mortality rates to account for El Nino	mortnino	-	10		Uniform Distr.	5	15	5	15	
14	loss nino	percentage lower survival rate due to non-bird predators from El Nino	lossnino	%	3		Uniform Distr.	1	5	1	5	

Par. No.	Short Par. Name	Full Parameter Name	Symbol	Unit	Best Estimate	Notes	Distribution Type	Distrib. Par. p1	Distrib. Par. p2	Min.	Max.	Proportion Class
15	min pop	percentage of maxbioma (Par. 12) as minimum population value in cohorts 4,5,6	minpop	%	10	Min Pop. equals maxbioma* prop(cohort) *minpop%/ weight (cohort) with prop(cohort) in P28, P29, P30	Uniform Distr.	5	20	5	20	
16	inpo nino	factor to increase minimum population values (cohorts 3,4,5,6) due to scatter in El Nino	inponino	-	2		Uniform Distr.	2	5	2	5	
17	techno	factor to minimum population value (cohorts 4,5,6) for better fishing technology	techno	-	0.5		Uniform Distr.	0.3	0.5	0.3	0.5	
18	weight1	average weight of anchovy in cohort 1	weight1	grs	5		Uniform Distr.	3	7	3	7	
19	weight2	average weight of anchovy in cohort 2	weight2	grs	10		Uniform Distr.	8	15	8	15	
20	weight3	average weight of anchovy in cohort 3	weight3	grs	20		Uniform Distr.	18	25	18	25	
21	weight4	average weight of anchovy in cohort 4	weight4	grs	30		Uniform Distr.	28	35	28	35	

Fig. 10.1 (**a–e**) SUSA output: state of knowledge quantifications

Par. No.	Short Par. Name	Full Parameter Name	Symbol	Unit	Best Estimate	Notes	Distribution Type	Distrib. Par. p1	Distrib. Par. p2	Min.	Max.	Proportion Class
22	weight5	average weight of anchovy in cohort 5	weight5	grs	40		Uniform Distr.	38	45	38	45	
23	weight6	average weight of anchovy in cohort 6	weight6	grs	50		Uniform Distr.	48	55	48	55	
24	start bio	start value for biomass as fraction of maxbioma (Par.12)	startbio	-	1.25		Uniform Distr.	0.75	1.75	0.75	1.75	
25	star pro1	start value of proportion of cohort 1 in total biomass	starpro1	-	0.11		Uniform Distr.	0.05	0.15	0.05	0.15	2
26	star pro2	start value of proportion of total biomass for cohort 2	starpro2	-	0	Distribution for conditional proportion	Uniform Distr.	0	1E-10	0	1E-10	2
27	star pro3	start value of proportion of total biomass for cohort 3	starpro3	-	0.29	Distribution for conditional proportion	Uniform Distr.	0.25	0.4	0.25	0.4	2
28	star pro4	start value of proportion of total biomass for cohort 4	starpro4	-	0.3	Distribution for conditional proportion	Uniform Distr.	0.4	0.6	0.4	0.6	2
29	star pro5	start value of proportion of total biomass for cohort 5	starpro5	-	0	Distribution for conditional proportion	Uniform Distr.	0	1E-10	0	1E-10	2
30	star pro6	start value of proportion of total biomass for cohort 6	starpro6	-	0.3		Uniform Distr.	0.0	1.0	0.0	1.0	
31	star bird	start value of guano-bird population	starbird	bird	2.50E+07		Uniform Distr.	15000000	40000000	15000000	40000000	

Par. No.	Short Par. Name	Full Parameter Name	Symbol	Unit	Best Estimate	Notes	Distribution Type	Distrib. Par. p1	Distrib. Par. p2	Min.	Max.	Proportion Class
32	anch bird	average number of anchovies per bird and period	anchbird	anchovy	1.20E+04		Uniform Distr.	11000	13000	11000	13000	
33	rinf mort	infant mortality rate (anchovy larvae)	rinfmort	-	0.998		Uniform Distr.	0.9955	0.9985	0.9955	0.9985	
34	eggs3	average number of eggs per female in cohort 3	eggs3	egg	9000		Uniform Distr.	8000	10000	8000	10000	
35	eggs4	average number of eggs per female in cohort 4	eggs4	egg	11500		Uniform Distr.	10000	13000	10000	13000	
36	eggs5	average number of eggs per female in cohort 5	eggs5	egg	15000		Uniform Distr.	13000	17000	13000	17000	
37	eggs6	average number of eggs per female in cohort 6	eggs6	egg	19500		Uniform Distr.	17000	22000	17000	22000	

Fig. 10.1 (continued)

10.2.7 State of Knowledge Dependences

Figure 10.2 lists the identified state of knowledge dependences. The information in this Figure is also part of the automatic documentation support offered by SUSA. Two sets of dependences, each due to proportionality relationships (see Sect. 3.5.2.3 and 4.4.1.8), are included in the Figure. This requires the Monte Carlo simulation to generate the random sample of possibly true values in two steps. SUSA

1st Par.No	Short Par.Name	2nd Par.No	Short Par.Name	Dependence Model	Characteristics
30	starpro6			Design Extension	#1.0-par(25)-par(26)-par(27)-par(28)-par(29)
9	loss6			Design Extension	#1.0-par(4)-par(5)-par(6)-par(7)-par(8)
34	eggs3	35	eggs4	Spearman's rank corr.	0.7
34	eggs3	36	eggs5	Spearman's rank corr.	0.7
34	eggs3	37	eggs6	Spearman's rank corr.	0.7
35	eggs4	36	eggs5	Spearman's rank corr.	0.7
35	eggs4	37	eggs6	Spearman's rank corr.	0.7
36	eggs5	37	eggs6	Spearman's rank corr.	0.7

Fig. 10.2 SUSA output: state of knowledge dependence quantifications

automatically performs the necessary steps as design extensions. It first samples for the involved uncertain data, according to their conditional distributions given in Fig. 10.1 and in the second step post-processes the sampled values according to the relationships in Sect. 4.4.1.8 to satisfy the proportions condition.

10.2.8 Model Results Obtained with Best Estimate Data Values and El Nino

Figures 10.3, 10.4, 10.5, and 10.6 show the

- Population of anchovies in cohort 6 (adult anchovies)
- Occurrence periods of the "best estimate" El Nino sequence
- Population of guano-birds
- Harvestable anchovy biomass (cohorts 3–6)

over the specified sequence of 60 consecutive periods of 4 months each (20 years).

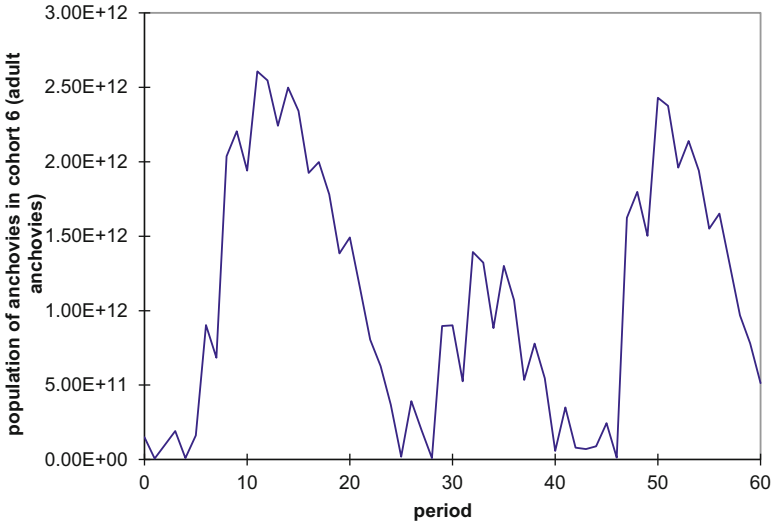


Fig. 10.3 The population of anchovies in age cohort 6 (adult anchovies) computed at intervals of 4 months and over the given time period of 20 years, using best estimate values of the uncertain data, best estimate El Nino sequence (it is the reason for the marked declines in population numbers) and an annual fishing quota of 5 Mt

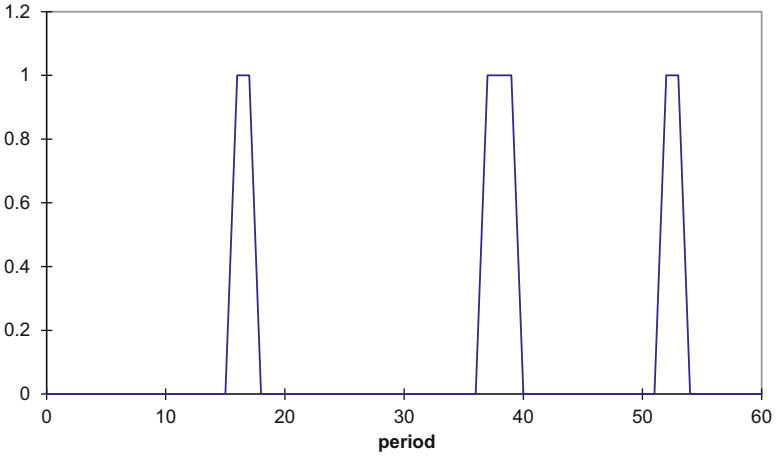


Fig. 10.4 Occurrence periods of the "best estimate" El Nino sequence. The width of the spike indicates the severity of the El Nino

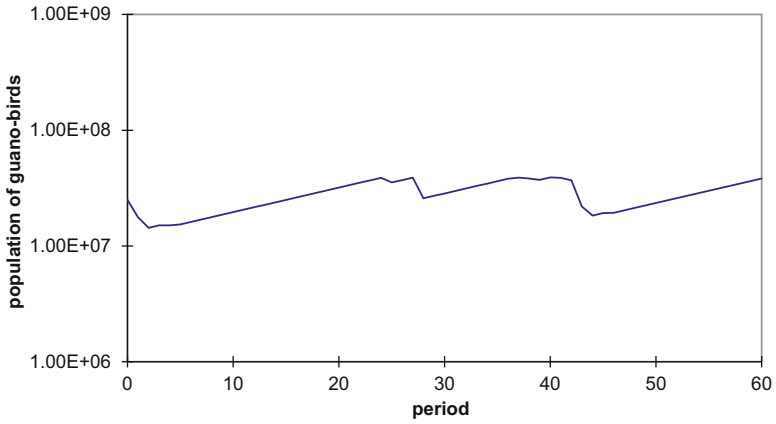


Fig. 10.5 The population of guano-birds computed with best estimate values of the uncertain data, best estimate El Nino sequence and an annual fishing quota of 5 Mt

10.2.9 Propagation of the Provisional State of Knowledge Quantifications Through the Model

One thousand sets of 37 data values each were selected at random (simple random sampling) according to a joint subjective probability distribution satisfying the provisional state of knowledge quantifications in Figs. 10.1 and 10.2. Each set was combined with a random El Nino event sequence defined over the specified 20 years. The model was run for each of these combinations. The 1000 sets of model results are a random sample from their unknown joint subjective probability distribution

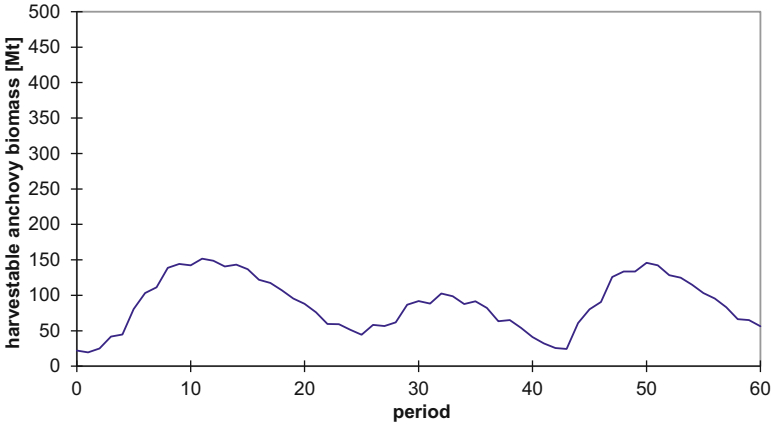


Fig. 10.6 The harvestable anchovy biomass computed with best estimate values of the uncertain data, best estimate El Nino sequence and an annual fishing quota of 5 Mt

that is the logical consequence of the propagation of the provisional state of knowledge quantifications through the computer model. Quantitative uncertainty statements are derived from this sample.

10.2.10 Uncertainty Statements for Selected Model Results

To give a first impression of the uncertainty of the model result "harvestable anchovy biomass", the first 10 time histories out of the simulated 1000 are shown in Fig. 10.7. They are followed by the 5%, 10%, 50%, 90% and 95% subjective probability lines in Figs. 10.8 and 10.9 as well as by the continuous connections of the local upper and lower endpoints of the (90%, 95%) two-sided statistical tolerance limits for the harvestable anchovy biomass (Fig. 10.10) and for the population of guano-birds (Fig. 10.11).

Comparing Figs. 10.6 and 10.10 shows that the harvestable anchovy biomass computed with best estimates of the uncertain data assumes values mainly from the lower end of the uncertainty range delimited by the endpoints of the two-sided statistical tolerance limits. The uncertainty range extends over about a factor of 8 between lower and upper endpoint of the two-sided (90%, 95%) limit, while the solution obtained with best estimates never lies more than a factor of 3 above the lower limit.

The comparison of Figs. 10.5 and 10.11 shows that the model result for the population of guano-birds, obtained with best estimates of the uncertain data, lies close to the upper end of the two-sided statistical tolerance limit. Contrary to the situation for the harvestable anchovy biomass, the uncertainty range extends here more to the lower values. The uncertainty importance measures will help explain this difference.

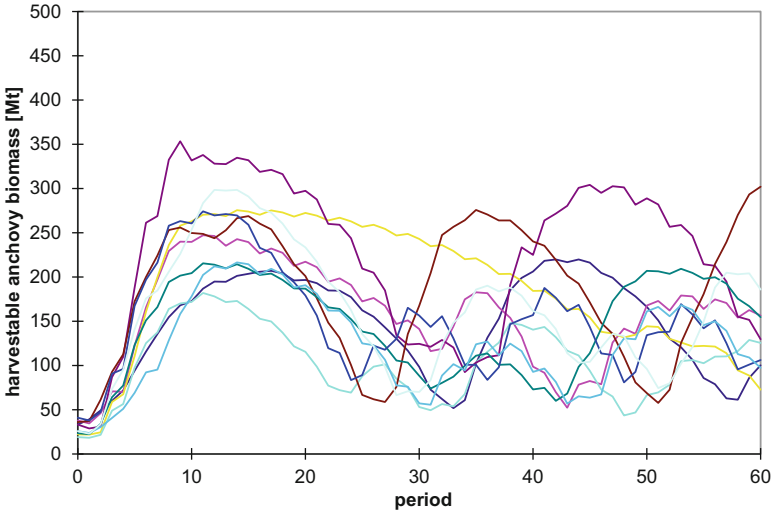


Fig. 10.7 Ten time histories of the harvestable anchovy biomass computed at intervals of 4 months over the time period of 20 years; they are the first ten time histories out of a random sample of size $N = 1000$ drawn according to the joint subjective probability distribution for all uncertain data including the uncertain El Nino event sequence

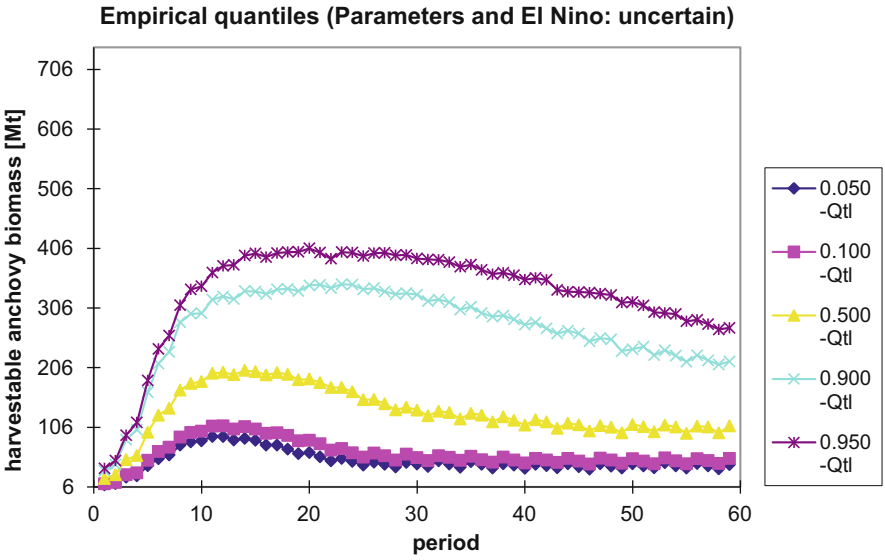


Fig. 10.8 The continuous connections of selected quantile values of the local (at intervals of 4 months) subjective probability distributions for the computed harvestable anchovy biomass in Mt; they are estimated using a random sample of size $N = 1000$ drawn according to the joint subjective probability distribution for all uncertain data including the uncertain El Nino event sequence

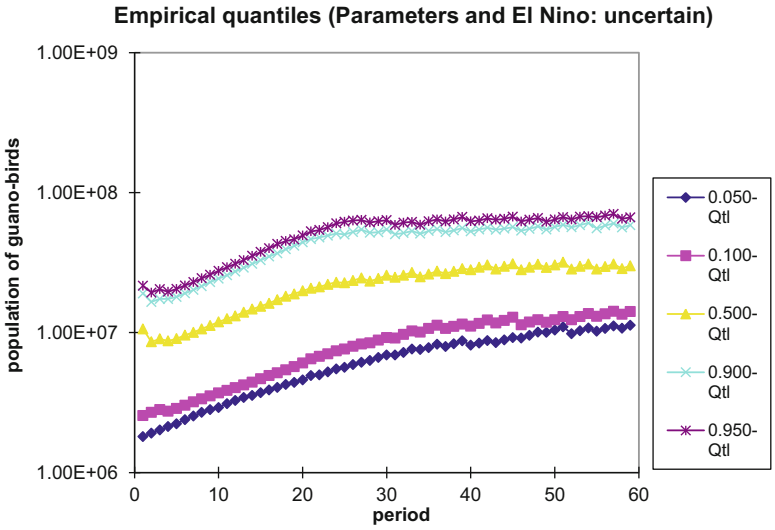


Fig. 10.9 The continuous connections of selected quantile values of the local (at intervals of 4 months) subjective probability distributions for the computed population of guano-birds

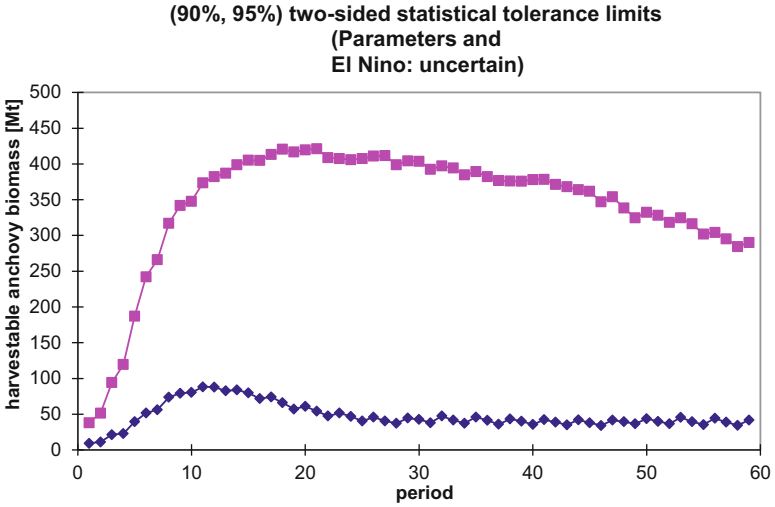


Fig. 10.10 The continuous connections of the local (at intervals of 4 months) (90%, 95%) two-sided statistical tolerance limits of the computed harvestable anchovy biomass in Mt; the limits contain at least 90% subjective probability at a confidence level of at least 95%

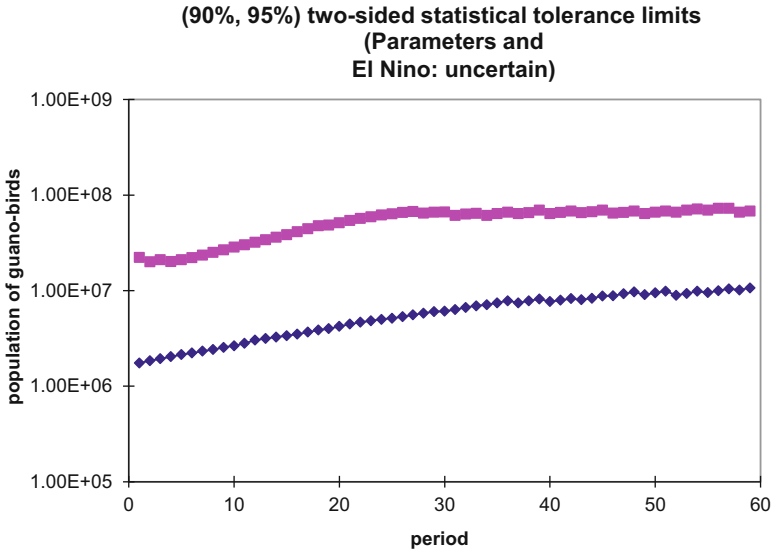


Fig. 10.11 The continuous connections of the local (at intervals of 4 months) (90%, 95%) two-sided statistical tolerance limits of the computed population of guano-birds. The limits contain at least 90% subjective probability at a confidence level of at least 95%

10.2.11 Uncertainty Importance Statements for Selected Model Results

Two groups of uncertain data in Fig. 10.1 are members of a proportion relationship (see Sect. 3.5.2.3). Some of these data exhibit strong state of knowledge dependence (in particular: The correlation between data nos. 4 and 5 is about −0.95) that leads to very high variance inflation factors for the standardized regression coefficients [see Eq. (6.45)]. SUSA issues a warning and indeed the regression coefficients obtained from raw as well as from rank transformed data (nos. 9 and 30 excluded from the regression model—see Fig. 10.2) turn out to make little sense. Consequently, the uncertainty importance analysis initially resorts to correlation coefficients and to the correlation ratios for ranking of the uncertain data.

The following figures show graphs of correlation coefficients and approximate correlation ratios for the model results

- Population of guano-birds
- Harvestable anchovy biomass

with respect to the uncertain data of Fig. 10.1. They are estimated using the 1000 values of the model result and the 1000 sets of values for the 37 uncertain data in combination with the 1000 El Nino event sequences.

Figure 10.12(a–d) shows the correlation coefficients for the model result "population of guano-birds" with respect to the uncertain data of Fig. 10.1.

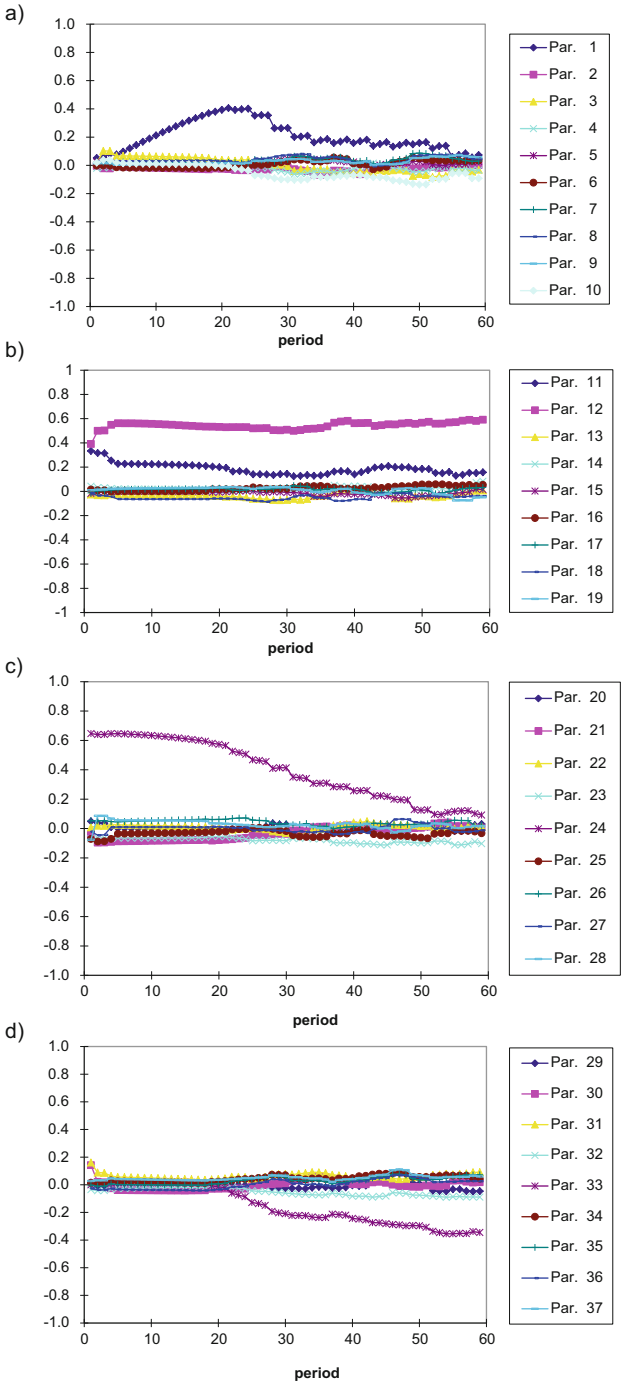


Fig. 10.12 (**a**–**d**) Correlation coefficients for the model result "population of guano-birds" with respect to the uncertain data of Fig. 10.1; the coefficients are estimated at each of the 60 points of time. Data nos. 1, 11, 12, 24 and 33 are identified as the main contributors to the uncertainty of the

Figure 10.13(a–d) presents the correlation coefficients obtained from rank transformed data. The latter emphasize associations between rank orders of data, while the correlation coefficients obtained from raw data measure associations between the actual values and are therefore more sensitive to outliers. However, the top ranked uncertain data and the directions of their influence on the computed population of guano-birds are the same for both uncertainty importance measures. The largest contributions to the uncertainty of the population of guano-birds come from uncertain data nos. 1, 11, 12, 24 and 33 over most of the simulated time span. Large values of nos. 1, 11, 12 and 24 tend to lead to large values of the population, while large values of no. 33 tend to lead to small values of the population of guano-birds. Unfortunately, there is no information available about how much of the uncertainty is explained by the top ranked uncertain data.

Figure 10.14(a–d) show the approximate correlation ratio. While this measure cannot provide directional information, it is capable of accounting for non-linear and even non-monotone relationships between the model result and the uncertain data.

The correlation coefficients and the approximate correlation ratios identify the same set of five main contributors to the uncertainty of the computed temporal evolution of the guano-bird population.

Figures 10.15, 10.16, and 10.17 present the same uncertainty importance measures for the model result "harvestable anchovy biomass".

Figure 10.15(a–d) shows the correlation coefficients for the model result "harvestable anchovy biomass" with respect to the uncertain data of Fig. 10.1, while the correlation coefficients obtained from rank transformed data are presented in Fig. 10.16(a–d). The top ranked uncertain data and the directions of their influence on the computed harvestable anchovy biomass are the same for both uncertainty importance measures. The largest uncertainty contributions come from data nos. 1, 11, 12, 24 and 33 over most of the simulated time span. Large values of nos. 1, 11, 24 and 33 tend to lead to small values of the computed harvestable anchovy biomass, while large values of no. 12 tend to lead to large values of the biomass.

The approximate correlation ratios are presented in Fig. 10.17(a–d)

The correlation coefficients obtained from raw as well as from rank transformed data and the correlation ratios identify the same set of five uncertain data as main contributors to the uncertainty of the computed temporal evolutions of the harvestable anchovy biomass and of the population of guano-birds. However, for some of the important uncertainty contributors the sign of the importance measure is changed. This change in sign explains why the uncertainty of the harvestable anchovy biomass extends more to larger computed values while that of the

Fig. 10.12 (continued) computed population of guano-birds. No. 1 is the growth rate of the bird population. No. 11 is the percentage of the bird population kept alive if only half the desired food supply is available and no. 12 is the anchovy biomass at which the anchovy fecundity starts to decline. No. 24 represents the start value for the biomass defined as an uncertain fraction of no. 12, while no. 33 is the mortality rate of the anchovy larvae

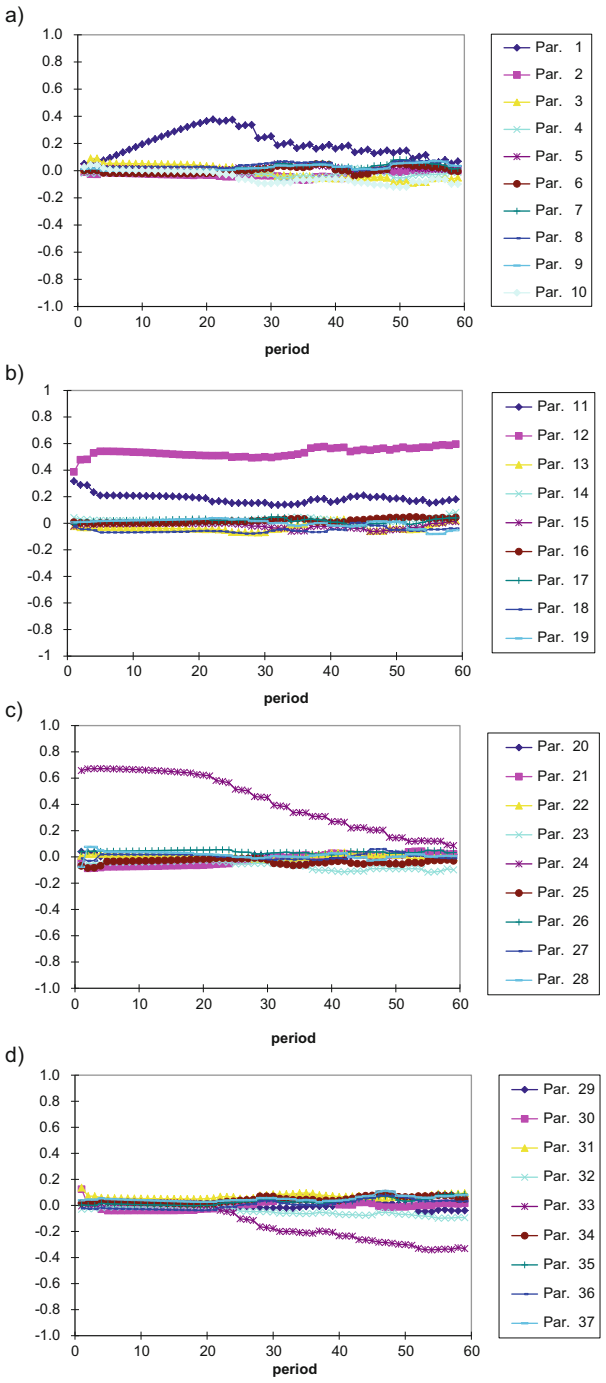


Fig. 10.13 (**a**–**d**) Rank correlation coefficients for the population of the guano-birds with respect to the uncertain data of Fig. 10.1

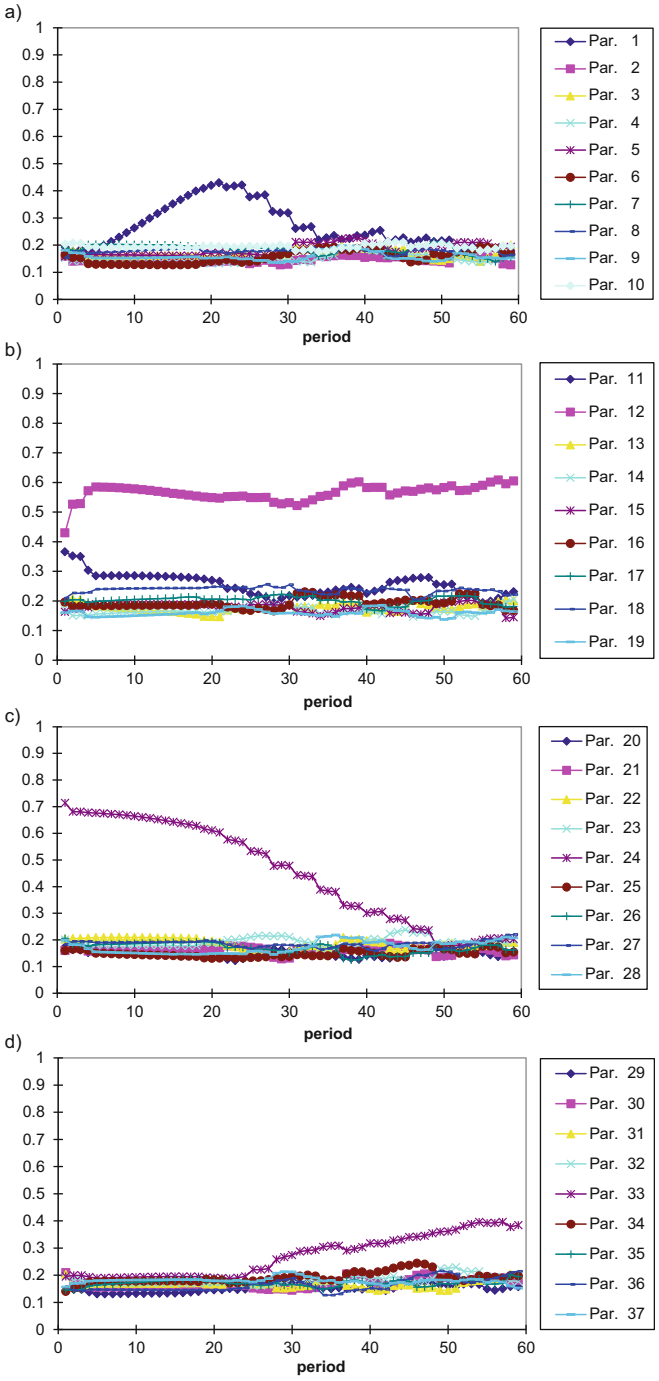


Fig. 10.14 (**a**–**d**) Approximate correlation ratios for the model result "population of the guano-birds" with respect to the uncertain data of Fig. 10.1

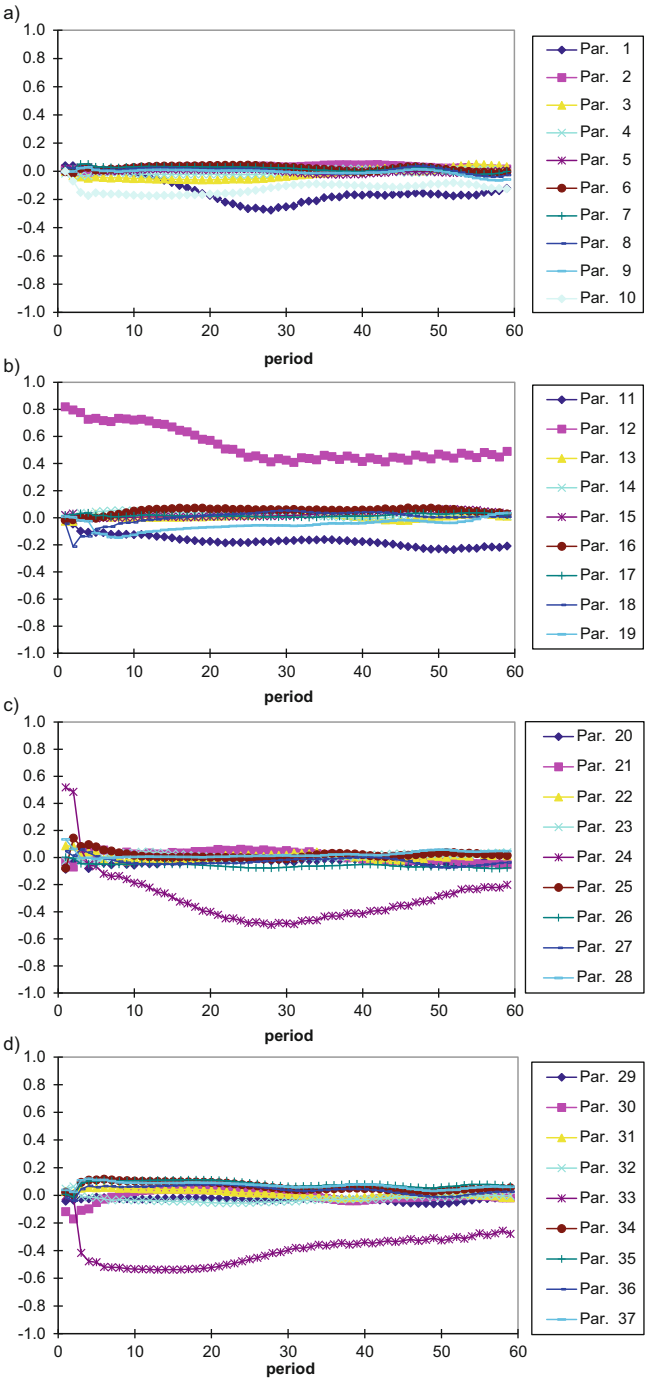


Fig. 10.15 (**a**–**d**) Correlation coefficients for the model result "harvestable anchovy biomass" with respect to the uncertain data of Fig. 10.1; the coefficients are estimated at each of the 60 points of time. Data nos. 1, 11, 12, 24 and 33 are identified as the main contributors to the uncertainty of the computed harvestable anchovy biomass. For their meaning, see the caption of Fig. 10.12 or see Fig. 10.1

population of guano-birds extends more to lower values if compared to the results obtained with best estimate data values and the best estimate El Nino sequence.

Regression coefficients cannot be obtained for the reason given above. Consequently, the coefficient of determination R^2 is not available to measure how much of the uncertainty in Figs. 10.10 and 10.11 is explained by the top ranked uncertain data. None of the uncertain data with large variance inflation factors of the regression coefficients is among the latter, and therefore, standardized regression coefficients have been obtained by admitting only the top ranked uncertain data into the regression model. The corresponding R^2 value is initially high for both model results (over or just below 0.8) and starts to decline after period 20. This decline is sharp for the model result "population of guano-birds". The R^2 value settles at around 0.5 at period 30 and fluctuates about this value for the rest of the time interval. The decline is smooth for the result "harvestable anchovy biomass". However, the R^2 value crosses the 0.5 line between period 40 and period 50 and continues to decline towards the end of the time interval. To get a clearer picture of the uncertainty importance over the time periods with low R^2 value, 1000 additional model runs were performed with only the 5 top ranked uncertain data sampled according to their subjective probability distributions and the El Nino event sequence varied but all other uncertain data set at their best estimate values. Figures 10.18 and 10.19 show the statistical tolerance limits for "population of guano-birds" and "harvestable anchovy biomass". Comparison with Figs. 10.10 and 10.11 demonstrates that the five top ranked uncertain data together with the uncertain El Nino sequence explain most of the uncertainty.

Figures 10.20 and 10.21 present the statistical tolerance limits for "population of guano-birds" and "harvestable anchovy biomass" if only the El Nino event sequence is varied and all uncertain data are set at their best estimate values. The resulting uncertainty range is moderate.

10.2.12 Conclusions

- Best estimate results and uncertainty analysis suggest that the fishing quota may be sustainable over the projected time period although the uncertainty of the computed harvestable anchovy biomass is rather large.
- The guano-bird population seems to be in a dynamic equilibrium with the anchovy population.
- Apart from the uncertainty about

  when an El Nino will strike and

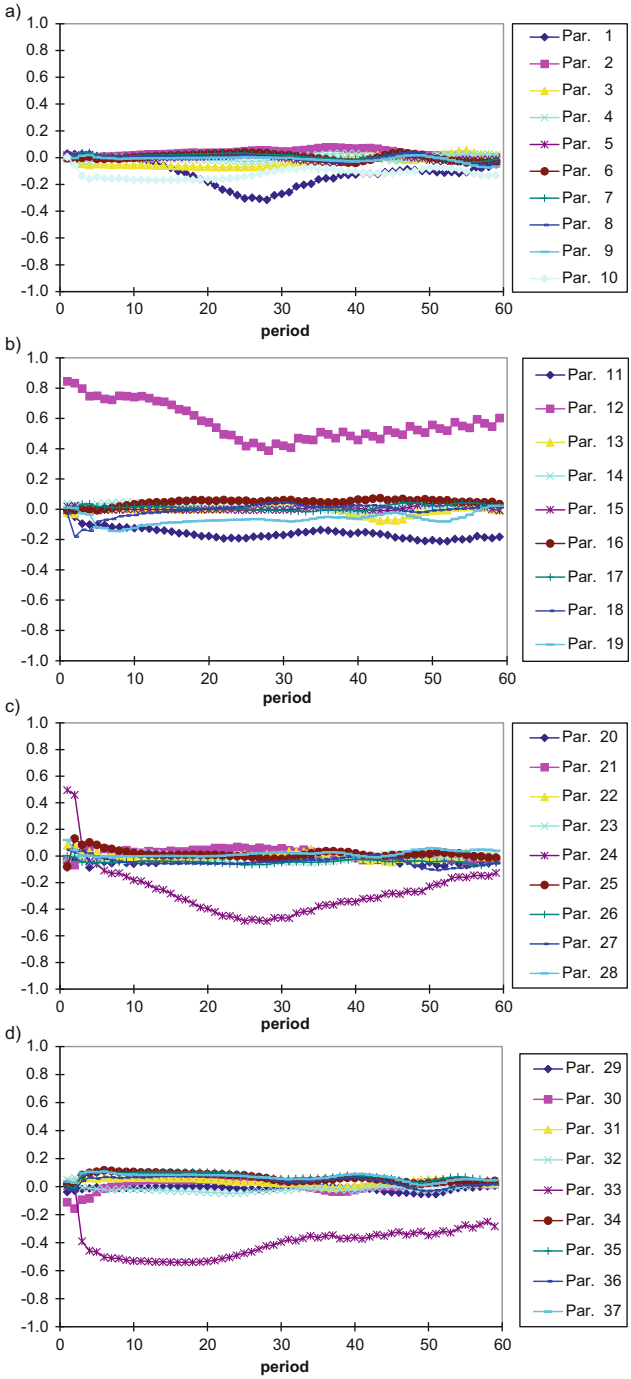


Fig. 10.16 (**a**–**d**) Rank correlation coefficients for the model result "harvestable anchovy biomass" with respect to the uncertain data of Fig. 10.1

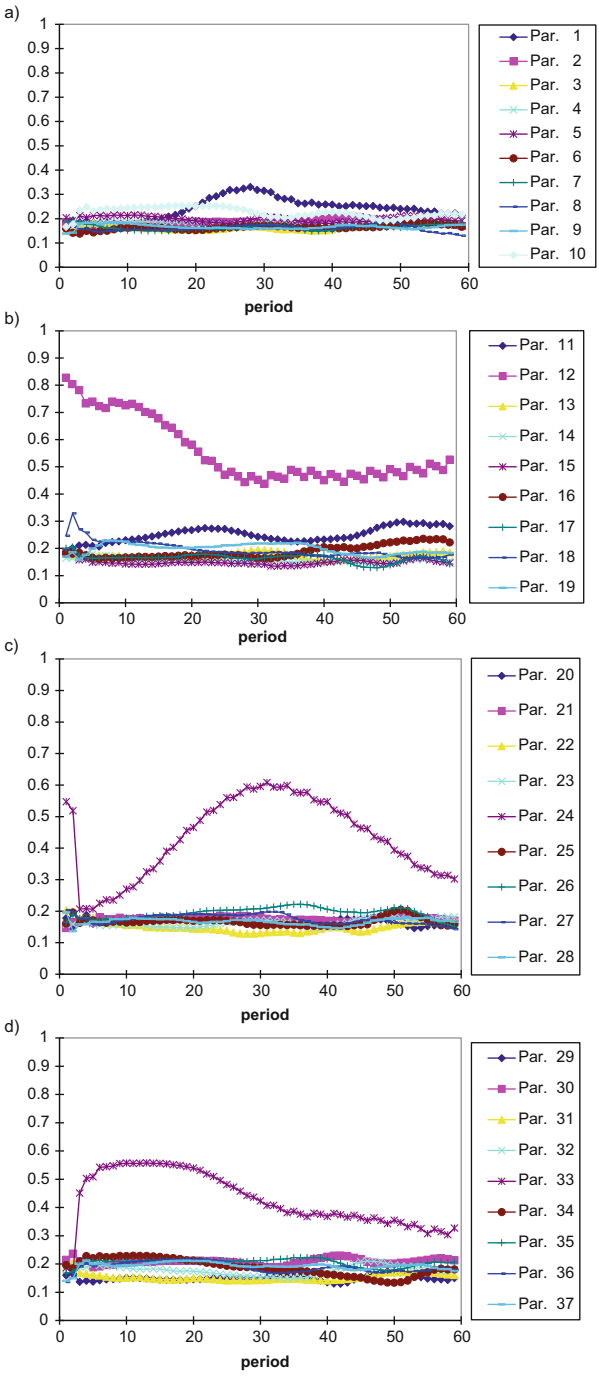


Fig. 10.17 (**a–d**) Approximate correlation ratios for the model result "harvestable anchovy biomass" with respect to the uncertain data of Fig. 10.1

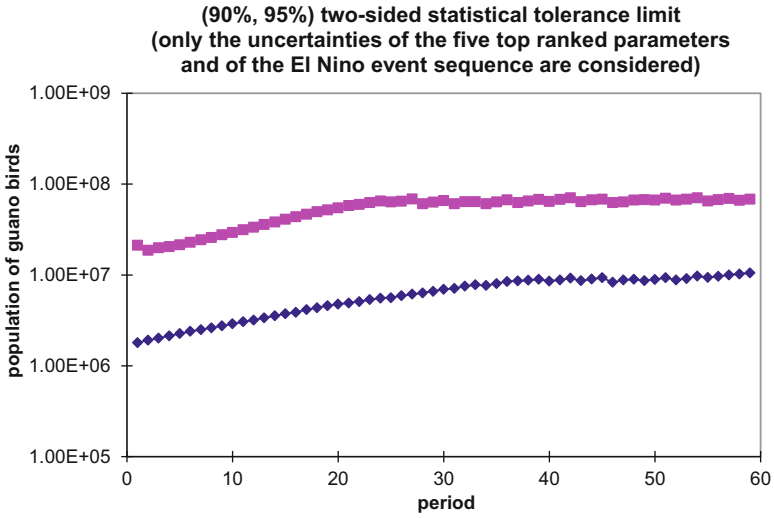


Fig. 10.18 The continuous connections of the local (at intervals of 4 months) (90%, 95%) two-sided statistical tolerance limits of the computed population of guano-birds. Clearly, most of the uncertainty, shown in Fig. 10.11, is explained by the five top ranked uncertain data and by the El Nino event sequence

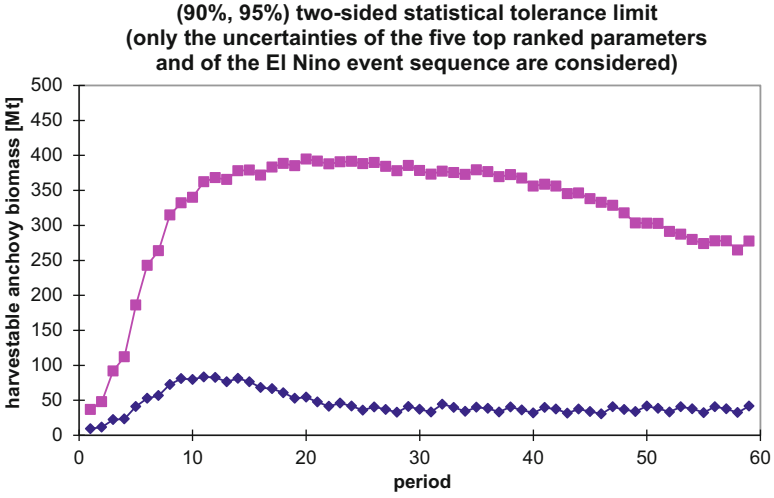


Fig. 10.19 The continuous connections of the local (at intervals of 4 months) (90%, 95%) two-sided statistical tolerance limits of the computed harvestable anchovy biomass. Clearly, most of the uncertainty shown in Fig. 10.10 is explained by the five top ranked uncertain data and by the El Nino event sequence

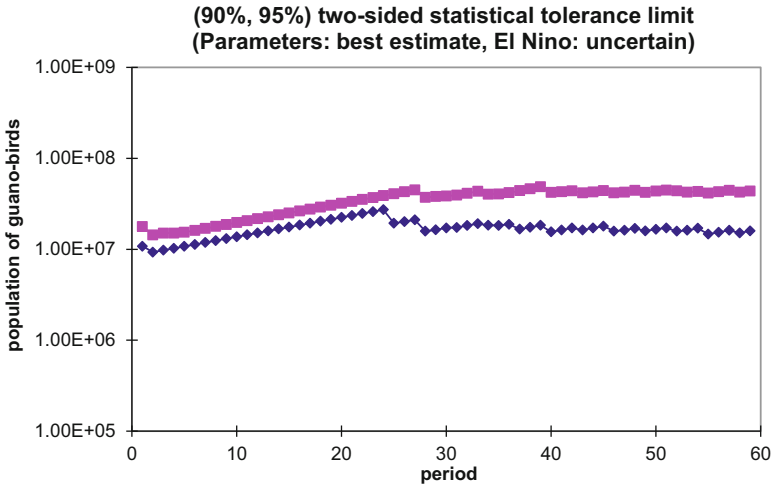


Fig. 10.20 The continuous connections of the local (at intervals of 4 months) (90%, 95%) two-sided statistical tolerance limits of the computed population of guano-birds; they are obtained using a random sample of size $N = 1000$ of El Nino event sequences, while all uncertain data are kept at their best estimate values

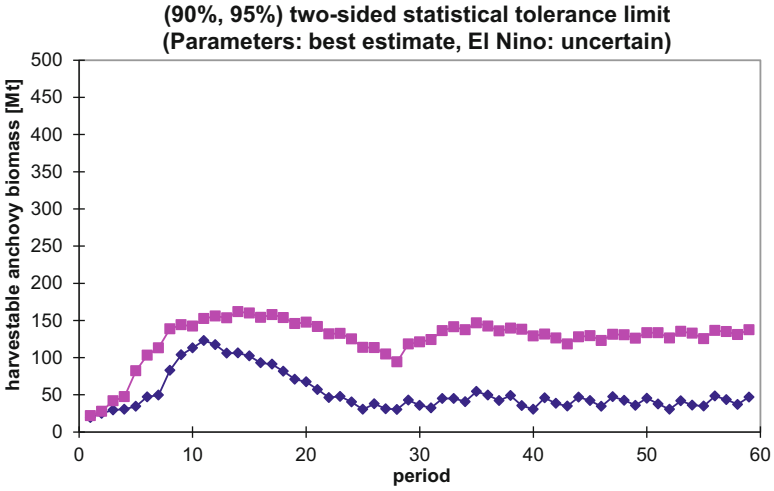


Fig. 10.21 The continuous connections of the local (at intervals of 4 months) (90%, 95%) two-sided statistical tolerance limits of the computed harvestable anchovy biomass. They are obtained using a random sample of size $N = 1000$ of El Nino event sequences, while all uncertain data are kept at their best estimate values

how severe it will be only five uncertain data are main contributors to uncertainty of the computed harvestable anchovy biomass. These are in the order of their uncertainty importance:
Biomass value at which fecundity starts to decline (par. no. 12)
Infant mortality rate of anchovies (par. no. 33)
Initial value of biomass (par. no. 24)
Exponent in food/survival relationship for guano-birds (par. no. 11)
Growth rate of guano-bird population (par. no. 1).

Improving the state of knowledge of these uncertain data would reduce the large uncertainty of the computed harvestable anchovy biomass is most effectively.

10.3 Uncertainty Analysis of Results from the Application of a Dose Reconstruction Model

A population of $I = 1000$ individuals was accidentally exposed to a carcinogenic contaminant. The release of the contaminant into the environment took place many years ago. Exposure of the individuals occurred via the intake of a specific foodstuff and during outdoor activities in the contaminated area.

10.3.1 The Assessment Question

In order to decide about compensation for the affected individuals, the decision-makers need to know the percentage of the exposed population with dose above a given limit value and who these individuals are. The dose value needs to be computed for each of the $I = 1000$ exposed individuals in order to answer these questions. The basic assessment question therefore is:

"What is the dose of individual i, $i = 1, \ldots, 1000$?". This question has a single true answer, namely the set of 1000 individual dose values. This set of values may be summarized by an empirical distribution and from this distribution in its complementary cumulative form the percentage of individuals with dose above the given limit value may be read.

The following notation is used in the subsequent subsections:

$\boldsymbol{d}$ the vector of the I unknown true values of the consumed quantity of the specific foodstuff

$\boldsymbol{d}^+$ the vector of the recalled (in interviews with the exposed individuals) values of the quantity consumed

$\boldsymbol{d}^*$ a vector of I possibly true values of the quantity consumed

$\boldsymbol{\delta}$ the vector of the unknown true recall errors
$\boldsymbol{\delta}^*$ a vector of possibly true recall errors
$\boldsymbol{\delta}^{\sim}$ an interim vector of possibly true recall errors
$\boldsymbol{m}$ the vector of the J unknown true contamination values of the foodstuff
$\boldsymbol{m}^+$ the vector of the measured contamination values
$\boldsymbol{m}^*$ a vector of J possibly true contamination values
$\boldsymbol{\mu}$ the vector of the unknown true measurement errors
$\boldsymbol{\mu}^*$ a vector of possibly true measurement errors
$\boldsymbol{\mu}^{\sim}$ an interim vector of possibly true measurement errors
$\boldsymbol{b}^*$ a vector of I possibly true individual exposure values due to ground contamination
$\boldsymbol{f}^*$ a vector of I possibly true individual-specific values of the coefficient used in the transfer of exposure into dose.

The following types of measurement error shall be considered:

- Berkson error: The **true** value is the sum or product of the **measured** value and a random variable that is statistically independent of the **measured** value.
- Classical error: The **measured** value is the sum or product of the **true** value and a random variable that is statistically independent of the **true** value.

Schafer and Gilbert (2006) present practical examples for both types of error together with a thorough explanation of the difference as well as of its consequences for regression analysis. Clearly, for both types, the true value is somewhere about the measured value. Its state of knowledge can be quantified by a subjective probability distribution that is based on the measured value plus all available information about the error distribution, i.e. about the random variable mentioned above. In the present example, one measurement result or recalled value is given per uncertain datum in the two sets of size I and J, respectively. Sampling independently from each of the I or J individual subjective probability distributions, to obtain a set of input values for one Monte Carlo simulation run of the uncertainty analysis, would be adequate if the errors are of the Berkson type. For this measurement error type, measured value and error are statistically independent. In the case of the classical measurement error, however, measured value and error are statistically dependent. Therefore, while each individually sampled value can be considered as possibly true, the set of input values cannot be considered as possibly true. Reason is that the statistical dependence was ignored by the sampling procedure, and therefore, the variance of the set of sampled values is bound to differ from that of the true values (see Sect. 3.3.1.1). Consequently, the set of independently sampled values is not suitable as input for a Monte Carlo simulation run of the uncertainty analysis. It requires the modelling of the statistical dependence between measured values and measurement errors, which characterizes the classical error, in order to obtain sets of input values that are suitable for the uncertainty analysis.

10.3.2 The Model

The model considers the total intake d_i of the foodstuff for each of the exposed individuals $i = 1, \ldots, 1000$. A recalled value d_i^+ is obtained in interviews with the individuals. It is subject to the so-called recall error δ_i. The error is of the additive classical measurement error type. The 1000 total intake values are uncertain input data of the model.

The model also considers $J = 50$ producers of the specific foodstuff. From each producer, one measured value m_j^+ of the contamination m_j, $j = 1, \ldots, 50$, is available. The 50 measurement values are subject to a multiplicative classical measurement error μ_j. The 50 contamination values are uncertain input data of the model.

For each individual, the model identifies at most three producers that supplied the individual with contaminated foodstuff. It also considers the proportion $c_{i,j}$ of the individual's total intake of the foodstuff that was supplied by each of the identified producers $j \in l(i)$ where $l(i)$ is the set of indices j of those producers that supplied the foodstuff to individual i. The values $c_{i,j}$ are uncertain input data of the model. Their uncertainty is neglected in this illustrative example.

The model also divides the contaminated area into $K = 5$ sub-areas of distinctly different average ground contamination, and it assigns each of the producers and individuals to their sub-area. $k(i)$ is the index of the sub-area where individual i lived at the time of exposure. The average ground contamination values are input data of the model and are derived from several measurements taken in the area. They are considered to be rather accurate. The ground contamination b_i, to which individual i was exposed, differs in the sense of a multiplicative Berkson error, from the average value for the sub-area in which individual i lived.

Additionally, the model considers a factor f_i that is to be used in the transfer of exposure into the dose for individual i. The individual values f_i vary, in the sense of a multiplicative Berkson error, about the mean value ϕ taken over many individuals. There is an estimate of ϕ available in the literature. It is uncertain how well it fits the mean over the exposed population. The values f_i are uncertain input data of the model, while the uncertain datum ϕ is the unknown mean value of the frequency distribution summarizing the variability of the transfer factor within the exposed population.

Furthermore, the model makes use of two additional data, namely r and t, in the transformation of exposure into dose. The values to be used for these two data are the same for all individuals but are imprecisely known. Their uncertainty is shared by the computations of the individual dose values.

The dose of individual i is computed as

$$q_i = \left[\left(\sum_{j \in l(i)} c_{i,j} m_j \right) d_i + b_i t \right] r f_i$$

See Sect. 10.3.8 for extensions of the assessment model.

As was mentioned in the introduction to this chapter, the model only serves illustration purposes. In practice, dose reconstruction models are rather complex and computationally demanding.

10.3.3 The Analysis Tool

The uncertainty and sensitivity analysis was performed with version 3.6 of the analysis software SUSA (Kloos 2015) .

SUSA supports the probabilistic modelling of the state of knowledge of uncertain data. It offers a large selection of distribution types, takes distribution truncations into account, derives distribution parameters from quantile values and plots the selected distributions. For the quantification of state of knowledge dependences, it accepts population measures of association, sample measures, conditional distributions, restrictions, complete dependence and functional relationships between uncertain data and provides scatter plots to illustrate the selected dependence quantification.

SUSA offers simple random sampling and Latin Hypercube sampling, derives distribution-free quantile estimates as well as statistical tolerance limits and performs statistical tests of distribution hypotheses and fitting of distributions to data. A choice of uncertainty importance measures is available for model output in the form of single values as well as of discretized functions of an independent variable such as time.

10.3.4 The Elicitation Process

There was no elicitation process as this example has been constructed for illustration purposes only.

10.3.5 The Potentially Important Uncertainties

The analysis considers the following uncertain data:

$P_1 - P_{1000}$	The individual recall errors δ_i
$P_{1001} - P_{1050}$	The errors μ_j of the contamination measurements of the foodstuff
$P_{1051} - P_{2050}$	The multiplicative Berkson error (factor to the average ground contamination) of the individual exposures b_i
$P_{2051} - P_{3050}$	The multiplicative Berkson error (factor to the average value ϕ) of the individual factor values f_i

P_{3051} The mean value ϕ around which the individual transfer factors f_i vary

P_{3052} Transformation parameter r (same value for all individuals)

P_{3053} Transformation parameter t (same value for all individuals)

Although true value and error are statistically independent in the classical measurement error situation, the value of their sample correlation may differ from zero. Therefore, the following uncertain data are additionally included in the analysis:

P_{3054} The coefficient of Pearson's sample correlation r($\boldsymbol{d}$, $\boldsymbol{\delta}$) between the true total quantities consumed and their recall errors

P_{3055} The coefficient of Pearson's sample correlation r(ln($\boldsymbol{m}$),ln($\boldsymbol{\mu}$)) between the natural logarithms of the true contamination values and of their measurement errors

10.3.6 The State of Knowledge Quantifications

$P_1 - P_{1000}$ Normal subjective probability distribution with mean value 0.0 and standard deviation 5.0 for each δ_i, $i = 1, \ldots, 1000$

$P_{1001} - P_{1050}$ Logarithmic normal (lognormal) subjective probability distribution with parameters 0.0 and 0.2 as mean value and standard deviation for the natural logarithm of each μ_j, $j = 1, \ldots, 50$; the mean value of the distribution for each μ_j is larger than 1.0 to account for measurement bias.

$P_{1051} - P_{2050}$ Lognormal subjective probability distribution with parameters -0.045 and 0.3 as mean value and standard deviation for the natural logarithm of each b_i, $i = 1, \ldots, 1000$

$P_{2051} - P_{3050}$ Lognormal subjective probability distribution with parameters -0.125 and 0.5 as mean value and standard deviation for the natural logarithm of each f_i, $i = 1, \ldots, 1000$

P_{3051} Lognormal subjective probability distribution with parameters -6.5 and 0.25 as mean value and standard deviation for the natural logarithm of ϕ

P_{3052} Lognormal subjective probability distribution with parameters 3.2 and 0.2 as mean value and standard deviation for the natural logarithm of r

P_{3053} Lognormal subjective probability distribution with parameters -5.5 and 0.5 as mean value and standard deviation for the natural logarithm of t

P_{3054} $T_{998}/(998 + T_{998}^2)^{1/2}$ where T_{998} is Student distributed with degree of freedom 998 (Rosner 1995)

P_{3055} $T_{48}/(48 + T_{48}^2)^{1/2}$ where T_{48} is Student distributed with degree of freedom 48

Par. No.	Short Par.Name	Full Parameter Name	Symbol	Unit	Reference Value	Best Estimate	Distribution Type	Distrib. Par. p1	Distrib. Par. p2	Minimum
1	TF	P3053	t	-	4.00E-03	4.60E-03	Log. Normal Distr.	-5.5	0.5	0
2	CF	P3051	Φ	-	0.0015	0.0016	Log. Normal Distr.	-6.5	0.25	0
3	REL	P3052	r	-	2.45E+01	2.50E+01	Log. Normal Distr.	3.2	0.2	0
4	RZ	Numerator for Student(998)	n_d	-			Normal Distr.	0	1	
5	RNM2	Nominator for Student(998)	a_d	-			Chi^2 Distr.	998		0
6	PZ	Numerator for Student(48)	n_c	-			Normal Distr.	0	1	
7	PNM2	Nominator for Student(48)	a_c	-			Chi^2 Distr.	48		0

Fig. 10.22 SUSA output: documentation sheet for the uncertain data varied in the outer simulation loop

T_{998} and T_{48} are each expressed using a standard normal distributed quantity and an independent Chi^2 distributed quantity with the degree of freedom equal to the number of data points minus 2.

Figure 10.22 presents the SUSA Output "Documentation Sheet" for the uncertain data varied in the outer simulation loop.

10.3.7 State of Knowledge Dependences

State of knowledge dependence between P_1, …, P_{1000} is introduced through the condition

$r(\boldsymbol{d}^+, \boldsymbol{\delta}^*) = g(r(\boldsymbol{d}, \boldsymbol{\delta}), \boldsymbol{\delta}^*, \boldsymbol{d}^+)$,

and state of knowledge dependence between P_{1001}, …, P_{1050} is introduced through the condition

$$r(\ln(\boldsymbol{m}^+), \ln(\boldsymbol{\mu}^*)) = g(r(\ln(\boldsymbol{m}), \ln(\boldsymbol{\mu})), \ln(\boldsymbol{\mu}^*), \ln(\boldsymbol{m}^+)).$$

r(.,.) stands for Pearson's sample correlation coefficient and the function g is given in the next subsection for possibly true values r*(.,.) to the unknown sample correlation coefficients r(.,.) needed in order to have $s^2_{d^*} = s^2_d$ and $s^2_{\ln(m^*)} = s^2_{\ln(m)}$, respectively (see Sect. 3.3.1.1).

10.3.8 Propagation of the State of Knowledge Quantifications Through the Model

The Monte Carlo simulation with SUSA proceeds in two nested simulation loops of sample size N for the outer loop (chosen to be 1000) and of sample size $I = 1000$ for the inner loop. In every simulation of the outer loop, SUSA provides one possibly true value each for ϕ, r, t, r($\boldsymbol{d}$, $\boldsymbol{\delta}$), r(ln($\boldsymbol{m}$),ln($\boldsymbol{\mu}$)) sampled according to the subjective probability distributions expressing the state of knowledge for P_{3051} to P_{3055} (see Fig. 10.22). These values are then used in the inner loop. The inner loop runs over the $I = 1000$ individuals.

The assessment model has been amended in order to

- Provide in the inner loop the sample values of the vectors $\boldsymbol{\delta}^{\sim}$, $\boldsymbol{\mu}^{\sim}$, $\boldsymbol{b}^{*}$, $\boldsymbol{f}^{*}$ drawn according to the subjective probability distributions specified for $P_1, \ldots, P_{3050}$.
- Post-process in the inner loop the vector $\boldsymbol{\delta}^{\sim}$ by permuting its components such that for the permuted vector $\boldsymbol{\delta}^{*}$ the sample correlation with $\boldsymbol{d}^{+}$ is

$$\mathrm{r}^{*}\left(\boldsymbol{d}^{+},\boldsymbol{\delta}^{*}\right) = \left\{ s_{\delta^{\sim}} + \mathrm{r}^{*}(\boldsymbol{d},\boldsymbol{\delta})\left[-\mathrm{r}^{*}(\boldsymbol{d},\boldsymbol{\delta}) s_{\delta^{\sim}} + \left\{ s_{\delta^{\sim}}^{2}\left[\mathrm{r}^{*2}(\boldsymbol{d},\boldsymbol{\delta}) - 1\right] + s_{d^{+}}^{2} \right\}^{\frac{1}{2}} \right] \right\} / s_{d^{+}}$$

  [see Eq. (13′) in Hofer (2008)].

  $\mathrm{r}^{*}(\boldsymbol{d},\boldsymbol{\delta})$ is the value sampled for P_{3054} by SUSA, $\boldsymbol{\delta}^{*}$ is the vector of the permuted components of $\boldsymbol{\delta}^{\sim}$, $s_{\delta^{\sim}}$ is the sample standard deviation of the components of $\boldsymbol{\delta}^{\sim}$ and $s_{d^{+}}$ is the sample standard deviation of the recalled values. The seven steps of the permutation algorithm are described in Sect. 3.3.1.1. The model run uses the possibly true vector $\boldsymbol{d}^{*} = \boldsymbol{d}^{+} - \boldsymbol{\delta}^{*}$ of consumed quantities of the contaminated foodstuff.
- Post-process in the inner loop the vector ln($\boldsymbol{\mu}^{\sim}$) by permuting its components such that for the permuted vector ln($\boldsymbol{\mu}^{*}$) the sample correlation with ln($\boldsymbol{m}^{+}$) is

$$\begin{aligned}\mathrm{r}^{*}\left(\ln\left(\boldsymbol{m}^{+}\right),\ln\left(\boldsymbol{\mu}^{*}\right)\right) = \Big\{ & s_{\ln\left(\mu^{\sim}\right)} + \mathrm{r}^{*}\left(\ln\left(\boldsymbol{m}\right),\ln\left(\boldsymbol{\mu}\right)\right)\Big[-\mathrm{r}^{*}\left(\ln\left(\boldsymbol{m}\right),\ln\left(\boldsymbol{\mu}\right)\right) s_{\ln\left(\mu^{\sim}\right)} \\ & + \Big\{ s_{\ln\left(\mu^{\sim}\right)}{}^{2}\left[\mathrm{r}^{*2}\left(\ln\left(\boldsymbol{m}\right),\ln\left(\boldsymbol{\mu}\right)\right) - 1\right] + s_{\ln\left(m^{+}\right)}^{2} \Big\}^{\frac{1}{2}} \Big] \Big\} / s_{\ln\left(m^{+}\right)}.\end{aligned}$$

  $\mathrm{r}^{*}(\ln(\boldsymbol{m}), \ln(\boldsymbol{\mu}))$ is the value sampled for P_{3055} by SUSA, ln($\boldsymbol{\mu}^{*}$) is the vector of the permuted components of ln($\boldsymbol{\mu}^{\sim}$), $s_{\ln\left(\mu^{\sim}\right)}$ is the sample standard deviation of the natural logarithms of the components of $\boldsymbol{\mu}^{\sim}$ and $s_{\ln\left(m^{+}\right)}$ is the sample standard deviation of the natural logarithms of the measured contamination values of the produced foodstuff. The model run uses the possibly true vector $\boldsymbol{m}^{*} = \exp\left(\ln(\boldsymbol{m}^{+}) - \ln\left(\boldsymbol{\mu}^{*}\right)\right)$ of contamination values.

A possibly true set of $I = 1000$ dose values is then computed in the inner loop using the sampled data values.

A new set of values is sampled in the outer loop, and this is followed by a run through the inner loop as described above. This sequence of simulations is repeated $N = 1000$ times for the purpose of uncertainty analysis.

10.3.9 Why Two Monte Carlo Simulation Loops?

If the error is additive, a value sampled at random according to the error distribution was added to the true value to give the measurement value. This added value is unknown just as the true value is. Any value sampled at random according to the error distribution could be the unknown added value. Subtracting it from the measurement value therefore renders a value that can be considered as possibly true. It would therefore be entirely appropriate to perform $I = 1000$ independent Monte Carlo simulations of sample size $N = 1000$ each instead of proceeding with two nested simulation loops, as long as the goal of the assessment was only to independently quantify the uncertainty of the dose received by each of the $I = 1000$ individuals in the exposed population. The $N = 1000$ dose values, computed for each individual, could then be compared to the given limit value and the subjective probability for compliance with (or violation of) the limit could be estimated for each individual's dose independently (see also the remark below).

Since the assessment question also asks for the percentage of the exposed population with dose value above the limit, $I = 1000$ independent Monte Carlo simulations are clearly not the correct approach. This approach would ignore the effect of the shared uncertainties and of the correlations between the recalled values $\boldsymbol{d}^+$ and their errors $\boldsymbol{\delta}$ as well as the measurement values $\boldsymbol{m}^+$ of the contamination and their errors $\boldsymbol{\mu}$. The answer to this question requires a possibly true set of $I = 1000$ dose values and not just $I = 1000$ independently sampled possibly true dose values. This is not merely a semantic difference. It is an important technical difference that has to do with the state of knowledge dependence of the individual dose values. This dependence is due to uncertainties that are fully or partially (like the two joint empirical distributions of $I = 1000$ as well as $J = 50$ measurement values and their errors)[1] shared by the computations of the individual dose values. At this point, it may be helpful to categorize the uncertain data P of a dose calculation for a cohort of I individuals as in Simon et al. (2015).

Due to the shared uncertain data and due to data with common uncertainty contributors, the states of knowledge of the I true dose values are dependent and are thus quantified by a joint subjective probability distribution. As a consequence, the dose values cannot be sampled independently according to their marginal subjective probability distributions if the goal of the assessment is to simulate a possibly true set of I dose values. Once a dose value has been selected for an individual, according to the corresponding marginal subjective probability distribution, the set of dose values for all other individuals in the exposed population can no longer be chosen at random according to their marginal subjective probability distributions. This becomes immediately clear if one considers, for instance, the uncertain data nos. 3051–3053. The dose value chosen for the first individual was

[1]The choice of possibly true vectors of measurement errors has to satisfy the condition that the variance of the difference between measurement values and possibly true errors must be a possibly true variance of the true values. In particular, it must not be inflated by the variance of the errors.

computed with specific values for these uncertain data, and therefore, the dose values for all other individuals would need to be chosen from only those that were computed with the same values for these uncertain data. If sampled independently, the set of I dose values would exhibit inter-individual variability that is inappropriately inflated by the variance contributions from shared uncertainties. This would render them useless for dose–response estimation.

$N = 1000$ possibly true sets of $I = 1000$ individual dose values each are obtained through a Monte Carlo procedure performed in two nested simulation loops [often called 2DMC (Simon et al. 2015)]. The method used in this example is the procedure described in Sect. 9.4. It is used here for the uncertainty analysis of a dose reconstruction although all uncertainties are of Type B. The reason is that partially or fully shared uncertainties require an outer simulation loop while individual-specific uncertainties are accounted for in the inner simulation loop.

In this example, the procedure using two nested simulation loops provides N possibly true sets of I dose values each as follows:

In the outer loop:

(a) Sample one value each for the uncertain data t, ϕ, r, n_d, a_d, n_c, a_c.
(b) Sample one value each for the uncertain data $m_j, j = 1, \ldots, J$; the input values m_j^+, $j = 1, \ldots, J$ are subject to classical measurement error following the same error distribution. Therefore, proceed as in Sect. 3.3.1.1 using the sampled values n_c and a_c to obtain the sample value $t_{48} = n_c / \left(\frac{a_c}{48}\right)^{\frac{1}{2}}$ and use it in (3.16) to compute the sample value $r^*(\ln(\boldsymbol{m}), \ln(\boldsymbol{\mu}))$. Use the values $m_j^* = \exp(\ln(m_j^+) - \ln(\mu_j^*))$ as possibly true values for $m_j, j = 1, \ldots, J$.
(c) Sample a value for the uncertain datum ϕ.
(d) The input values d_i^+, $i = 1, \ldots, I$ are subject to classical measurement error following the same error distribution. Therefore, proceed as in Sect. 3.3.1.1 using the sampled values n_d and a_d to obtain the sample value of $t_{998} = n_d / \left(\frac{a_d}{998}\right)^{\frac{1}{2}}$ and use it in (3.16) to compute the sample value $r^*(\boldsymbol{d}, \boldsymbol{\delta})$. Use the values $d_i^* = d_i^+ - \delta_i^*$ as sample values that are possibly true for d_i, $i = 1, \ldots, I$ in the inner simulation loop.

Then, in the inner loop, for $i = 1, \ldots, I$

(e) Sample a value f_i^* for the uncertain datum f_i.
(f) Take the value d_i^* for d_i and sample b_i^* for b_i.
(g) Compute the dose value using the sampled values for the uncertain data t, r, n_d, a_d, n_c, a_c, ϕ, f_i, b_i, d_i and $m_j, j \in \mathrm{l}(i)$ (i.e. the contamination values with respect to the suppliers of the foodstuff to individual i).

At the end of the inner loop:

(h) The possibly true set of I dose values may now be used to determine a possibly true percentage of the individuals with dose value above the limit. To this end, the I dose values are arranged in the form of a complementary cumulative distribution function.

Repeat steps (a), (b), (c), (d) and (h) and the execution of the inner simulation loop N times, if N is the chosen sample size of the outer Monte Carlo simulation. The results are N possibly true sets of I dose values each. These N sets may be used to quantify the effect of dose uncertainty on the state of knowledge of the true percentage of individuals with dose value above the limit.

The results from the procedure using two nested simulation loops also answer the first question mentioned above, namely to determine for each individual the subjective probability for the dose value to exceed the given limit value. A Monte Carlo sample of N possibly true dose values can be obtained for any specific individual by simply collecting the N dose values computed for the individual over the N repetitions of the outer simulation loop. Thus, the method addresses questions about the uncertainty of the true dose for any specific individual as well as questions about the uncertainty of the true set of N dose values for the exposed population. The computational effort involved in the two nested simulation loops is not different to the approach running I independent Monte Carlo simulations of size N. In both cases, $I*N$ individual dose calculations are required.

Uncertainty importance measures can be derived for the mean value, standard deviation, the percentage of individuals with dose above a given limit value and any other distribution characteristic of the probability distribution summarizing the variability within the set of I individual dose values. This task uses the N sets of I dose values and the N sets of values sampled for the uncertain data in the outer Monte Carlo simulation loop. The ranking indicates where the state of knowledge should be improved in order to reduce dose uncertainty most effectively.

A ranking of uncertain data may also be obtained for any individual's dose value by using the N dose values collected over the outer loop and the N values of all involved uncertain data sampled in the outer and inner loop of the simulation.

10.3.10 Answering the Assessment Questions

The assessment questions of Sect. 10.3.1 can only be answered by subjective probability distributions that quantify the combined influence of the uncertain data involved in the dose reconstruction.

For instance, Fig. 10.23 answers the question "What is the dose of individual $i = 22$?". From the distribution, it becomes evident that there is only a negligible subjective probability for the dose to be larger than the limit value $d_{lim} = 1000$. An answer to the question would be:

"The dose of individual $i = 22$ is smaller or equal to 400 with subjective probability of approximately 0.95."

The 95% quantile was estimated from a sample of $N = 1000$ dose values. The sample was obtained by Monte Carlo simulation, i.e. it was drawn at random according to the subjective probability distribution for the dose of individual $i = 22$. This subjective probability distribution results from the propagation of the combined state of knowledge for the involved uncertain data through the model.

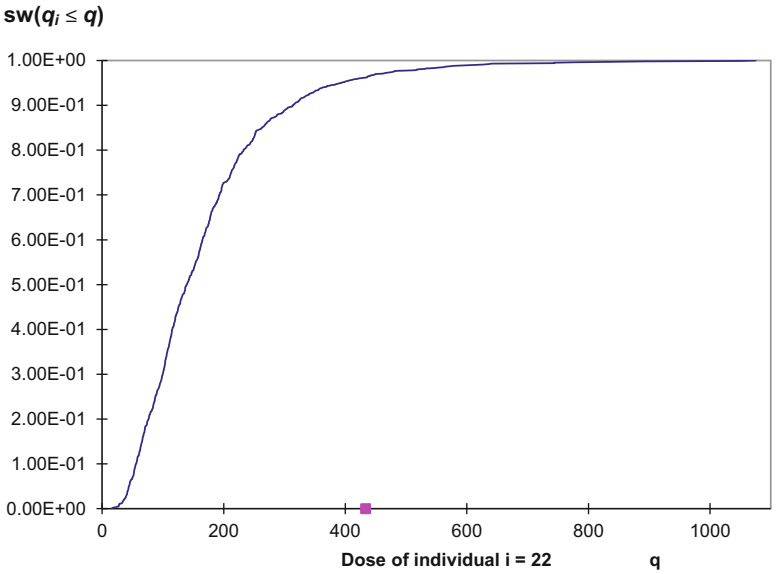


Fig. 10.23 Empirical subjective probability distribution for the dose of individual $i = 22$; the distribution shows the combined influence of the uncertain data on the reconstruction of the dose value. An upper (95%, 95%) statistical tolerance limit of the dose value is shown on the abscissa. At a confidence level of at least 95%, the subjective probability is at least 0.95 for the dose of individual $i = 22$ not to exceed 434

In order to account for the possible estimation error (also called sampling error), the upper (95%, 95%) statistical tolerance limit, indicated on the abscissa of Fig. 10.23, may be used in answering the assessment question. Its meaning is:

"At a confidence level of at least 95% the dose of individual $i = 22$ is smaller or equal to 434 with subjective probability of at least 0.95."

or

"At least 95% of the dose values of individual $i = 22$ that have to be considered as possibly true, given the quantified state of knowledge at the level of the uncertain data, are below 434 at a confidence level of at least 95%."

These uncertainty statements refer to the true dose of individual $i = 22$. In subchapter 7.2, it was explained that this reference to the true value is based on the following assumptions:

- The encoded computer model is not seriously flawed.
- The most important contributors to uncertainty of its results have been accounted for in the uncertainty analysis (the combined effect of all those not accounted for may be neglected).
- The states of knowledge at the level of parameters, models and input data have been appropriately expressed by a joint subjective probability distribution.

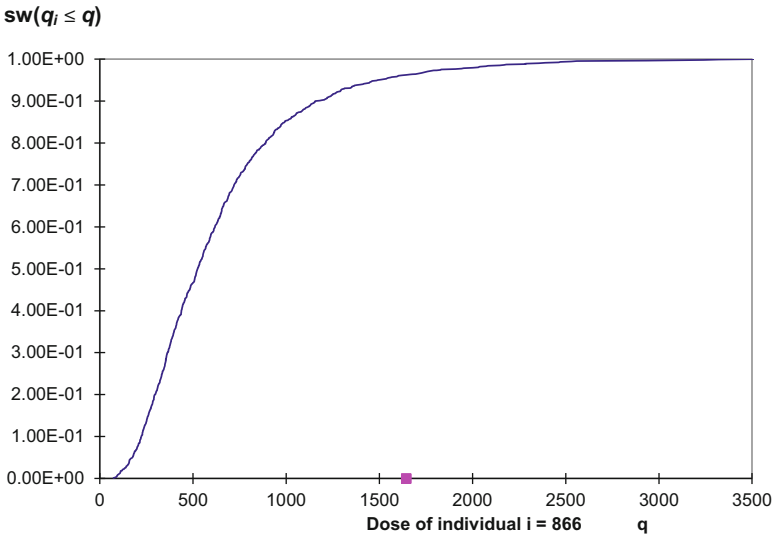


Fig. 10.24 Empirical subjective probability distribution for the dose of individual $i = 866$; the distribution shows the combined influence of the uncertain data on the reconstruction of the dose value. An upper (95%, 95%) statistical tolerance limit of the dose value is shown on the abscissa. At a confidence level of at least 95%, the subjective probability is at least 0.95 for the dose of individual $i = 866$ not to exceed 1645

– Any changes to the interpretation of "subjective probability" (see Sect. 3.4.1) are of only minor consequence.

If there is doubt whether the first assumption does apply, then the uncertainty statements may refer only to the computed dose for individual $i = 22$. If the second assumption does not apply, then the uncertainty statements are only of limited use. The uncertainty statements may be useless (or only of illustrative value, as is the case for the present example) in case the third assumption does not apply.

The corresponding information for individual $i = 866$ is shown in Fig. 10.24. Here, the estimate of the subjective probability for the dose value to exceed the limit value $d_{lim} = 1000$ is between 0.14 and 0.15. An upper (95%, 95%) statistical tolerance limit of the dose is again indicated on the abscissa.

Remark

The subjective probability for a dose value above d_{lim} may exceed a decision limit for several individuals. The dose value may not be above d_{lim} for all of these individuals in the same possibly true set of dose values. Since only one such set can be true and since the state of knowledge does not permit to decide which, compensation will have to be paid to all of those individuals with subjective probability larger than the decision limit for a dose value above d_{lim}.

The situation with only two clearly separated possibly true wind directions that could have prevailed at the time of the release of the contaminant may serve as an

example. It is unknown which of the two directions is the true wind direction, and therefore, the area affected by the contaminant cloud is unknown. Dose values above d_{lim} may have been computed for individuals from both areas under the assumption that the wind was blowing in their direction. The subjective probability for a dose value above d_{lim} may be above the decision limit for several individuals from both areas. Yet, it is obvious that only individuals from one area could have received a dose above d_{lim}. However, it will remain unknown which of the two areas.

The question asking for the percentage of individuals with dose above the limit value $d_{lim} = 1000$ is answered by the subjective probability distribution in Fig. 10.25. A two-sided (90%, 95%) statistical tolerance limit is indicated on the abscissa. It says:

"At a confidence level of at least 95%, the percentage of individuals with dose above the limit value $d_{lim} = 1000$ is between 0.1% and 5.7% with subjective probability of at least 0.90." Again, this reference to the true percentage (instead of only to the computed percentage) is appropriate only under the condition that the assumptions listed above hold.

The uncertainty range shown in Fig. 10.25 is rather wide. It will most likely be considered as too wide for decision-making. The uncertainty importance measures will provide guidance as to where the state of knowledge needs to be improved in order to reduce this substantial uncertainty most effectively.

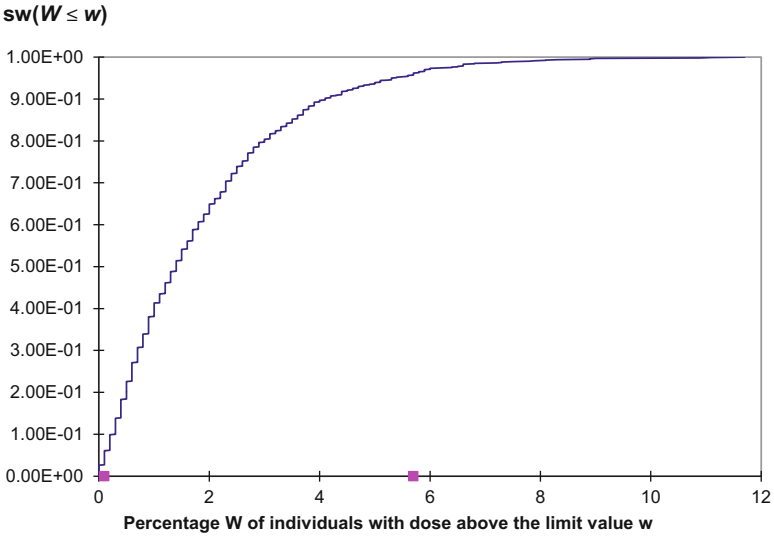


Fig. 10.25 Empirical subjective probability distribution for the percentage of individuals with dose above the limit value; the distribution shows the combined influence of the uncertain data on the percentage. A two-sided (90%, 95%) statistical tolerance limit of the percentage is shown on the abscissa. At a confidence level of at least 95%, the subjective probability is at least 0.90 for the percentage of individuals with dose above the limit value to lie between 0.1% and 5.7%

It should be noted that the answer to the question asking for the percentage of individuals with dose above d_{lim} is not affected by the dilemma discussed in the remark made above. Reason is that each possibly true percentage value is computed from only one run of the outer simulation loop and therefore uses one possibly true set of dose values.

The compilation of uncertain data in Sect. 10.3.5 contains 1000 recalled and 50 measurement values subject to error. Inevitably, there is some sample correlation between the true values and their recall errors and between the true values of the measurands and their measurement errors. This correlation increases or reduces the sample variance of the recalled or measurement values, depending on the sign of the sample correlation coefficient. The measurement values and their errors are correlated as well since the measurement values are $\boldsymbol{x} = \boldsymbol{z} + \boldsymbol{\varepsilon}$. $\boldsymbol{x}$ is the vector of measurement values, $\boldsymbol{z}$ the vector of true values and $\boldsymbol{\varepsilon}$ the vector of errors. If one were to use $\boldsymbol{z}^* = \boldsymbol{x} - \boldsymbol{\varepsilon}^*$ as "possibly true" values with $\boldsymbol{\varepsilon}^*$ the vector of possibly true errors sampled independently of $\boldsymbol{x}$ according to the subjective probability distribution for the measurement error, one would ignore this correlation in the Monte Carlo simulation. As a consequence, one would arrive at sets of "possibly true" values with sample variance even larger than that of the measurement values since $s^2_{Z^*} = s^2_X + s^2_{\varepsilon^*}$. The effect of this variance inflation on the reconstructed dose values can be seen from a comparison of the scatter plots in Figs. 10.26 and 10.27. It becomes also apparent from a comparison of Figs. 10.25 and 10.28.

The increased inter-individual variability shown by each of the N dose vectors in Fig. 10.27 makes them useless for dose–response estimation since the higher variability will lead to a lower estimate for the dose–response coefficient.

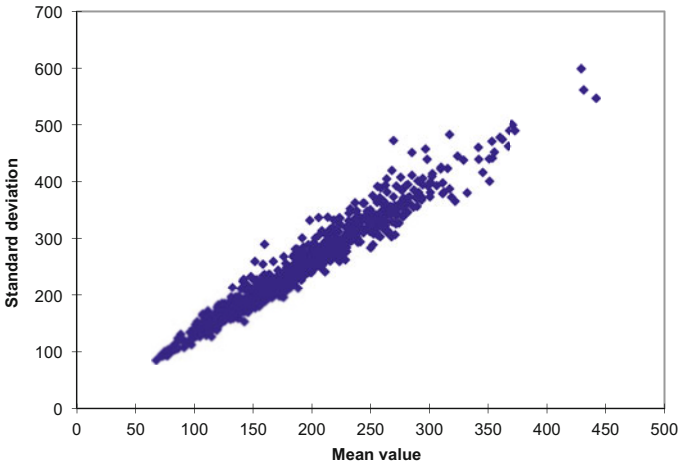


Fig. 10.26 Scatter plot of mean value versus standard deviation of the $N = 1000$ sampled dose vectors

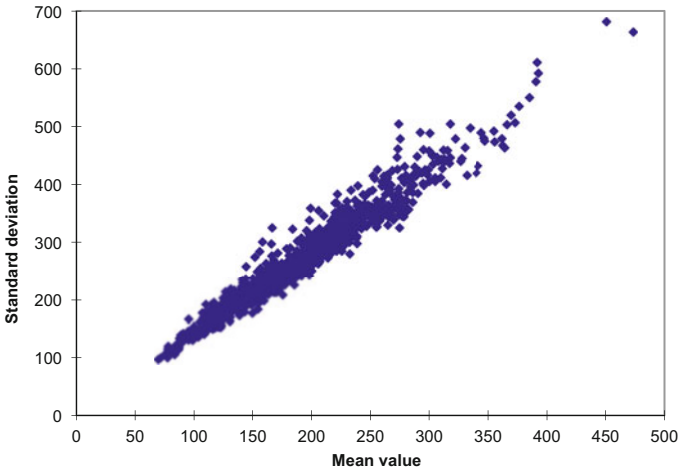


Fig. 10.27 Scatter plot of mean value versus standard deviation of the $N = 1000$ sampled dose vectors. The dependence between measurement or recalled values and their errors was ignored this time. As a consequence, the cloud of data points is shifted upwards to higher standard deviation values if compared to Fig. 10.26

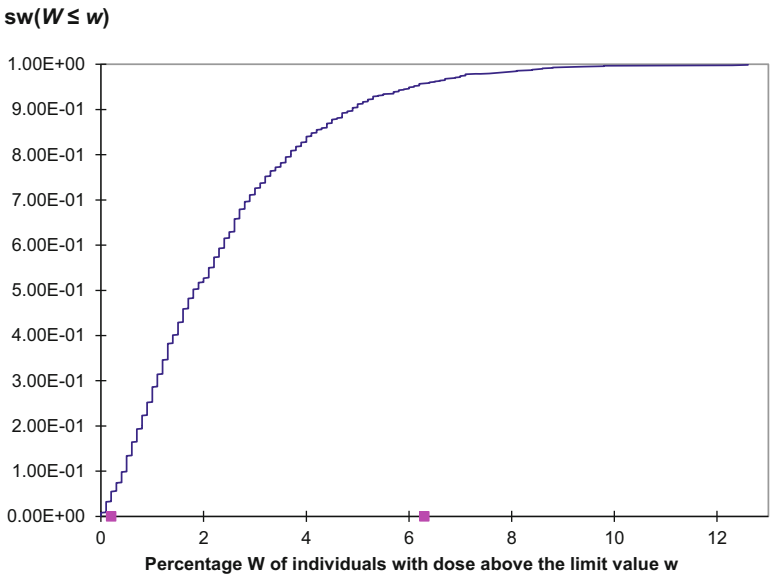


Fig. 10.28 Empirical subjective probability distribution for the computed percentage of individuals with dose above the limit value; the distribution shows the combined influence of the uncertain data on the percentage. The dependence between measurement or recalled values and their errors was, however, ignored. This caused the shift to higher standard deviation values illustrated by Fig. 10.27, and consequently, the subjective probability distribution for the computed percentage of individuals with dose above the limit value is shifted to higher percentages

10.3.11 Uncertainty Importance Statements for Selected Model Results

Figure 10.25 indicates that the computed percentage of individuals with dose above the limit value is rather uncertain. Within the random sample of size $N = 1000$ (1000 dose vectors of $I = 1000$ dose values each), the percentage ranges from 0% to over 11%. It will be of interest to know where the main uncertainty contributions to this wide range come from. Figure 10.29 shows Pearson's correlation coefficients as uncertainty importance measures for those uncertain data that are handled by SUSA (see Fig. 10.22). The uncertain data nos. 4–7 are used to represent the two sample correlation coefficients P_{3054} and P_{3055}. Obviously, the main uncertainty contributions come from the uncertain data nos. 2 (P_{3052}) and 3 (P_{3053}). They have the same value for all individuals (shared uncertainty). The sign of their uncertainty importance measure is positive. This indicates that changing the values of these uncertain data leads to a change of the computed percentage in the same direction. Improving the state of knowledge for these two uncertain data would reduce the uncertainty of the percentage of individuals with dose above the limit value most effectively.

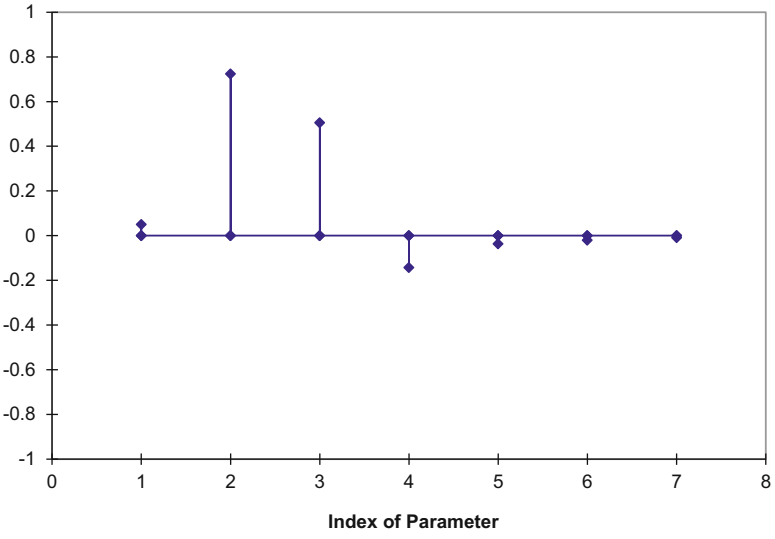


Fig. 10.29 Correlation coefficients as uncertainty importance measures for the percentage of individuals with dose above the limit value, obtained with respect to the uncertain data handled in the outer simulation loop by SUSA (see Fig. 10.22). The uncertain data nos. 2 and 3 are clearly the main contributors to the uncertainty of the computed percentage. The larger their value, the larger is the percentage of individuals with dose above the limit

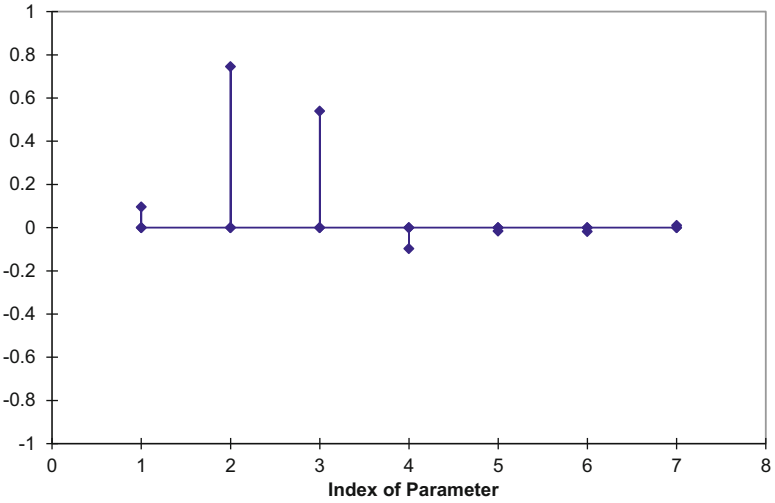


Fig. 10.30 Standardized regression coefficients as uncertainty importance measures for the percentage of individuals with dose above the limit value, obtained with respect to the uncertain data handled in the outer simulation loop by SUSA (see Fig. 10.22). Again, the uncertain data nos. 2 and 3 are the main contributors to uncertainty. The larger their value, the larger is the computed percentage of individuals with dose above the limit. The value of the coefficient of determination R^2 is 0.83

Comparison with the standardized regression coefficients in Fig. 10.30 shows a slight change for the uncertain datum no. 4. The only spurious correlation of any consequence in the sample of 1000 sets of data values exists between nos. 4 and 2. The correlation coefficient for no. 4 is partly due to this spurious correlation.

The coefficient of determination R^2 is 0.83, saying that a little less than 83% of the variance of the subjective probability distribution in Fig. 10.25 may be explained by a linear relationship between the computed percentage of individuals with dose above the limit value and the values thought to be true for the uncertain data nos. 2 and 3.

The situation is clearly different for the dose values to individuals no. 22 and no. 866 where the R^2 values are only 0.18 and 0.21 for linear least squares regression models in the uncertain data of Fig. 10.22 (varied in the outer Monte Carlo loop). A separate calculation of Pearson's correlation coefficient between the individual dose values and the corresponding individual components of the vectors $\boldsymbol{\delta}^*$, $\boldsymbol{b}^*$ and $\boldsymbol{f}^*$ (used in the inner Monte Carlo loop) was performed applying the scatter plot option in SUSA. By far the largest correlation coefficient is computed for the individual transfer parameters f_{22} and f_{688}, respectively, namely 0.857 for the dose of individual no. 22 and 0.88 for the dose of individual no. 866. Different to the uncertainty of the percentage of individuals with dose above the limit, the uncertainty of the individual dose in Figs. 10.23 and 10.24 could be most effectively reduced by improving the state of knowledge of the uncertain individual transfer parameters.

References

Caulkins, J., Marrett, R., & Yates, A. (1985). Peruvian anchovy population dynamics. *COMAP Inc., UMAP Journal, VI*(3), 1–23.

Hofer, E. (2008). How to account for uncertainty due to measurement errors in an uncertainty analysis using Monte Carlo simulation. *Health Physics, 95*(3), 277–290.

Kloos, M. (2015). Main features of the tool SUSA 4.0 for uncertainty and sensitivity analyses. ESREL 2015, European Safety and Reliability Conference, Zurich.

Rosner, B. (1995). *Fundamentals of Biostatistics* (4th ed.). New York: Duxbury Press.

Schafer, D. W., & Gilbert, E. S. (2006). Some statistical implications of dose uncertainty in radiation dose-response analysis. *Radiation Research, 166*, 303–312.

Simon, S. L., Hoffman, F. O., & Hofer, E. (2015). The two-dimensional Monte Carlo: A new methodologic paradigm for dose reconstruction for epidemiological studies. *Radiation Research, 183*, 27–41.