Methods for uncertainty analysis

OPERA-PU-GRS7321

Radioactive substances and ionizing radiation are used in medicine, industry, agriculture, research, education and electricity production. This generates radioactive waste. In the Netherlands, this waste is collected, treated and stored by COVRA (Centrale Organisatie Voor Radioactief Afval). After interim storage for a period of at least 100 years radioactive waste is intended for disposal. There is a world-wide scientific and technical consensus that geological disposal represents the safest long-term option for radioactive waste.
Geological disposal is emplacement of radioactive waste in deep underground formations. The goal of geological disposal is long-term isolation of radioactive waste from our living environment in order to avoid exposure of future generations to ionising radiation from the waste. OPERA (OnderzoeksProgramma Eindberging Radioactief Afval) is the Dutch research programme on geological disposal of radioactive waste.

Within OPERA, researchers of different organisations in different areas of expertise will cooperate on the initial, conditional Safety Cases for the host rocks Boom Clay and Zechstein rock salt. As the radioactive waste disposal process in the Netherlands is at an early, conceptual phase and the previous research programme has ended more than a decade ago, in OPERA a first preliminary or initial safety case will be developed to structure the research necessary for the eventual development of a repository in the Netherlands. The safety case is conditional since only the long-term safety of a generic repository will be assessed. OPERA is financed by the Dutch Ministry of Economic Affairs, Agriculture and Innovation and the public limited liability company Electriciteits-Produktiemaatschappij Zuid-Nederland (EPZ) and coordinated by COVRA. Further details on OPERA and its outcomes can be accessed at www.covra.nl.

OPERA-PU-GRS7321
Title: methods for uncertainty analysis
Authors: D-A Becker, J.B. Grupa, J. Wolf
Date of publication: 13 December 2013
Keywords: Safety Case, Safety Assessment, Uncertainty Analysis

Summary

This report gives a general description of the methods for uncertainty analyses that can be used in the safety assessment performed in OPERA.

Samenvatting

Dit rapport geeft een algemeen overzicht van de methodes voor onzekerheidsanalyse die toegepast kunnen worden in de veiligheidsstudie die verricht wordt in het OPERA onderzoeksprogramma.

1. Introduction

1.1. Background

The five-year research programme for the geological disposal of radioactive waste – OPERA- started on 7 July 2011 with an open invitation for research proposals. In these proposals, research was proposed for the tasks described in the OPERA Research Plan [VER 11]. This report is a result of one of the research projects in OPERA, OPAP-I, started in June 2012, carried out by a consortium of NRG, SCK.CEN, GRS and TNO.

1.2. Objectives

This report is the first result of the research Task 7.3.2 *Methods for uncertainty analysis*. The work elaborates on useful and feasible state-of-the-art techniques for uncertainty analysis. Proper methods have been worked out and documented in a way that they can be integrated in the OPERA modelling environment developed in Task 7.2.4 *Integrated model for safety assessment*. Guidelines for Task 7.2.5 *Parameterization of PA models* have been proposed to identify combinations of extreme parameter values that may influence the overall PA outcome (sensitivity analysis). Also, some attention has been given to the process on how to organize, perform, integrate and communicate uncertainty calculations and their outcomes.

1.3. Realisation

The study presented in this report is performed by GRS with support from NRG. The study is based on the experience in previous projects, in particular VSG, PAMINA, EVEREST and OPLA-1A.

1.4. Explanation contents

Chapter 2 gives a general discussion of uncertainties in the safety Case. Chapter 3 describes the concepts for dealing with uncertainties. Chapter 4 gives the mathematical concepts for probabilistic uncertainty and sensitivty analysis. Chapter 5 gives a brief discussion of the methodology from the viewpoint of the overall safety assessment. Chapter 6 concludes the report by listing the steps that are taken in the uncertainty analysis foreseen in OPERA.

2. Uncertainties in the Safety Case

The goal of a safety case for a final repository project is to prove that the facility will be safe in every respect. This comprises considerations about the near and far future. It is a principal fact, however, that statements about the future can never be more than likelihood statements. Although, by this reason, a strong proof of safety is principally impossible, the remaining uncertainty can be assessed and should be kept as small as possible. This has to be done by carefully identifying and quantifying the primary uncertainties that can have an influence on the overall uncertainty of the safety statement and properly assessing this influence.

An overview of the treatment of uncertainties in the disposal programmes of several European countries has been compiled within the PAMINA project [PAM 08, PAM 11].

The OPERA safety case is based on the IAEA guide [IAEA 12], which requires that the safety case provide, among other things, i*dentification of uncertainties in the behaviour and performance of the disposal system, analysis of the significance of the uncertainties, and identification of approaches for the management of significant uncertainties*.

This chapter gives an introduction to the problem and explanations of the relevant terms.

2.1. Aleatory and epistemic uncertainties

Generally, uncertainties in describing the physical reality of some system can be classified in two categories:

- Uncertainties that are due to the physics of the system under consideration are called *aleatory*.
- Uncertainties that are due to the lack of knowledge about the system under consideration are called *epistemic*.

Aleatory uncertainties are an intrinsic property of the physical system. They result from effects that cause a principally unpredictable behaviour. Although such effects are ubiquitous in physics on a microscopic level, they are usually of limited importance for the macroscopic description. While, for example, one is unable to predict the exact travel path and time of a specific particle in a porous medium, the macroscopic transport can be well described using integrative measures like concentration, advection, diffusion and dispersion. Nevertheless, aleatory uncertainties occur also on the macroscopic level, either because it is practically impossible to collect sufficiently detailed information about the system to allow a less uncertain prediction of its behaviour, or because insufficiently predictable future events like ice ages might have a strong influence to some important element of the system. Aleatory uncertainties cannot be reduced, and in uncertainty management the main task about them is to identify and quantify them properly and to assess their influence.

Epistemic uncertainties, on the other hand, have nothing to do with the physics of the system but with our incomplete knowledge about it. Possibly, not all relevant effects are completely understood and described properly by the applied models. Even if the system and the influencing effects are well understood, however, there are usually some parameters we do not know much about, because they have not (yet) been or cannot be measured with sufficient accuracy. In practice, uncertainties of the epistemic type are usually more relevant than the aleatory ones. They can be reduced by focusing research efforts to dedicated investigations, which is why a reliable sensitivity analysis is important for them.

While in principle, aleatory and epistemic uncertainties are two different things, the distinction between them is often not unique in practice. The hydraulic conductivity of the relevant radionuclide travel path, for example, will depend on the path itself, which may be subject to aleatory uncertainty, as well as on poorly known hydraulic properties of the geological layers. The uncertainty of container lifetime can be regarded as aleatory, but it can be reduced by focusing research to improving the containers, which requires this uncertainty to be handled as epistemic. In fact, many, if not most uncertainties have some characteristics of both types. Since purely aleatory uncertainties are rare in a safety case and epistemic uncertainties generally require a more sophisticated handling, it is a possible approach to simply handle all uncertainties as if they were epistemic.

2.2. Sources of uncertainties

Generally, three main sources of uncertainties in the safety case are distinguished:
- scenario uncertainties,
- model uncertainties,
- parameter uncertainties.

2.2.1. Scenario uncertainties

To assess the safety of a deep underground repository (DGR) using a contaminant transport model, it is necessary to define the principle system development, the scenario, before modelling the details. Scenario analysis is an important task within the safety case. It is done by identifying and assessing the relevant features, events and processes (FEPs) and combining them to a number of scenarios that are considered more or less likely to occur. This procedure is subject to a variety of uncertainties, concerning, for example, the relevance of features or processes or the probability and consequences of future events, and resulting from the uncertainty of information about the different components of scenario development:
- initial conditions,
- internal FEPs and couplings between them,
- external FEPs,
- time scales.

Consequently, there will be considerable uncertainty about the relevance of the derived scenarios. A characteristic property of such scenario uncertainties is that in most cases they are hard to quantify.

2.2.2. Model uncertainties

As the assessment of the effects of a DGR to the environment in the far future can only be done by numerical calculations, it is necessary to describe the relevant processes defined in the identified scenarios using adequate mathematical and numerical models. Reliable models for simple physical effects like diffusion or advection are available, but in reality, many effects occur together and influence each other, requiring complex coupled models. In the development of such models, different sources of uncertainty occur:
- poor or incomplete knowledge or understanding of physical processes,
- incomplete understanding of interactions and mutual influences,
- simplified representation of the system,
- simplified mathematical representation of individual processes and their interactions
- errors in the numerical codes,
- human errors in executing the calculations.

It is therefore always questionable to what extent a model describes the real conditions correctly. A proper validation of models is often practically or actually impossible. Therefore, the applied models are an important source of uncertainty in the safety case, which people tend to ignore once they have gained some trust in their models by calculating results that look plausible enough. It should always be kept in mind that numerical results are not by themselves true descriptions of the nature.

2.2.3. Parameter uncertainties

The numerical models used for calculating future impacts to the environment have to be supplied with input data. Even if the model itself is adequate for providing valid safety statements about the system, this requires proper selection of the values of all model parameters, most of which are, by various reasons, more or less uncertain. Each model parameter corresponds with some physical property of the real system, but this is, in general, not a simple 1:1 relationship. As well as it is possible that several system properties are represented in the model by one common parameter, different parameters are sometimes influenced by the same system property. Therefore, the uncertainties of the parameters can be coupled in a complex manner. Generally, the quantification of parameter uncertainties and their dependencies is a challenging task. Parameter uncertainties are not necessarily the most important source of uncertainties in the safety case, but once quantified, they can be handled using well-established mathematical procedures.

Figure 2-1 gives a schematic impression of the sources of uncertainties on different levels and their contribution to overall uncertainty. Predicting the real development correctly means finding the right way through the tree to the one red balloon that represents the reality best.

2.3. Identification and quantification of uncertainties

To be handled properly, all potentially relevant uncertainties in the safety case have to be identified and quantified in an adequate manner. A reliable statement about the overall uncertainty of the safety assessment can only be given if all relevant uncertainties of scenarios, models or parameters are taken into account and quantified in such a way that they actually represent our lack of knowledge about the reality. Realistic quantification of uncertainties is a task of its own in a safety case, and not at all an easy one. If uncertainties are underestimated, the resulting safety statement might appear more reliable than it really is. Strong overestimation of uncertainties, however, can lead to very uncertain and therefore nearly meaningless results.

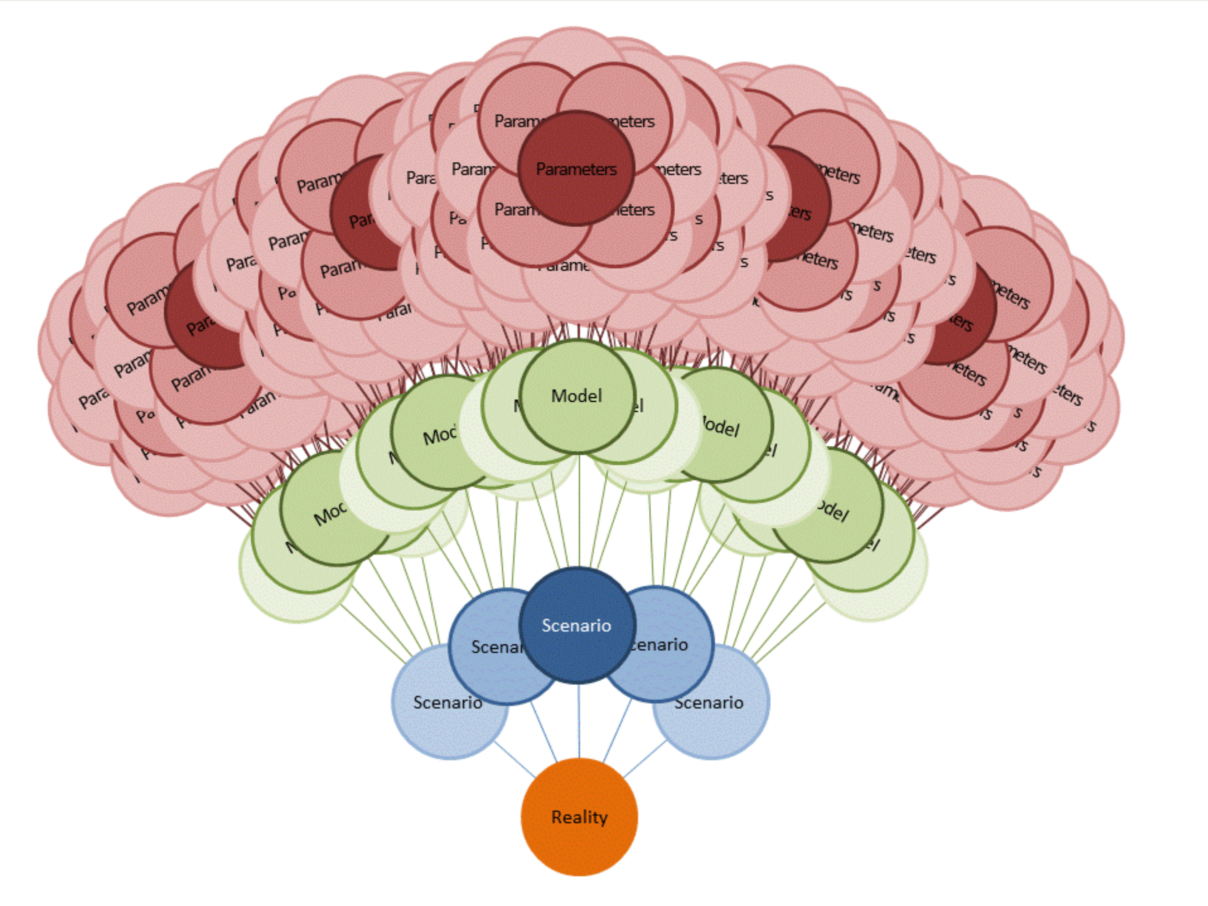

Figure 2-1 Levels of uncertainty

2.3.1. Scenario uncertainties

Scenario uncertainties result from the uncertainties of the various FEPs the scenarios are composed of. These uncertainties themselves can be of the aleatory or the epistemic type and are often hard to quantify objectively. Moreover, even if the uncertainty of a FEP is known, it can be unclear how it affects the probability of the scenario to occur. The assessment of scenario probabilities by expert judgement, taking into account the probabilities of the relevant FEPs, is the core task of scenario analysis. More than a rough categorisation, however, is hardly possible.

2.3.2. Model uncertainties

Proper quantification of model uncertainties is even harder, if not practically impossible. We use the models we regard as valid, normally without being able to specify a probability that they are really valid for all computational situations under consideration. The trust in a model may increase if it is properly validated under various conditions, but especially for the typical complex coupled models applied in performance assessment a real validation is at least expensive and time-consuming, if not practically impossible. Moreover, even if a model can be validated under certain conditions, this does not provide a quantitative measure for the model uncertainty.

If different physical situations can occur that are described by different model options, a quantification of their probabilities by expert judgement can be possible. Such model uncertainties can be mapped to and treated as parameter uncertainties

2.3.3. Parameter uncertainties

Compared with scenario and model uncertainties, the quantification of parameter uncertainties looks like a straightforward task on the first sight, but it is nevertheless demanding. Since each parameter is measured on a one-dimensional scale, its uncertainty can be described by a probability density function (pdf). The integral of the pdf over an interval of the parameter axis gives the probability that the actual parameter value lies within this interval. There are a number of types of pdfs, like uniform, normal or exponential distributions, which require different statistical parameters for specification. So the quantification of the uncertainty of a physical parameter means finding the right distribution type as well as the correct statistical parameters. If there is enough experience about the parameter in question, maybe from measurements under different conditions, it should be relatively simple to find the adequate pdf. In many cases, however, there is very little knowledge about the values the parameter can take on. This means that the uncertainty of the parameter is rather high, but a concrete pdf is hard to specify. The pdfs used should always be substantiated and consistent with the actual knowledge about the parameters (see chapter 3.5).

2.4. Uncertainties in the context of Performance Assessment

Performance Assessment (PA) is that part of the safety case that deals with assessment of the long-term safety of the DGR system by numerical model calculations. Long-term here means a principally unlimited time frame, because safety cannot be confined to a limited future. Nevertheless, there is a common consensus that an assessment period of 1 million years is normally sufficient, since the radiological hazard potential of the typical inventory of a repository for high-level waste (HLW) will decay to a level that is comparable to or below natural values within this time.

It is obvious that a prediction over 1 million years will be subject to an uncertainty that can hardly be assessed. Therefore, the claim of Performance Assessment is not to predict the future development of the system correctly, but to provide a statement about its safety. If, for example, the uncertainty of a parameter affects some details of the system behaviour without influencing the safety of the system as a whole, this uncertainty is irrelevant for the safety statement. The applied models do not have to be realistic descriptions of the future development of the system, but should make sure, as far as possible, that they do not underestimate detrimental effects to man and environment. The uncertainty of the model calculation results should therefore not be misinterpreted as the uncertainty of the safety assessment.

Nevertheless, as a precondition for estimating the uncertainty of the safety statement reliably, it is necessary to investigate the effects of all uncertainties to the model output thoroughly. This is done in three phases:

- identification and quantification of all uncertainties that might have a significant influence to the model results,
- uncertainty analysis: assessment of the overall uncertainty of the model results,
- sensitivity analysis: identification of those uncertainties that are actually most relevant for the model output uncertainty.

For all of these tasks, different approaches, concepts and methods are available, which are described in the following chapters.

3. Concepts for dealing with uncertainties

Figure 3-1 is taken from [IAEA 12] and shows the components of a safety case and their interactions. Component F, management of uncertainty, is an essential part, which affects all other major components. For explanation of this component, the mentioned IAEA guide refers to the IAEA General Safety Requirements, part 4 [IAEA 09], which says:

Uncertainty and sensitivity analysis shall be performed and taken into account in the results of the safety analysis and the conclusions drawn from it.

4.58. […].There will always be uncertainties […] that will depend on the nature of the facility or activity and the complexity of the safety analysis. These uncertainties have to be taken into account in the results of the safety analysis and the conclusions drawn from it.

4.59. Uncertainties in the safety analysis have to be characterized with respect to their source, nature and degree, using quantitative methods, professional judgement or both. Uncertainties that may have implications for the outcome of the safety analysis and for decisions made on that basis are to be addressed in uncertainty and sensitivity analyses. Uncertainty analysis refers mainly to the statistical combination and propagation of uncertainties in data, whereas sensitivity analysis refers to the sensitivity of results to major assumptions about parameters, scenarios or modelling.

As a consequence from these requirements, practicable concepts have to be defined for characterizing and handling of all kinds of relevant uncertainties. According to Figure 3-1, management of uncertainties is relevant for the following components of the safety case:

- C – System description,
- D – Safety assessment,
- G – Limits, controls and conditions,
- H – Integration of safety arguments.

Since this report is written from the view of Performance Assessment, it does not deal with the uncertainties of components G and H, but only with those affecting components C and D. However, we do not classify the uncertainties according to the components but distinguish between the three kinds of uncertainties introduced in chapter 2, scenario uncertainties, model uncertainties and parameter uncertainties.

In the following, the principal concepts for identifying and handling uncertainties in the safety case are introduced and explained. In performance assessment, a combination of all of them should be applied.

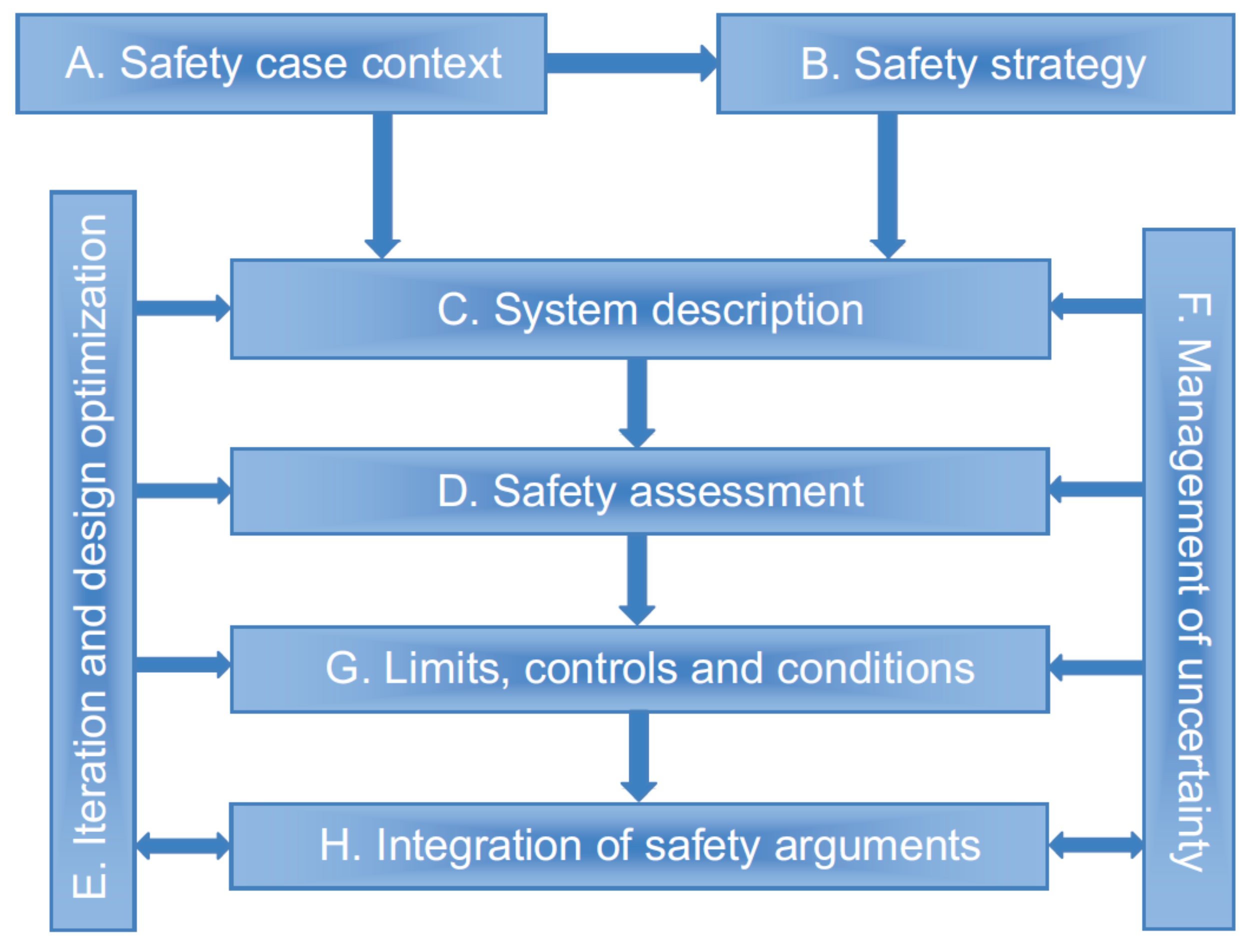

Figure 3-1 Components of the safety case [Source: IAEA SSG-23]

3.1. Scenario analysis

The site and the repository system will undergo exactly one evolution, which will be governed both by climatic and geological processes at the site and processes induced by the repository construction and the emplacement of the waste. This real evolution cannot be predicted in all details. The resulting uncertainty with regard to the future evolution of the repository system can be reduced only to a limited extent by additional research and site investigations. For example, it can be assumed that several cold times with permafrost formation will occur in Northern Europe within the next one million years, which may be associated with glaciation of the site. An exact prediction, how many of these cold times will occur is not possible.

Therefore, on the basis of a systematic assessment of relevant influencing factors, a limited number of possible evolutions have to be derived with the objective to identify and describe in detail the most relevant scenarios.

A systematic scenario analysis methodology was developed in the project VSG [BEU 12]. It aims at deriving one reference scenario and a number of alternative scenarios. At large, the scenarios are supposed to represent the reasonable range of repository system evolutions to cover the scenario uncertainties. The methodology allows the assignment of probability classes to the scenarios pursuant to the German regulatory framework [BMU 10]. This methodology is described in the following as an example for how the uncertainties in scenario development can be handled systematically.

3.1.1. FEP catalogue

The individual scenarios are characterised by FEPs that may influence the future evolution of the repository system and their associated characteristics. The relevant information is given in a site-specific FEP catalogue, providing detailed information for each FEP. The catalogue allows selecting directly all FEPs that are relevant for the reference scenario and the alternative scenarios.

Each FEP entry in the catalogue comprises general information, a description of the circumstances at the site and site-specific impacts, a classification of the conditional probability of occurrence, details on the impairment of the initial barriers' functionality, and information regarding the time frame of action. The direct interdependence with other FEPs must be specified and explained, thereby distinguishing between initiating FEPs, resulting FEPs, affecting FEPs and affected FEPs, respectively. If possible, probable and less probable characteristics of the FEP are indicated. Sometimes it is only possible to describe a characteristic but no probabilities can be attributed to it. This may be due to scarce data or information, or to a situation, where only bounding values are of interest with respect to the scenario analyses. In those cases representative characteristics are described.

3.1.2. Scenario development

Scenario outlines for a disposal in clay have already been developed since the earliest safety assessments for geological disposal. E.g. in the EC PAGIS study of 1988, normal evolution scenarios and two altered evolution scenarios (climatic changes and faulting) were identified for two reference sites; in Boom clay and in Oxford clay. Since then the list of scenarios has been growing in the various national and international programmes.

Since the OPERA outline of a disposal concept in clay (OPERA-PG-COV008) is a generic design it is not possible to identify site specific scenarios, so the use of public available scenario descriptions is expected to be adequate. Nevertheless, a FEP screening process has been undertaken in order to identify potential additional alternative scenarios. This screening method is typically a 'top-down' method for developing scenarios, as descried in SSG-23. The method is based on analyses of how the safety functions of the disposal system may be affected by possible events and processes.

The screening procedure resulted in the confirmation of the initial set of nine scenarios, but also leads to the identification of six so-called "what-if cases": cases that need to be analysed in more detail the impact of specific FEPs on the safety functions. The procedure and results are further described in OPERA-PU-NRG011.

The following, typical "bottom-up" method has been applied by GRS in the VSG. The scenario development commences at two starting points that ensue directly from the guiding principles for deriving the safety concept:

- A number of initial barriers are identified that constitute a subset of all barriers acting in the repository system via diverse modes of operation and, partly, in different time frames. The initial barriers embrace (parts of) the host rock and (a subset of) the technical barrier system. FEPs that could impair the functionality of the initial barriers provide the first starting point for scenario development.
- In addition all possible system evolutions need to be considered which involve a release of radionuclides from the waste forms. Those FEPs which are related to the mobilisation of radionuclides and their transport are the second starting point for scenario development.

The reference scenario results from considering all probable FEPs that

- may impair the functionality of the initial barriers (initial FEP), and

- determine the mobilisation of radionuclides from the waste and their subsequent transport, both in the gas phase and in the liquid phase.

If appropriate information is available in the FEP catalogue, the probable or representative characteristics of these FEPs are taken as a basis. Otherwise, the characteristics result from the direct (first level) interaction with other FEPs. In this case, always the probable or representative characteristics of the controlling FEPs of the first level are assumed. If these FEP themselves are controlled directly by other FEPs (second level), their characteristics are used. Further levels are only included, if the aspects are not yet covered by FEPs in the first or second level. Owing to the method applied, the reference scenario is probable. The procedure to identify the characteristics of FEPs for the reference scenario is depicted in Figure 3-2.

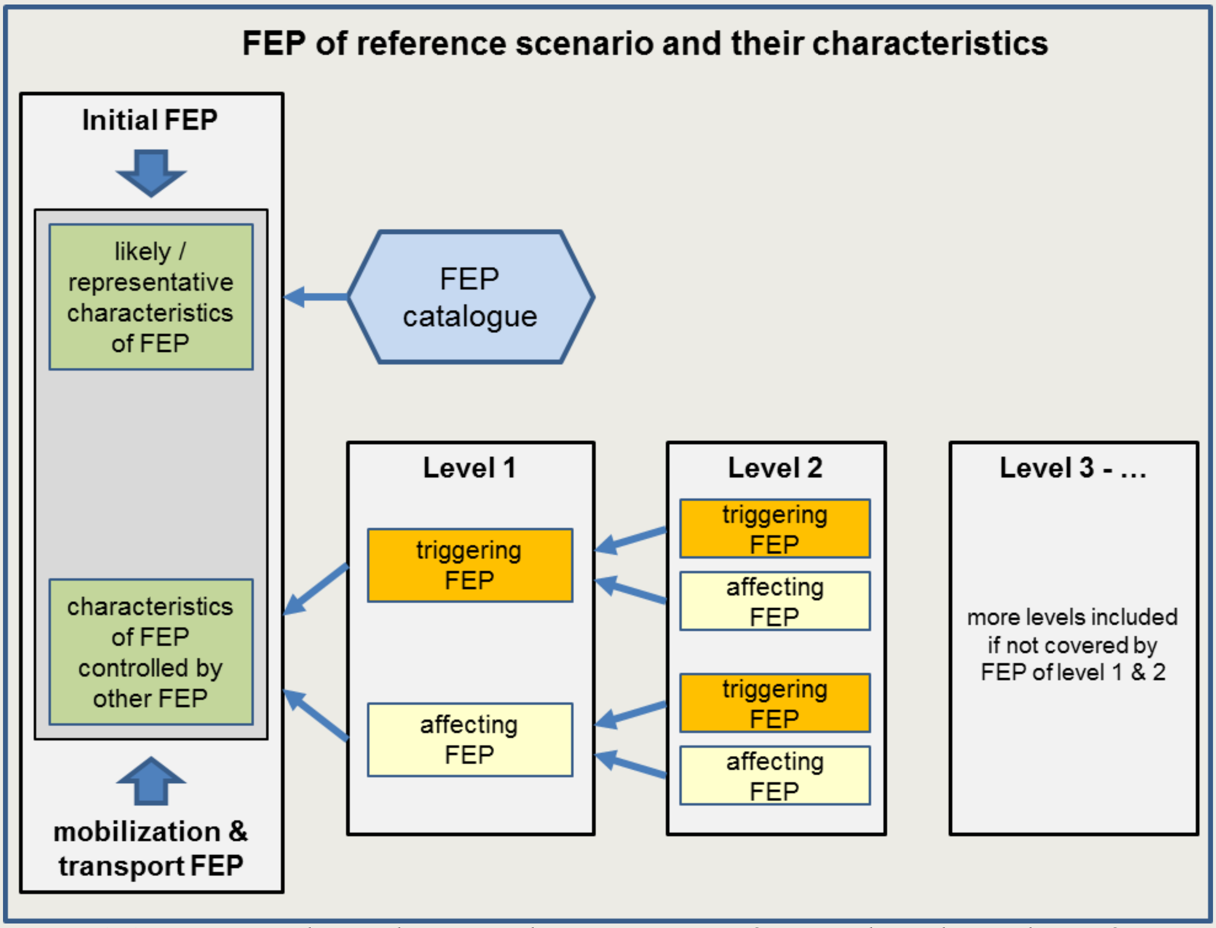

Figure 3-2 Approach to deriving characteristics of FEPs describing the reference scenario

Specific assumptions concerning the reference scenario are an important element of scenario development. They provide a means to deal in a transparent and traceable way with particular uncertainties, some of which may be minimised in the future while others may never be reduced at all. In particular, the latter pertains to the future climatic evolution. Therefore, a certain climatic evolution with a series of different types of cold times must be defined for the reference scenario. Other specific assumptions deal with situations, where no proof has been furnished yet with regard to producibility and functionality of engineered barriers or other technical components. Alternative specific assumptions constitute a starting point for deriving alternative scenarios.

The reference scenario embraces a set of probable evolutions of the repository system that is as large as possible. Alternative scenarios are evolutions which differ in exactly one aspect from the reference scenario. Alternative scenarios are developed from the following starting points:

- evolutions resulting from alternatives concerning the specific assumptions for the reference scenario,
- evolutions resulting from less probable characteristics of the FEPs that may impair the functionality of the initial barriers,
- evolutions resulting from less probable characteristics of the FEPs describing mobilisation and transport of radionuclides, and
- evolutions resulting from less probable FEPs.

If possible, information is directly taken from the FEP catalogue concerning less probable characteristics of FEPs that may impair the functionality of the initial barriers or that describe mobilisation and transport of radionuclides. Otherwise, the characteristics are directly controlled by other FEPs in a similar way as shown in Figure 3-2.

It is feasible that similar alternative evolutions result from the different starting points. In this case, various evolutions may be abstracted into one representative alternative scenario that covers the characteristics of the various evolutions.

3.2. Conservative approach

In each phase of developing scenarios, establishing models or collecting data one is confronted with uncertain knowledge about the system and its physical or chemical behaviour. Since, however, the final goal is safety of the system and often the direction of influence is obvious, one can eliminate or reduce a lot of uncertainties by choosing conservative options or values. That means that by fixing such a value or option to the unfavourable end of its range of uncertainty it is made sure that its detrimental influence that might jeopardise safety will not - or probably not – be underestimated.

Conservative assumptions, instead of realistic ones, are often applied in scenario analysis, model development and parameter determination. This is one of the reasons for PA results not being predictions of the real system evolution but safety-oriented estimations. It should be kept in mind, however, that conservatism has to be applied with care. Over-conservatism can lead to results that impair the safety statement or even exceed the permissible limits and make the whole performance assessment useless. Therefore, even if the direction of influence is clear, one should normally avoid selecting extreme values or options regardless of their probability, but try to find a reasonable compromise between conservatism and probability. The remaining uncertainty can then be handled by applying a probabilistic approach.

Many effects in the repository system are coupled in a complex way, which is often hard to understand and difficult to model. The uncertainties of a specific model or parameter can therefore have various impacts to the PA results and even cause opposing trends, depending, for example, on the values of other model parameters. In such cases, a conservative choice is principally impossible. There is some danger that this is overlooked in practice and an alleged conservatism is introduced to the PA model or the parameter set, which actually does not exist. This can distort the results. Before deciding to use a conservative model or value one should thoroughly analyse the situation, taking into account all possible factors that might have an influence. Conservative assumptions should only be applied where it is sufficiently clear that they shift the PA results to the unfavourable side under all possible circumstances.

Conservative assumptions are integrated in the model that is foreseen for further investigation and become a part of this system. By finally establishing the model and its parameterisation for each relevant scenario, a number of standard or reference calculation cases are defined. These do neither necessarily represent the most probable evolution of the system nor the worst case, but should be designed to allow safety-oriented statements of the kind that it is improbable that the detrimental impacts to the environment will be higher than calculated. Such statements are still subject to uncertainties, which have to be specified and addressed by adequate investigations, keeping in mind that these will yield statements about the uncertainty of the conservative system, not about the uncertainty of the reality.

3.3. Deterministic investigation

Both model and parameter uncertainties can be addressed by conducting deterministic investigations in performance assessment. This means that several distinct variations of the standard calculation cases are considered, which should cover, as far as possible, the range of uncertainty.

Deterministic investigations are a good means for studying the system behaviour under the influence of uncertain models or parameters in detail, but do not provide a quantitative estimation of the effects of uncertainties. Parameter intervals can be explored by using the extreme values, or different model options can be investigated, regardless of their actual probabilities. The benefit of such studies is twofold:

- They provide a deeper insight to the behaviour of the model. This can improve the general understanding of the processes and their interactions, but also reveal modelling errors by showing implausible results.
- They give an impression of the model output under extreme conditions, which are normally considered improbable or even not expected to occur at all. Investigations with parameters or options that lie outside the expected bandwidth are called *what-if-studies*. Such studies can be useful for showing that the system will not violate the safety requirements even under assumptions that are unreasonably pessimistic.

A deterministic investigation typically addresses only one parameter or model option, based on one of the standard calculation cases. Varying several parameters or options at a time could lead to blurred results and impair the gain of insight. The effects of different parameters and options to the calculation results, however, are not independent of each other. Simply setting all parameters and options to that value that appears most unfavourable in a deterministic investigation will normally not lead to the global “worst case”. This can only be achieved by varying all parameters together.

3.4. Probabilistic analysis

While deterministic investigations are worthwhile for improving the understanding of the qualitative behaviour of the model, for a quantitative assessment of the uncertainty of the model output and its sensitivity against parameter variations, taking into account the parameter distributions and option probabilities, a probabilistic approach is necessary. The general idea of probabilistic uncertainty and sensitivity analysis is to execute a high number of model runs with sets of parameter values that are distributed over the parameter space according to their actual distributions.

A probabilistic approach can only handle quantifiable parameters and is therefore most adequate for parameter uncertainties. Model uncertainties can be taken into account by mapping them to artificial parameters that switch between different model options according to their values. In principle, also scenario uncertainties could be handled this way, but this could lead to hardly understandable results, and so they are normally kept out of the probabilistic analysis. That means that each relevant scenario has to be investigated separately, including a probabilistic analysis, if considered necessary.

A probabilistic investigation requires the availability of reliable probability density functions (pdf) for all parameters. For achieving results that reflect the actual uncertainty of the model calculations the pdfs should be consistent with the knowledge about the parameters. A proper quantification of uncertainties therefore requires evaluation of available sources and an objective scheme for transforming knowledge into a mathematical function, the pdf.

Once the pdfs for all uncertain parameters are established, they can be used for drawing a sample of parameter sets, which can be used as a basis for the probabilistic investigations. For each parameter set a model run is performed. The results of all model runs are then evaluated using specific methods.

There are two principally different tasks of probabilistic analysis. While uncertainty analysis means assessing the overall uncertainty of the model output under the influence of the uncertainties of all considered parameters, sensitivity analysis is the investigation of how the different input uncertainties affect the uncertainty of the model output. Sensitivity analysis is the mathematically more demanding task, but it is an important part of the safety case. A reliable sensitivity analysis provides good insight to the mathematical behaviour of the model, can trigger research activities by focusing them to the most influential parameters, and can identify model errors by detecting implausible sensitivities.

Probabilistic analysis is generally considered the most powerful tool for assessing uncertainty and sensitivity of the model and should be executed with care.

3.5. Determination of pdfs

A proper quantification of uncertainties in the form of probability density functions (pdfs) is an essential part of the uncertainty management and a pre-requisite for probabilistic uncertainty and sensitivity analysis. It should be avoided simply to use standard pdf types like normal or uniform distribution with more or less arbitrarily chosen distribution parameters without thinking about the actual knowledge about the parameter in question. A review of methods for assigning pdfs by expert judgement has been made in PAMINA [PAM 09a]. For practical purposes, a simple systematic procedure for pdf development has also been proposed [PAM 08b]. The following is a very short summary of this procedure.

Firstly, available sources of information about the parameter have to be identified and assessed according to their quality. For levels of quality have been defined:

- direct measurements (3),
- model representation (2),
- analogy consideration (1),
- plausibility (0).

If several independent sets of information are available, they should be merged under assessment of their reliability. Data sets that seem unreliable should be excluded from the investigation. Lower-quality data should be used to confirm and substantiate higher-quality data. It is recommended to merge data from the two highest available quality levels. By adding their levels one gets a number between 0 and 5. The further procedure depends on this number.

Case 0: Only a plausibility interval is known. No parts of the interval should be weighted higher than others, and consequently, the PDF must be a uniform distribution between the interval bounds.

Case 1: Some analogies are available, supported by plausibility limits. In this case the pdf has to be transferred from the analogy data or model. The resulting pdf has then to be re-calibrated to make sure that the plausibility limits are not exceeded.

Case 2: A model representation for the parameter under consideration exists. A model will always allow the derivation of a pdf. For calibration of the pdf only the plausibility limits can be used.

Case 3: A model representation for the parameter under consideration exists. The pdf has to be derived from the model. Its calibration can be based on analogies instead of simple plausibility. One can try to transfer the model so that it describes, as well as possible, the analogous situation and then re-transfer the pdf to the actual situation.

Case 4: Though it seems to be a comfortable situation to be in possession of directly measured data supported by analogues, this case can become problematic and include considerable subjectivity. The pdf has to be derived from the measured data, which can be impossible if only a few values or even a single one are available. First, it has to be decided whether the amount of data is sufficient to derive a pdf, which then can be compared with the analogue. If there is no strong discrepancy, the pdf is confirmed, otherwise uncertainty might be higher than the data suggest and the pdf needs to be recalibrated. If, however, the directly measured data do not suffice to derive a pdf one should try to use the analogue for derivation of a pdf, transfer it to the actual parameter and calibrate it with the measured data. If even that is not possible, one should take a uniform distribution in an adequate interval. If no detailed background is available to perform a qualified expert judgement, one will have to derive a sufficiently large interval from the analogue. A triangular distribution is probably the best choice in this case.

Case 5: Measured data are supported by a model representation. In this case the pdf can be derived from the model and calibrated with the measured data. If the model and the measurements obviously do not fit together and it cannot be decided by expert judgment which is more reliable, one should generally prefer the measured data. Disregarding the less reliable source of information leads to a lower case, and the PDF determination should be performed accordingly.

In some of the cases the pdf has to be determined from data points. If there are plenty of points, the pdf can be read off directly. If the number of data points is low (4 or less), there is very little information and it is recommended to identify a plausibility interval and to use either a uniform distribution or a triangular distribution with its peak at the mean of the given data. The situation that requires the most sophisticated procedure results from a medium number of data points. In this case, only (log-)uniform or (log-)normal distributions should be considered, as there is too little information to choose a more sophisticated distribution type, and too much for relying on plausibility assumptions, which could justify a triangular distribution. It should first be decided whether a linear or a logarithmic scale is to be used. The following criterion can be applied: if the median of the data is closer to the geometric than to the arithmetic mean, a logarithmic-scale pdf should be used. As a null-hypothesis, a (log-)uniform distribution should be assumed and parameterized adequately. Only if this hypothesis has to be rejected, which can be tested, e.g., by the Kolmogorov-Smirnov-test [CON 99], a (log-)uniform distribution should be tried.

4. Probabilistic methods for uncertainty analysis

This chapter is dedicated to the description of the technical and mathematical methods for probabilistic uncertainty and sensitivity analysis. It is assumed that a numerical model is available and all input uncertainties to be taken into account can be represented by statistically distributed parameters. It is further assumed that a pdf is given for each of the parameters under consideration.

Mathematically, it has to be distinguished between random variables (normally denoted by capital letters) with their characteristics like expectation or variance and their realisations in the form of n-tuples of values (denoted by indexed lower-case letters). Since the explanations in this chapter are meant as an overview for practical application, we only use the latter concept here and show some formulas. The letter n is principally used for the sample size, the letter k for the number of parameters. For an exact mathematical formulation of the underlying theory and further details we refer to the pertinent literature [CON 99, YAT 03]. A practical introduction to the subject is given in [ROC 08].

4.1. Sampling

Generating a sample of parameter values is the first step in a probabilistic analysis. A sample of size n is a collection of n sets of values to be used for the calculation runs, taking account of the parameter distributions. Different more or less sophisticated sampling techniques have been developed, which lead to different kinds of samples with differently homogeneous coverage of the parameter space and differently pronounced random elements. It cannot be generally said which one is the best; this depends on the problem and the intended evaluation. Well-established sampling schemes for use in uncertainty and sensitivity analysis studies are:

- simple random sampling,
- stratified random sampling (e. g. Latin Hypercube sampling, LHS),
- quasi-Monte-Carlo sampling (low discrepancy sequences),
- specific sampling (needed for specific sensitivity analysis methods).

These are shortly described in the following.

4.1.1. Simple random sampling

Simple random sampling means that each individual set of parameter values is selected randomly and independently of all others. In practice, true random numbers are hardly available, and a software pseudo-random number generator will be used. Such generators need a start value, called seed. As long as the same seed is used, the generator will always yield the same sequence of pseudo-random numbers and with that the same sample. If different samples for the same problem are needed, different seeds have to be used.

The main characteristic of random sampling is that there is neither a system nor a memory in the sampling procedure. Each point is drawn without knowledge of the positions of the other points. Although, for some purposes like a pure uncertainty analysis this is exactly what is needed, it typically leads to clusters and holes in the parameter space. Simple random sampling is therefore usually not a very good choice if a homogeneous coverage of the parameter space is required.

4.1.2. Stratified random sampling, LHS

For sensitivity analysis, a good sample should cover the total parameter space as well as possible, so that all possible parameter combinations are explored according to their probability of occurrence. A better coverage can be reached by applying a stratified sampling scheme. This means that the ranges of the parameter values are sub-divided into separate strata from which appropriate numbers of values are randomly drawn. By making sure that each stratum is considered in the sampling a more homogeneous coverage is reached.

There are different techniques of strata definition and sampling. A well-known, often applied and approved technique is Latin Hypercube sampling (LHS). For each parameter the interval of possible values is divided into *n* non-overlapping and adjacent strata of equal probability:

$$\int_{a_i}^{b_i} f(x)dx = \frac{1}{n},$$

where f is the pdf and a_i and $b_i = a_{i+1}$ denote the lower and upper limits of stratum i. Then one value is drawn randomly from each stratum. Finally the drawn values of all parameters are combined by randomly permuting the stratum numbers for each parameter. This procedure provides a sample that is better adequate for sensitivity analysis but still shows a random structure.

4.1.3. Quasi-Monte-Carlo sampling

If the requirement of randomness is abandoned one can essentially improve the homogeneity of coverage of the parameter space by using low-discrepancy sequences instead of (pseudo-)random sequences. Discrepancy is a measure for the deviation of a sequence from the ideal equidistribution. Low-discrepancy or quasi-random sequences are constructed specifically to minimize discrepancy. These sequences are calculated using deterministic algorithms and leave only little freedom or room for randomness. Sobol sequences are a well-established type of low-discrepancy sequences [SOB 67]. A parameter sample based on a low-discrepancy sequence is called a quasi-Monte-Carlo sample. Such samples can be advantageous for sensitivity analysis, but one should keep in mind that they are neither random nor pseudo-random and therefore inadequate for investigations that explicitly require randomness.

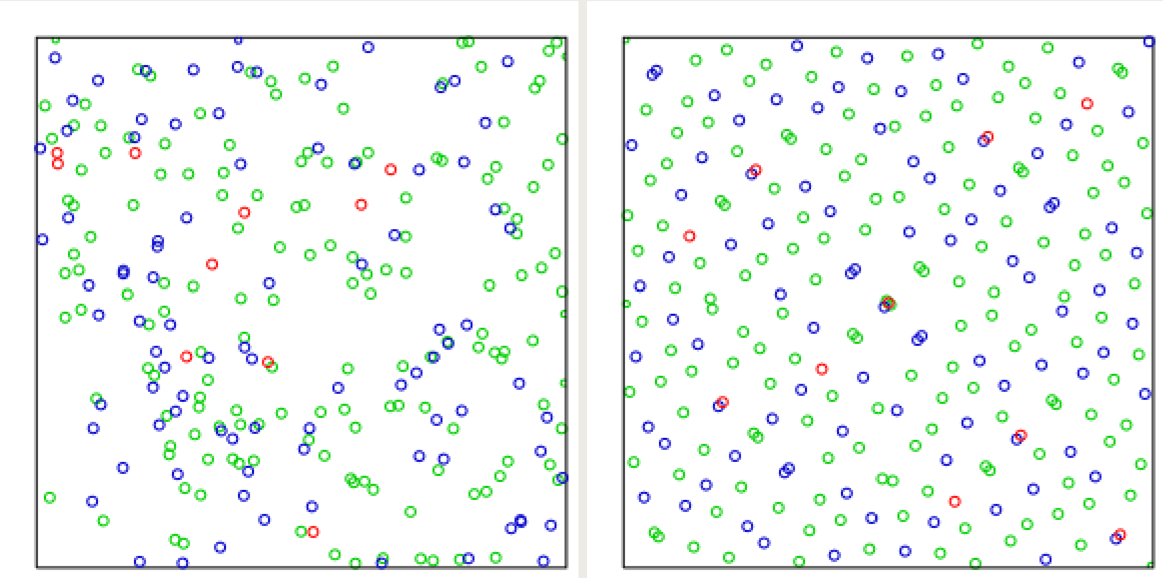

Figure 4-1 Comparison of random (left) and Sobol sequences (right). 256 points were drawn (red=1,..,10, blue=11,..,100, green=101,..,256). Points from Sobol sequence are more evenly distributed. [Source: Wikipedia]

4.1.4. Specific sampling

Some sensitivity analysis methods need specific sampling schemes and even an ordered sequence of evaluation. This typically leads to a rather inhomogeneous coverage of the parameter space. Moreover, such specific samples often have the drawback of non-reusability, which means that if an evaluation with a different sample size is to be performed, a totally new sample has to be drawn and all model runs with the old sample are useless for the new evaluation. Therefore, such sampling should only be performed if there is a good reason for applying the method that requires it.

4.2. Parameter dependencies

The input parameters of the model are not necessarily independent of each other. It has to be distinguished between strong and statistical dependencies. Strong dependency means that two parameters are coupled via a unique mathematical relation, which can be due to specific system properties that are not considered in the model itself. Such dependencies can always be resolved by adequate technical measures, so that the dependent parameters appear as a single one.

Statistical dependencies, however, require adequate handling in the probabilistic analysis. Such dependencies occur if two parameters are influenced by the same uncertain effect, but are additionally subject to their own uncertainties, so that they show a common tendency. A low porosity of a geological layer, for example, normally means that also the permeability is low, but since there is no universally valid porosity-permeability-relation, this is not a strong dependency. Such dependencies are often due to our inability to describe the natural interactions in detail.

Although theoretically, statistical dependencies can follow any mathematical relation, they are normally expressed in practice by linear correlation. For two statistically distributed parameters X and Y with the realisations $(x_1,\dots,x_n)$ and $(y_1,\dots,y_n)$ the empirical correlation coefficient is defined as

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}},$$

where the bar denotes the mean: $\overline{x} = \sum_{i=1}^{n} x_i$. The correlation coefficient is always between -1 and 1. Positive values mean that higher *x*–values mostly appear together with higher *y*-values, negative or inverse correlation means that higher *x*-values are correlated with lower *y*-values. An absolute value of 1 means that there is a strong linear or inversely linear relationship between *x* and *y*. If the correlation coefficient is 0 the variables are called uncorrelated, which does, however, not mean that they are necessarily statistically independent of each other. Even if the parameters are theoretically uncorrelated, there is always a random correlation, so that a correlation coefficient of exact zero will hardly ever be found.

The usual problem is not to calculate the correlation coefficient of two given parameter tuples, but the opposite one: we know or assume that two parameters are statistically correlated in reality and have to take account of that in the sampling procedure. As long as the sequence order is irrelevant for the evaluation, this can be achieved by drawing the parameter values independently according to their respective pdfs and then re-sorting and combining them appropriately. Adequate algorithms are available for introducing a given correlation matrix into a parameter sample.

In practice, it will hardly ever be possible to give a substantiated justification for a specific value of the correlation coefficient of two parameters we consider statistically coupled. It does not make much sense normally, to quantify such a dependency with a correlation coefficient below 0.5, since such a correlation will hardly be visible and have only a minor influence to the results of the probabilistic analysis.

4.3. Uncertainty analysis

Uncertainty analysis is the investigation of the overall uncertainty of the model output under the influences of all input uncertainties. Once all parameter uncertainties have been quantified properly by appropriate pdfs, possible correlations have been defined and an appropriate sample has been drawn, probabilistic uncertainty analysis is a straightforward task. After all *n* model runs have been performed, a set of *n* output values is available, which allows all kinds of statistical evaluation. If the model calculates a time-dependent output, one can either select a specific point in time or take the absolute maximum of each run.

Principally, for uncertainty analysis only the output of the model is relevant, since it already reflects the effects of all input uncertainties, regardless of the number of uncertain parameters and how they interact. For an uncertainty analysis there is no need to adapt the sample size and the number of runs to the number of parameters.

4.3.1. Statistical measures

Numerical information about the uncertainty of the model output can be obtained by calculating statistical measures like

- mean,
- standard deviation,
- maximum and minimum,
- median,

- quantiles (e. g. 5 % and 95 %).

All of these are well-known measures that can be calculated using standard software tools and do not need more explanation. Such values are helpful for quantifying the overall uncertainty, but do not give deeper insight to the model behaviour.

4.3.2. Graphical uncertainty analysis

A graphical representation is often more meaningful than some calculated numeric values. Scatterplots allow a two-dimensional visualisation of the model output values versus some other characteristic. DGR performance assessment models normally yield a time-dependent output for each run. If one intends to evaluate the maxima of all runs, the time component gets lost by simply calculating statistical measures. In a scatterplot, however, it can be made visible by using the second axis. A very descriptive presentation results from plotting the maximum value versus the time of its occurrence. Additional information, for example the radionuclide responsible for the maximum, can be colour-coded in such a plot. An example is shown in Figure 4-2.

Another type of plot that often gives interesting information about the model output is that of the relative frequency of output values. It can be plotted either as a column diagram, the height of each column representing the frequency of values that lie in a specific interval, or as a curve of cumulated frequency, which is normally presented in its inverse form. These curves indicate which percentage of the output values lie above the value on the x-axis. Examples for both types of presentations are shown in Figure 4-3.

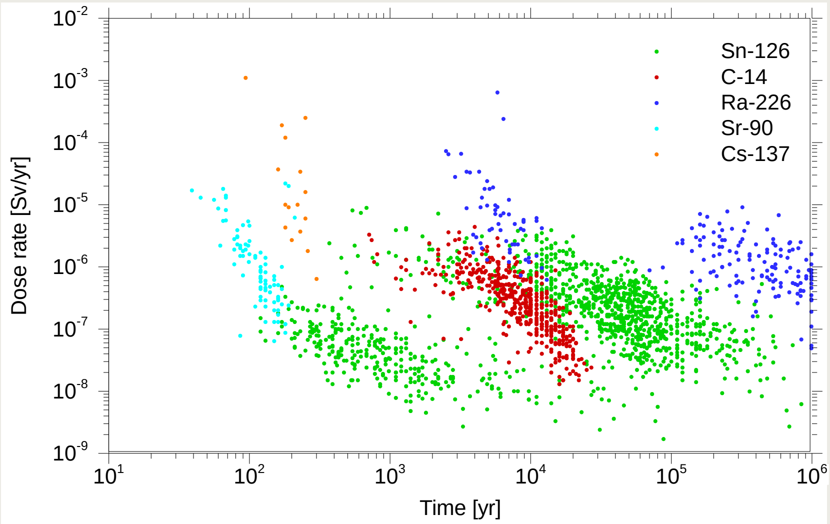

Figure 4-2 Scatterplot of maxima vs. time of occurrence with colour-coding of radionuclides

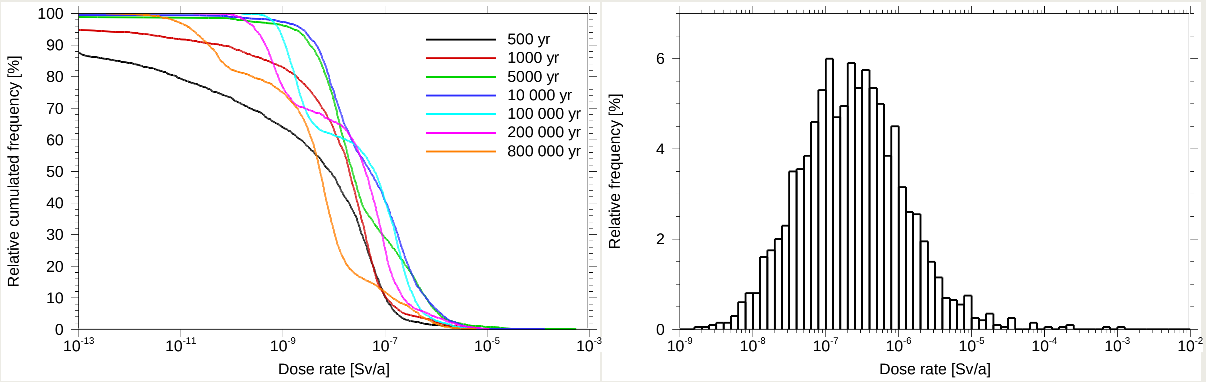

Figure 4-3 Relative frequency plots as cumulated curves for several points in time (left) and as a column plot for the maximum (right)

4.3.3. Compliance with criteria

For the detrimental effects of a DGR there are in most countries formal criteria requiring that some specific limit must not be exceeded. This can, of course, only be proved by applying the model, taking into account all relevant uncertainties. A strong criterion could therefore cause two kinds of problems:

- Since the distribution of model output values can theoretically extend to infinity, with a finite number of model runs one can never be sure to have actually calculated the highest possible output. A proof of compliance with a strong criterion would therefore be theoretically impossible.
- The requirement that even in very improbable cases the criterion must be met could lead to the necessity of unreasonably expensive measures to avoid situations that most likely would never occur anyway.

Therefore, criteria are normally given in a statistical form that allows a low probability of exceeding the limit value. But even the requirement that, for example, 95 % of all possible model output values have to remain below the limit would still require an infinite number of model runs. For this reason, such criteria are formulated with two probability values, *p* and *q*:

The probability that the model yields an output value not exceeding the limit value must be at least *p* with a confidence of at least *q*. That means that there is a probability of *q* or higher that the model will yield values below the limit in at least a fraction of *p* of all possible cases.

For a number *n* of independent model runs, the confidence can be calculated as

$$q = 1 - F(k; n, 1 - p) \ ,$$

where *k* is the number of limit exceedances (*k* = 0 is required if no exceedance is allowed) and *F* denotes the cumulative binomial distribution

$$F(k; n, 1-p) = \sum_{i=0}^{k} \binom{n}{i} (1-p)^i p^{n-i} \ .$$

Such a *p*/*q*-criterion can easily be checked by evaluating a relatively low number of model runs, as long as the following points are made sure:

- The parameter sets for the model runs to evaluate have to be purely random (or pseudo-random) and independent of each other. Model runs based on a non-random sample must not be used.
- The model runs for evaluation have to be finally selected prior to knowing their results. It is inadmissible to replace a limit-exceeding run by another one afterwards.
- The results of all evaluated model runs remain below the limit value.

If actually an exceedance of the limit is found, it is not sufficient to run the model once more in order to reach the required number of non-exceeding runs. In such a case a considerably higher number of runs become necessary.

The minimum numbers of model runs that guarantee compliance with the criterion can be calculated using the formula given above. For no exceedances and for one exceedance they are listed in Table 4-1 for different values of p and q.

Table 4-1 Minimum numbers of model runs to fulfil the *p*/*q* criterion with no exceedances (black) or one exceedance (red)

	q = 90 %	q = 95 %	q = 99 %
p = 90 %	22 (38)	29 (46)	44 (64)
p = 95 %	45 (77)	59 (93)	90 (130)
p = 99 %	230 (388)	299 (473)	459 (662)

4.4. Sensitivity analysis

Sensitivity analysis means the investigation of the influences of parameter uncertainties to the uncertainty of the model output. This is necessary to identify those parameters that have the highest influence on the variability of the model output and should therefore be given the most attention. While parameters to which the model is nearly insensitive need only a rough quantification within their range of uncertainty, those that turn out to cause highly sensitive model reactions are obviously important for the system behaviour should be handled with specific care. If the overall uncertainty of the model output seems too high to fulfil a formal criterion, it should be tried to reduce the uncertainty ranges of the most sensitive parameters.

Sensitivity studies can be performed by varying a specific parameter within its uncertainty interval and leaving all other parameters fixed to their standard value. This is called local sensitivity analysis and is normally done on a deterministic basis with only a few model runs in order to better understand the model behaviour. In contrast, assessing the sensitivities under the influences of the uncertainties of all parameters together is called global sensitivity analysis, which can normally only be effectively performed as a probabilistic investigation. The methods described in this chapter address this kind of sensitivity analysis.

4.4.1. Regression- and correlation-based methods

A group of methods for probabilistic sensitivity analysis is based on linear regression or correlation. These two kinds of methods are mathematically related and yield similar results.

The idea of the regression method is to approximate the actual model by a multi-linear one and to interpret the coefficients of the individual parameters as sensitivity measures.

If the input parameter values of the i-th run are denoted by x_{1i}, …, x_{ki} and the model output by y_i, the linear estimator has the form

$$y_i = b_0 + \sum_{j=1}^{k} b_j x_{ji} + \varepsilon_i$$

with error terms ε_i, which have to be minimised using least squares method. Then the coefficients b_j are a measure for the sensitivity of the model output against variations of the parameters x_j. To allow a unified assessment of these values, the parameters are transformed such that they get the expectation 0 and the standard deviation 1. Then the coefficients are called standardised regression coefficients (SRC). They are always in the range between -1 and 1. The coefficient of model determination, R^2, is defined as the correlation coefficient of the estimated values (without the error correction term) and the real values. If it is close to 0, the estimation is rather bad, whereas a value near 1 indicates a close-to linear behaviour of the model and for this reason a good performance of linear sensitivity analysis methods.

A similar sensitivity measure can be obtained by calculating the linear correlation coefficients between the model output Y and any input parameter X_j. These coefficients are named after Pearson (PEAR).

$$R_{\mathrm{PEAR}}(X_j, Y) = \frac{\sum_{i=1}^{n} (x_{ji} - \bar{x}_j)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_{ji} - \bar{x}_j)^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

For both SRC and PEAR a positive coefficient means that the result increases if the parameter does so; a negative value indicates an inverse correlation. A coefficient of 1 or -1 means a strong linear direct or inverse relationship between input and output. If the coefficient is 0 the parameters are uncorrelated, which means that the output is insensitive to the parameter.

If the input parameters are correlated among themselves, accidentally or on purpose, their influences to the model output are coupled. The methods described so far are unable to resolve this coupling and describe the total influence of the input parameters including that resulting from couplings with other parameters. To separate these influences, two regression ansatzes can be made for the parameter under investigation and for the model output:

$$y_i = b_0 + \sum_{p=1,\, p \neq j}^{k} b_p x_{pi} + \varepsilon_i \qquad x_{ji} = c_{j0} + \sum_{j=1}^{k} c_{jp} x_{pi} + \delta_i$$

The partial correlation coefficient (PCC) is the correlation coefficient between the errors ε_i and δ_i. It is a measure for the model sensitivity to the parameter reduced by external influences. It can be shown that in case of uncorrelated parameters these values are identical to the SRCs.

4.4.2. Rank transformation

The correlation- and regression-based methods work best on models with a nearly linear behaviour. DGR models, however, are typically non-linear and can show a rather complex behaviour, which can lead to low R^2-values and a low performance of the linear methods. This problem can be mitigated by performing a rank transformation on the model input and output. The rank $\mathrm{rk}(x_{ji})$ of a parameter value x_{ji} is its number in an ordered list of all n values x_{j1}, … , x_{jn}. The highest value is assigned the rank 1, the lowest value the rank n. If several values happen to be equal, they are assigned the same rank number, which is calculated as the mean of the ranks they share. By replacing each value by its rank, a monotonic relation is transformed into a linear one.

Rank transformation often (but not always!) leads to an increased R^2 and a better performance of linear sensitivity analysis methods. It has to be kept in mind, however, that there is a considerable loss of information in this kind of transformation. Therefore, the results of a rank-based sensitivity analysis may be more reliable, but are generally less meaningful than those of a value-based analysis.

If the SRCs are calculated on the ranks, they are called standardised rank regression coefficients (SRRC). The rank-based correlation coefficients are named after Spearman (SPEA). It can be shown that – as long as all ranks are different – they can be calculated as

$$R_{\mathrm{SPEA}}(X_j, Y) = 1 - \frac{6\sum_{i=1}^{n}(\mathrm{rk}(x_{ji}) - \mathrm{rk}(y_i))^2}{n(n^2-1)}$$

The PCC concept applied on rank basis yields the partial rank correlation coefficients (PRCC), which, of course, are mathematically equal to the SRRCs in case of uncorrelated parameters.

4.4.3. Non-parametric tests

A number of statistical tests have been developed for investigating the sensitivity of a model to variation of a parameter. In contrast to the linear regression- or correlation-based methods described so far, such tests do not assume a close-to-linear relationship or some other specific kind of model behaviour and are therefore called non-parametric. A typical and often applied non-parametric test is the Kolmogorov-Smirnov test or Smirnov test [CON 99]. This test checks whether there is a significant influence of an input parameter on the model output.

For the Smirnov test, the total of all parameter sets of the sample is separated into two subsamples according to the 90%-quantile of the output. That means that the 10 percent of input parameter sets that lead to the highest output values are separated from the others. The distributions of the parameter under investigation in both subsets are compared with each other. If there is no significant difference, the model can be assumed to be rather insensitive to the parameter. The test is performed by calculating the maximum absolute difference between the empirical distributions of the parameter in both subsamples. The hypothesis of equal distributions is rejected with significance a if this difference exceeds the 1-a quantile c(a) of the Kolmogorov distribution as listed below [Source: Wikipedia].

α	0.10	0.05	0.025	0.01	0.005	0.001
$c(\alpha)$	1.22	1.36	1.48	1.63	1.73	1.95

It was observed that the Smirnov test can yield parameter rankings that differ essentially from those derived from correlation or regression coefficients.

4.4.4. Variance-based sensitivity indices

If the model under consideration is neither linear nor monotonic, which is often the case when dealing with complex repository structures, the linear methods perform rather poorly, and the rank-based sensitivity measures have only a limited quantitative meaning.

Variance-based methods [SAL 00] follow a different approach and do not require linearity of the model. The variance of a statistically distributed parameter is the mean squared deviation from its mean value. To assess the influence of a parameter X_j to the model output Y the expectance of Y is calculated under the condition that X_j remains constant. The variance of this value under variation of X_j is then calculated and divided by the total variance of Y:

$$\frac{\mathrm{Var}_{X_j}[\mathrm{E}(Y|X_j)]}{\mathrm{Var}(Y)}$$

This value is a quantitative measure for the sensitivity of the output to the parameter X_j. It is called the first-order sensitivity index. There are different methods to calculate these indices. A universal, but computationally expensive method is directly based on Sobol's theory of decomposition of the total variance into terms of increasing dimensionality, which yields not only the first-order indices but also all higher orders, describing the influence of a parameter to the output in coactions with other parameters. Of specific interest are the total-order indices, which take account of all possible parameter interactions.

Other methods for calculating the sensitivity indices have been developed. One of them is the Fourier Amplitude Sensitivity Test (FAST) [SAL 97, SAL 00]. The idea is to scan the parameter space by means of periodical functions with interference-free frequencies and to recover these frequencies in the model output by performing a Fourier analysis. Whereas the classical FAST yields only the first-order indices, the extended FAST (EFAST) method also calculates the total-order indices within the same evaluation. This is achieved by varying groups of parameters with the same frequency or harmonics for some periods so that they show up together in the Fourier analysis. In recent years, more computationally efficient methods for calculating variance-based sensitivity indices of first order have been developed, like the EASI method [PLI 10].

In comparison with the linear methods the variance-based methods have some specific advantages. In particular, they allow quantitative assessment and comparison of the parameter sensitivities, even with highly non-linear and non-monotonic models. A drawback, however, is the high number of model runs that is often necessary to get reliable results. Moreover, there seem to be some limitations of these methods in practical application. A big disadvantage of FAST/EFAST is that very specific and non-reusable samples are needed. Moreover, these methods perform rather poorly if applied to problems that depend on discrete or quasi-discrete parameters. A second drawback of the variance-based methods in general is that, though they do not require linearity of the model, the variance is calculated on a linear scale, and if the output varies over several orders of magnitude, high values are essentially overvalued, which can result in a low robustness of the calculated results.

4.4.5. Graphical sensitivity analysis

Graphical methods of sensitivity analysis provide a qualitative overview of parameter sensitivities on first sight and are often very helpful. For such an analysis a number of calculation runs have to be performed on the basis of a parameter sample. Any kind of sampling is possible, although very inhomogeneous samples can produce blurred or misleading results.

Several methods of graphical sensitivity analysis have been proposed, three of which are described here:

- contribution to sample mean (CSM) plots,
- conditional cobweb plots,
- mean rank plots.

For the CSM plot the model runs are sorted according to the values of the parameter under investigation in increasing order. Then for each fraction of all model runs the cumulated relative contribution to the mean of all output values is plotted. These curves always start at (0,0) and end at the point (1,1). While a curve close to the diagonal indicates a low influence of the parameter to the model output, a noticeable bent curve indicates a high sensitivity. If the curve lies below the diagonal, the model output is mainly determined by high parameter values, a curve above the diagonal indicates the opposite. If the CSM curves for several parameters are plotted in one figure, the sensitivities can be easily compared. An example is shown in Figure 4-4.

For a conditional cobweb plot a specific subset of the model runs is selected, for example the 10 percent with the highest model output. The model parameters and the output value are each represented by a vertical line in a diagram, representing the respective range of values or their ranks. For each run of the subset, the relevant parameter values are connected by a line. This leads to a mess of lines, which appears thinned or condensed at the most sensitive parameters. Cobweb plots, however, become useless if the number of evaluated runs – and with that the number of intersecting lines – becomes too large.

In order to get a clearer figure compared with the messy cobweb plots, Cormenzana and Bolado proposed a kind of presentation called mean rank plots [PAM 09b]. For a defined subset of the runs the mean ranks of the parameter values and the model output value are connected by a single line. In this manner, the lines for several different subsets, say the upper and the lower 10 %, can be plotted in one figure, which is very illustrative. Parameters of low sensitivity are met by the line near the middle, which is, in terms of ranks, one half of the number of runs. An example with 20 lines is shown in Figure 4-5.

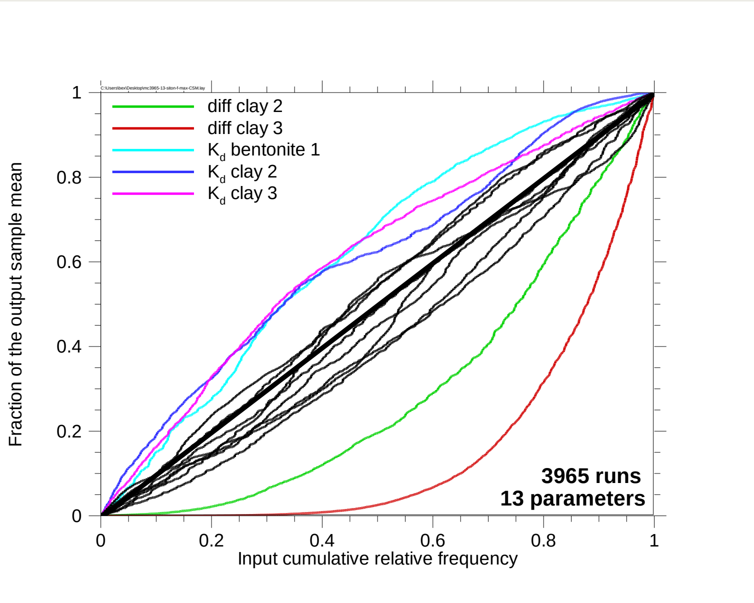

Figure 4-4 CSM plot of a DGR model. 13 parameters have been considered, five of which seem more important (plotted in colour).

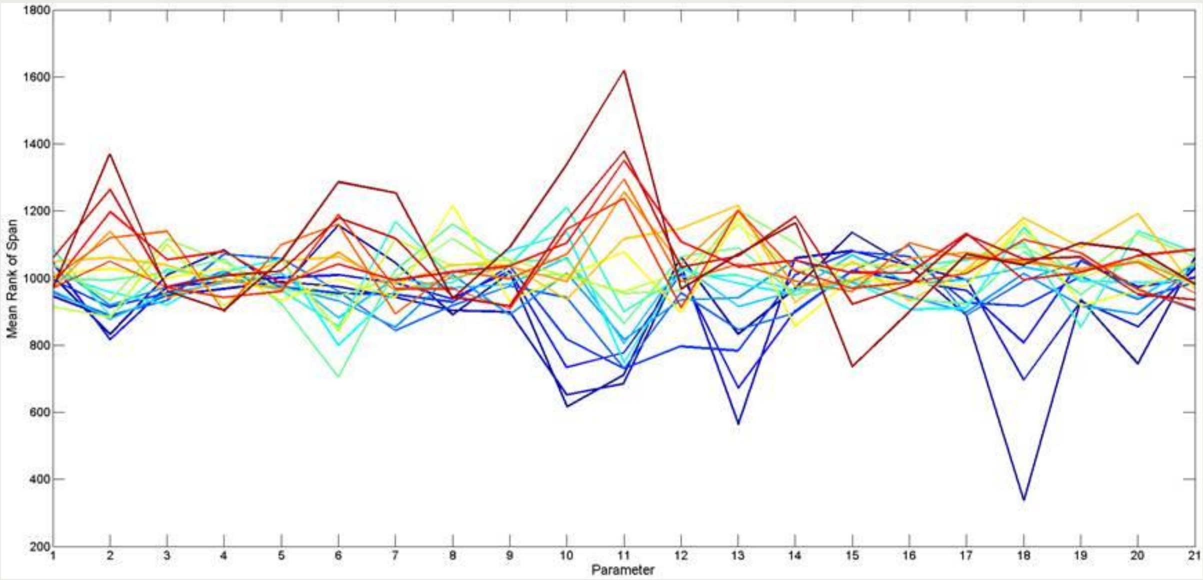

Figure 4-5 Mean rank plot of a DGR model. The 2000 model runs have been divided into 20 subsets according to the model output value. For each subset one line is presented. The model is most sensitive to the parameters with the highest spans.

5. Discussion

Strategies

This document discusses various techniques for uncertainty analysis. However, it is not obvious how to choose between these techniques. For making a choice, a strategy must be used. [NEA 13] describes a number of strategies to handle uncertainties. The following strategies are common in safety assessment:

1. Addressing the uncertainty explicitly.
2. Demonstrating that the uncertainty is irrelevant to the safety assessment.
3. Bounding the uncertainty
4. Ruling out the event or process being uncertain
5. Using an agreed stylised approach.

Although this list of strategies is useful, in practice the strategy is more complex.

In the safety assessment process used in OPERA, a scenario analysis is a key step as a preparation for the quantitative analyses. In OPERA, scenario analysis is treated in Task 7.1. From the viewpoint of strategy, it is recognised that it is uncertain how the disposal system will evolve. This uncertainty is addressed explicitly (strategy 1) by developing a set of scenarios that each may be a description of the system evolution.

To keep the number of scenarios manageable, strategies 2 and 4 are applied in the process of identifying scenarios: irrelevant events and processes are screened out, event and processes that are extremely unlikely are ruled out.

Last but not least, inside each scenario on or more assessment cases are formulated in such way that they are bounding cases for scenario uncertainties (strategy 3).

Once the set of scenarios is defined and calculation models have been developed, the same sequence is repeated. Uncertainties in the results of the calculations are addressed explicitly, but only the uncertainties of a limited number of models and parameters can be treated explicitly by a probabilistic uncertainty and sensitivity analysis. Models and parameters are screened out by demonstrating that uncertainty is irrelevant. And if possible without losing relevant information about the uncertainties, bounding values are used.

Therefore, in practise, the strategy is to address uncertainties explicitly, to the extent that is useful and practical.

Other types of uncertainties

For some at present uncertain parameters it makes no sense to use a probabilistic approach. This is the case, for example, for the waste inventory. At present there is a relatively well known uncertainty about the waste inventory that will accumulate in the Netherlands for final disposal as of 2130. If this uncertainty were included in the probabilistic analysis, the results of this probabilistic analysis would lose their relevance in 2130 when the waste inventory is definitely known. In such cases it is better to establish some 'waste scenarios', for which separate analyses are performed. The same goes for all "technical parameters" relating to the facility design, the host rock and the geosphere.

How to identify the parameters that are of interest for the probabilistic process

The identification of the parameters that should be considered in a probabilistic investigation is not trivial and requires some expertise. Since a high number of probabilistically varied parameters require a lot of work for pdf quantification and a high numerical effort to perform sensitivity analysis, the number should be kept as low as possible. Parameters that allow a clearly conservative, but not over-conservative quantification can be left out of the probabilistic investigation, as long as there are no specific reasons to include them. Often, however, it is not fully clear whether a parameter will under all circumstances influence the model results in the same direction, so that a conservative value can be selected. In such cases it has to be assessed in an adequate manner whether it makes sense to include the parameter in the probabilistic analysis. A graphical screening can be very helpful in this task and is probably the better choice compared with other screening methods like Morris screening [SAL 00]. For orientation a limited number of runs can be performed, nevertheless taking into account all questionable parameters. If no proper pdfs are known, uniform or log-uniform distributions between reasonable limits should be used. A CSM or mean rank plot will then normally show, which parameters are probably important and which do not have significant effects to the model output. The proper sensitivity analysis should then be restricted to those parameters that appear most important, if possible no more than about 10.

<u>What can be obtained with an uncertainty and sensitivity analysis</u>
A well understood result of a sensitivity analysis is firstly a good basis for deepening the understanding of the system behaviour. A robust list of parameter sensitivities can improve the confidence in the model results. Secondly, it is important to know about the parameters that have a high influence to the model output. Further research can then be concentrated on reducing the most important uncertainties.
The latter statement is not straightforward to execute. For example, the OPLA-1A studies show that the transport trough the aquifer system is the dominant source of uncertainty in the far future dose rates. However, this formulation does not reflect that there is very little uncertainty about the processes in the waste, EBS and host rock (salt dome), which is more important. Moreover, the scenario considered is an altered evolution scenario: (unlikely) brine intrusion in the system. The most important uncertainty therefore, is related to the occurrence of brine intrusion.

It is expected now, that through the inclusion of the performance indicator, more nuance is obtained in the results of the uncertainty and sensitivity analysis.

6. Conclusion

The following concepts for dealing with uncertainties have been discussed:
- Scenario analysis
- Conservative approach
- Deterministic investigation
- Probabilistic analysis

For the uncertainty analysis process, the following steps are foreseen:
1. establish a set of scenarios (Task 7.1)
2. establish for each scenario a number of assessment cases (i.e. various inventories, focus on specific waste types)
3. establish models and (conservative) parameter values for the assessment cases.
4. for each assessment case, identify the parameters that are of interest for the probabilistic process
5. establish pdf's for these parameters
6. perform the analyses
7. evaluate using one or more of the techniques described (e.g. regression- and correlation-based methods, or rank transformation, etc.)

A careful evaluation of the results of the uncertainty and sensitivity analysis is required to show the robustness of the conclusions that can be drawn from the safety assessment.

7. References

[BEU 12] Beuth, T. et al.: *Szenarienentwicklung: Methodik und Anwendung. Bericht zum Arbeitspaket 8, Vorläufige Sicherheitanalyse für den Standort Gorleben*. GRS-280, Köln 2012

[BMU 10] Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU): *Sicherheitsanforderungen an die Endlagerung wärmeentwickelnder radioaktiver Abfälle*. Stand: 30. September 2010, Bonn 2010

[CON 99] Conover, W. J.: *Practical nonparametric statistics*. Wiley, 1999

[IAEA 09] IAEA: *Safety Assessment for Facilities and Activities*. IAEA GSR Part 4, IAEA, Vienna 2009

[IAEA 12] IAEA: *The Safety Case and Safety Assessment for the Disposal of Radioactive Waste*. IAEA SSG-23, Vienna 2012

[NEA 13] NEA: *The Nature and Purpose of the Post-closure Safety Cases for Geological Repositories*. NEA/RWM/R(2013)1, 2013

[PAM 08a] Marivoet, J. et al.: *Task reports for the first group of topics: Safety Functions, Definition and Assessment of Scenarios, Uncertainty Management and Uncertainty Analysis, Safety Indicators and Performance/Function Indicators*. EC-PAMINA Del. 1.1.1, 2008

[PAM 08b] Becker, D.-A. et al.: Protocol for Assessing Parameter Uncertainty. EC-PAMINA M 2.2.A.2, 2008 (unpublished, available from GRS Braunschweig)

[PAM 09a] Bolado, R. et al.: *Review of expert judgment methods for assigning PDFs*. EC-PAMIINA M 2.2.A.3, 2009

[PAM 09b] Cormenzana López, J. L. and Bolado Lavin, R.: *Sensitivity/uncertainty analyses – Application to a repository in granite*. EC-PAMINA M2.1.D.8, 2009

[PAM 11] Bailey, L. et al.: *European Handbook of the state-of-the-art of safety assessments of geological repositories-Part 1*. EC-PAMINA Del. 1.1.4, 2011

[PLI 10] Plischke, E.: *An effective algorithm for computing global sensitivity indices (EASI)*. Reliability Engineering & System Safety 95, 354-360 (2010)

[ROC 08] de Rocquigny, E, Devictor, N and Tarantola, S. (ed.): *Uncertainty in Industrial Practice – A guide to quantitative uncertainty management*. Wiley 2008

[SAL 00] Saltelli, A. (ed.): *Sensitivity Analysis*. Wiley, 2000

[SOB 67] Sobol, I. M.: *Distribution of points in a cube and approximate evaluation of integrals*. U.S.S.R Comput. Maths. Math. Phys. 7: 86–112, 1967

[VER 11] Verhoef, E., Schröder, T.J., Research Plan, OPERA-PG-COV004, COVRA N.V., 2011.

[YAT 03] Yates, D.; Moore, D; Starnes, D.: *The Practice of Statistics* (2nd ed.). New York, Freeman 2003

OPERA

Meer informatie:

Postadres
Postbus 202
4380 AE Vlissingen

T 0113-616 666
F 0113-616 650
E info@covra.nl

www.covra.nl