EUROPEAN COMMISSION

Community research

PAMINA
Performance Assessment Methodologies in Application to Guide the Development of the Safety Case

(Contract Number: FP6-036404)

TREATMENT OF PARAMETER UNCERTAINTY IN PA DELIVERABLE (D-N°:2.2.A.1)

Author(s):
D.-A. Becker, O. Destin, H. Nordman, R. Bolado, L. Duro, J.L. Cormenzana, M.A. Cuñado, A. Vetešník, R. Avila, K. Stenberg

Date of issue of this report : **16/12/2009**

Start date of project : **01/10/2006** Duration : 36 Months

Project co-funded by the European Commission under the Euratom Research and Training Programme on Nuclear Energy within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
RE	Restricted to a group specified by the partners of the [PAMINA] project	
CO	Confidential, only for partners of the [PAMINA] project	

Foreword

The work presented in this report was developed within the Integrated Project PAMINA: **P**erformance **A**ssessment **M**ethodologies **IN A**pplication to Guide the Development of the Safety Case. This project is part of the Sixth Framework Programme of the European Commission. It brings together 25 organisations from ten European countries and one EC Joint Research Centre in order to improve and harmonise methodologies and tools for demonstrating the safety of deep geological disposal of long-lived radioactive waste for different waste types, repository designs and geological environments. The results will be of interest to national waste management organisations, regulators and lay stakeholders.

The work is organised in four Research and Technology Development Components (RTDCs) and one additional component dealing with knowledge management and dissemination of knowledge:

- In RTDC 1 the aim is to evaluate the state of the art of methodologies and approaches needed for assessing the safety of deep geological disposal, on the basis of comprehensive review of international practice. This work includes the identification of any deficiencies in methods and tools.

- In RTDC 2 the aim is to establish a framework and methodology for the treatment of uncertainty during PA and safety case development. Guidance on, and examples of, good practice will be provided on the communication and treatment of different types of uncertainty, spatial variability, the development of probabilistic safety assessment tools, and techniques for sensitivity and uncertainty analysis.

- In RTDC 3 the aim is to develop methodologies and tools for integrated PA for various geological disposal concepts. This work includes the development of PA scenarios, of the PA approach to gas migration processes, of the PA approach to radionuclide source term modelling, and of safety and performance indicators.

- In RTDC 4 the aim is to conduct several benchmark exercises on specific processes, in which quantitative comparisons are made between approaches that rely on simplifying assumptions and models, and those that rely on complex models that take into account a more complete process conceptualization in space and time.

The work presented in this report was performed in the scope of RTDC 2.

All PAMINA reports can be downloaded from http://www.ip-pamina.eu.

Executive Summary

Input parameters for long-term safety assessment models are always subject to specific uncertainties, which have to be taken into account in an adequate manner. One possibility of doing this is to evaluate the model a number of times with different values for the input parameters. For this purpose, the parameter uncertainties have to be quantified.

The work performed in the PAMINA task described in this report aimed at demonstrating the influences of parameter uncertainties on typical models as well as identifying methods and techniques for determination and validation of probability density functions (PDFs). Several individual studies have been performed, each addressing a specific issue. Each of the studies is independent of the others, but altogether they provide an overview of the specific problems in this context as well as some approaches to resolving them.

Table of Contents


PAMINA

1. Introduction

The treatment of uncertainties is a very essential issue in the field of repository performance assessment (PA). In order to get an impression of the credibility of the calculated results it is necessary to analyse the influences of the various uncertainties. Such uncertainties result from the model itself, from the considered scenarios or from the input parameters. This document deals with parameter uncertainties.

Two different approaches to handling parameter uncertainties can be distinguished. One possibility is to choose pessimistic values for all parameters, based on expert assessment. The model results will then represent the 'worst case', which is unlikely to be exceeded with regard to disadvantageous environmental effects by the real system evolution. Such deterministic studies are often supported by 'what-if'-investigations, considering some extreme situations or parameter values. Some models, however, show a non-linear behaviour and do not allow a uniquely pessimistic choice of all parameters. In such cases the parameter uncertainties should be investigated more systematically by performing a higher number of model runs with varying combinations of input parameters. This approach, which often uses probabilistic techniques, requires a quantification of the parameter uncertainties. PAMINA task 2.2.A mainly focused on this issue.

Especially in the field of long-term performance assessment for radioactive waste repositories, all model input parameters are more or less uncertain. This can be due to a scatter of physical values that is, by principle, irreducible and has to be accepted as it is. Uncertainties of this kind are called aleatory. Another source of parameter uncertainty is our insufficient knowledge of the system. The value of a parameter may be physically well-defined, but we do not know it, due to lack of appropriate measuring techniques, insufficient number or poor exactness of measurements, or similar reasons. Such uncertainties are called epistemic. It is essential that they can be reduced by additional scientific effort. Parameter uncertainties can often not be uniquely classified as one or the other type, since most of them are a bit of both. In the field of long-term PA, however, the epistemic character dominates in most cases.

The calculation of statistical characteristics of the output of a given model under the influence of given parameter uncertainties is called uncertainty analysis. It is performed in order to assess the uncertainty of the model output and to get an impression of how reliable the results really are. If the goal is to identify those parameters that predominantly affect the model output and that are worth of putting effort in reducing their uncertainty, one has to perform a sensitivity analysis, quantifying the model output uncertainty against the uncertainties of individual input parameters.

Both uncertainty and sensitivity analysis can be performed as deterministic or probabilistic analysis. While in deterministic investigations one tries to get a feeling for the model behaviour by varying specific parameters individually between separate values, regardless of

their probability, a probabilistic analysis is done by running the model a big number of times with statistically distributed values for all parameters.

Deterministic investigations have a more qualitative character and are performed to improve the general understanding of the system. Probabilistic analyses, however, are designed as exercises to quantify the output uncertainty exactly. For such analyses, probability distribution functions (PDFs) for all input parameters under investigation have to be established. These should reflect, as well as possible, the actual uncertainty of the parameters, because otherwise, the analysis will not yield the actual uncertainty of the model output.

While for aleatory uncertainties the PDFs should be clear or derivable from models, for epistemic uncertainties establishing correct PDFs can be a difficult task, because it means a quantification of knowledge, which is inevitably subjective to a certain extent. As mentioned above, epistemic uncertainties play a dominant role in the field of long-term PA. Therefore, to get unassailable results, it is essential to minimise the subjectivity of the process.

The report at hand summarises the work performed in the framework of the PAMINA project concerning the treatment of parameter uncertainties and especially the process of establishing PDFs, as well as their influence to the model output. The investigations performed by the individual partners complement one another and provide a basis for well-funded uncertainty analyses.

VTT performed a deterministic uncertainty analysis exercise with a realistic repository model system. Several parameters were varied without regarding probabilities. This investigation shows how valuable information about the functioning of the system can be derived from deterministic parameter variations.

GRS developed a protocol for determining PDFs for specific parameters under a given knowledge situation. The procedure is designed to minimise subjectivity. It should be seen as a guidance on the problem of establishing PDFs that helps justifying the individual decisions and making them traceable.

BEL V evaluated PDFs for several key radionuclide transport parameters, focusing on the potential impact on PA results of different PDF shapes. The aim of the study was to investigate the potential influence of an arbitrary use of PDFs on the performance indicators commonly used to assess the performance of a disposal system.

Facilia demonstrated how Bayesian methods can be used for deriving probability distributions of model parameters using empirical data and other prior information. Three methods were studied: 1) direct updating for deriving the PDF of a parameter using empirical data available in combination with PDF available for an analogue, 2) hierarchical models – for deriving the PDF of the parameter of interest using available information for several analogues and 3) updating of the coefficients of a regression model, which can be used to

generate the PDF of a parameter of interest from information about explanatory variables in the model.

NRI worked on devising the general procedure to treating uncertainty in parameters of a computational model used in PA. A PDF derivation tool was developed that allows construction of a best-fitting PDF in case of sufficient data.

JRC in cooperation with AMPHOS21 and ENRESA focused on the subject of expert elicitation. In case of epistemic uncertainties the PDFs have to reflect our level of knowledge, which is often hard to quantify. Expert judgements are necessary to get a reliable assessment of the situation. An expert elicitation protocol was developed and applied to a test case, which aimed at determining PDFs for radionuclide solubilities.

In the following chapters summaries of the individual pieces of work are given with their main results. Detailed documentations are provided by the individual organisations.

2. Parameter uncertainty in the Finnish KBS-3 concept (VTT)

In several safety analyses for a KBS-3 concept it has been noted that a lot of parameters are not important and some parameters are. For instance in the safety assessment TILA-99 [Vieno and Nordman, 1999] there are about 100 calculation cases with different parameters, but between several cases there are only small differences in the release rates of radionuclides. That is why VTT has carried out a comprehensive evaluation of the importance of the parameters and the effect of their uncertainty in certain cases.

2.1 Importance of near-field parameters

In addition to equivalent flow rates from the near field and retardation parameters the diameter of the defect has been varied. The different diameters, flowrates and corresponding acronyms of cases are:

- 1 mm diameter, acronym Sh1 for default flow and Sh1 Q for high flow
- 4 mm diameter, acronym Sh4 for default flow and Sh4 Q for high flow
- 100 mm diameter, acronym Lh for default flow and Lh Q for high flow

2.1.1 Effect of defect size

The postulated defect diameters were 1 mm, 4 mm and 100 mm. The chosen high flow case decreases the effect of the far-field thus emphasizing the significance of the defect size. The effect of increasing the defect size from 1 mm to 4 mm results in increase of release rates of most of the nuclides from the near field by a factor of 10...15 which is just slightly less than the ratio of the defect areas (factor of 16). Further increase of the defect area by a factor of 625 shows a reduced increase of release rates compared to the ratio of areas. In the large hole case (diameter 100 mm) other resistances in the near field limit then the releases - e.g. flowrate to geosphere or diffusion resistance in bentonite.

2.1.2 Source term and solubility limit

An important parameter is the IRF (instant release factor) of a nuclide as a part of the source term. For some relevant nuclides like I-129 the dose rate is linearly dependent on the IRF. But for some nuclides like C-14 the IRF is not dominating as a sufficient majority of their inventories is in the UO2 matrix or in the metal parts of a fuel element. The importance and uncertainty of IRF has been evaluated compared to degradation rate of metal parts and UO2 matrix. Also the importance of solubility limit is strongly coupled to the source term from fuel elements. For some nuclides like Zr-93 and Tc-99 (for Tc in reducing conditions) the solubility limit is always applicable and thus diminishes the roles of the IRF and other source terms.

Some additional examples of the importance of the solubility limit are presented in Table 1.

Table 1 Onset of solubility limits.

Nuclide	Solubility limit reached	Comment on cases
U, Zr, Np	Always	For all cases and assumptions.
Ra	Mostly	Ra-226 becomes solubility limited due to in-growth between 30 000 – 300 000 years depending on the case.
Pu	Mostly	In all small defect cases solubility limited. In some large defect cases Pu is limited for some time
Pa	In Sh cases only (1 mm or 4 mm diameter)	Pa is solubility limited later in time due to in-growth and release from the fuel in small defect cases. The earliest times are in the cases of high fuel degradation rate.

2.1.3 Effect of retardation parameters

The effect of groundwater salinity on near-field release rates is presented in Table 2 and Figure 1. Figure 1 shows the near-field release rates of Sr-90 and Ra-226 in cases Sh4 Q and Sh4 Q Sal (defect diameter 4 mm, high flow and brackish/dilute or saline chemistry). Sr and Ra are sensitive to the assumed variations in the chemistry (saline/non-saline) of groundwater. The Kd values in saline chemistry are about 4-5 times lower.

Table 2 Maximum release rates from near field. Effect of salinity in the release rates for the near-field in Sh4 case (defect diameter 4 mm)

Radionuclide		Sh4		Sh 4 saline	
		t_{max} [a]	Bq/a	t_{max} [a]	Bq/a
	Sr-90	2.86E+01	2.00E+06	2.86E+01	9.89E+06
	Ra-226	1.00E+06	9.09E+02	6.02E+04	2.96E+03

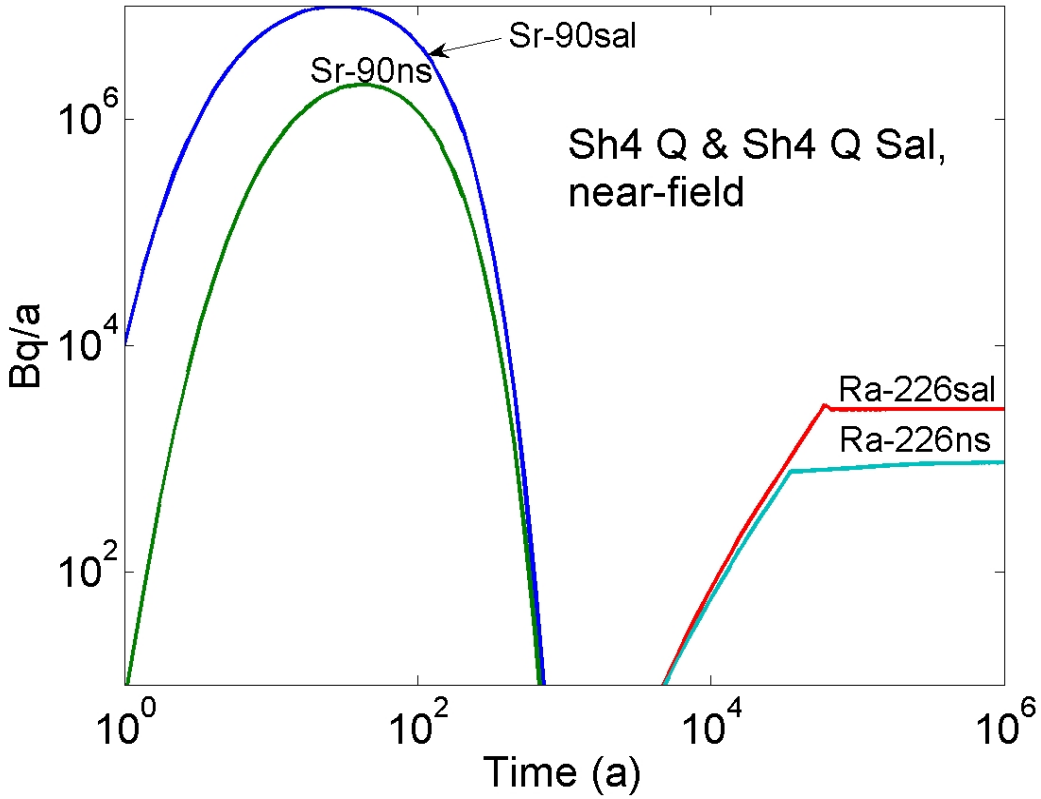


Figure 1 Effect of salinity on Ra-226 and Sr-90 release rates from near field in the small defect and high flow case Sh4 Q.

If plutonium species is pentavalent the retardation is very low compared to neutral species. The K_d values for pentavalent plutonium are about three orders of magnitude smaller compared to other species of plutonium. The near-field release rates for both species with high flow and a large defect are presented in Table 3.

Table 3 Maximum release rates from near field. Effect of Pu-239 speciation on the release rates for the near-field in Lh Q case (high flow and defect diameter 100 mm).

Base value		Pentavalent	
t_{max} [a]	Bq/a	t_{max} [a]	Bq/a
5.52E+05	3.37E+03	1.86E+04	1.29E+06

The release rate increases with a factor of about 400 when plutonium is pentavalent. This result is quite self evident.

2.1.4 Effect of near-field flow rates

The behaviour of the system regarding radionuclide transport was checked against possible changes in the flow distribution in the near field. A set of supplementary calculation cases was formed in order to study the impact of changes in the near-field flow distribution. Flow rates and transport resistance were varied by one order of magnitude.

In the analysis of these cases solubility, sorption, and data for buffer and backfill for four radionuclides with two neutral species (Ra 226, C 14), one anion (Cl-36) and one cation (Cs 135), and for two non-sorbing generic species (neutral and anion, both stable) are used. A unit release of 1 GBq (delta pulse) to the canister interior is assumed for the generic neutral and anionic species. Only large defect case is analysed as in the Sh1 cases, the hole is the limiting factor.

Some observations are provided below.

- For the anionic species (like IRF of I-129) the maximum release rate increase by a factor of 3 when flow rates increase. For Cl-36 the increase is very small (about 10 %) as the release from zircalloy metal parts is the dominating phenomenon. The release from these metal parts is assumed to take 10 000 years.
- The time of maximum release rate of Ra-226 is about at the time when its solubility limit is reached. In the case where the flow rates to the geosphere are 10 times higher, the release rate is 8 times higher.
- The two other non-anionic species, neutral and C-14, being non-sorbing with Kd values as zero, behave in a different way from Ra-226 as they are not solubility limited. The time of the maximum release rate of C-14 is when corrosion of metal parts ends. The release rate increase is about a factor of 3-4 when flow is 10 times higher.
- The IRF fraction of Cs-135 inventory governs the early behaviour of the release rate and can be fastest transported into the fracture intersecting the deposition hole if the flow rates are high. In high flow case the release rate is about 8 times higher.

2.2 Importance of geosphere parameters

An important thing to be evaluated is the importance of the so called u parameter which represents the transport resistance of geosphere. The u parameter consist of retardation data on rock matrix and ground water flow conditions. The importance of this geosphere parameter depends on the half life of a nuclide and also on the shape of the release pulse from near field. If the fuel degradation time is long and canister life time also is long the relative importance of geosphere can be minor.

2.2.1 Conceptual model

The model considers advection along fractures and matrix diffusion is the only phenomenon assumed to lead to dispersion and retardation. After entering the rock matrix, the radionuclides also are affected by sorption. Diffusion and sorption in the rock matrix are modelled according to the porous medium and linear equilibrium sorption approaches. The essential characteristics of a transport route can thus be presented by means of a single parameter (u) taking into account the distribution of groundwater flow in the fracture system and the effects of matrix diffusion.

For the case where a constant water phase concentration C_o of a stable species is prevailing at the inlet of the fracture beginning from t = 0, the water phase concentration at the distance of L in the fracture is [Vieno and Nordman, 1999]:

$$C_f(L,t) = C_o \operatorname{erfc}\left[\frac{u}{\sqrt{t}}\right],$$

where u is a parameter describing the transport properties of the migration route for the given species

$$u = \left[\varepsilon_p D_e R_p\right]^{1/2} \cdot \frac{W L}{Q}$$

ε_p is the porosity of the rock matrix (-),

D_e is the effective diffusion coefficient from the fracture into the rock matrix (m^2/s),

Rp is the retardation factor of the species in the rock matrix (-),

W is the width of the flow channel i.e. the width, over which the flow is measured (m),

L is the transport distance (m),

Q is the flow rate in the channel or over the given width (m^3/a),

t is the time (a).

The first factor of the u-parameter is related to diffusion and sorption in the rock matrix. The retardation factor, R_p, depends on the rock properties and the volume-based distribution coefficient, K_d (m^3/kg).

In the above equations the matrix diffusion is unlimited. The traditional matrix porosity and effective diffusion coefficients for the far field are shown in Table 4 (data from [Vieno and Nordman, 1999]). The matrix diffusion is limited to 10 cm from the fracture.

Table 4 The traditional matrix porosity and effective diffusion coefficients for the far field.

Parameter	Distance from fracture	Species	Brackish/dilute
Porosity	0 – 1 cm	Anions	0.1%
		Neutral and cationic species	0.5%
	1 - 10 cm	Anions	0.02%
		Neutral and cationic species	0.1%
Effective diffusion coefficient	0 – 1 cm	Anions	10^{-14} m^2 s^{-1}
		Neutral and cationic species	10^{-13} m^2 s^{-1}
	1 - 10 cm	Anions	10^{-15} m^2 s^{-1}
		Neutral and cationic species	10^{-14} m^2 s^{-1}

The WL/Q has traditionally had values between 5000..50 000 a/m, but in this work some more extreme values have been also used.

Effect of u parameter on certain important nuclides has been analysed. In these cases the release pulses from near-field are similar to those of the high flow and large defect case Lh Q [Nykyri et al., 2008]. The FTRANS code has been used [Vieno and Nordman, 1999].

C-14, half life $5.7 \cdot 10^3$ a

C-14 is typically considered to be a non sorbing neutral (organic) or slightly sorbing anion (carbonate) species, meaning that its retardation in geosphere is typically regarded negligible. If the data in Table 4 is used and WL/Q is 5000 a/m, the C-14 pulse from near field does not decrease at all. The maximum release from geosphere occurs typically just after 1000 years when the metal parts containing 33 % of C-14 inventory have been

corroded. The analyses showed that for C-14 a small Kd value is sufficient to effectively decrease the peak value of the near field pulse within the geosphere. N.B. The near field pulse is quite flat as it's a result from evenly corroding metal parts.

I-129, half life $1.6 \cdot 10^7$ a

I-129 is non sorbing and anion so the retardation in geosphere is typically negligible. If the allowed rock matrix depth is increased from 0.1 m to 3 m the I-129 release to biosphere decreases only slightly (~ 10 %). Quite extreme values of WL/Q are needed in order to get considerable decrease in near-field release in the geosphere migration process. In addition the total integrated release does not change, since matrix diffusion only affects the timing of release. The peak value from near field resulting from instant release fraction is decreased by a factor of only 4 if a very high value of 5 000 000 a/m for WL/Q was used. (Traditionally WL/Q has values between 5000…50 000 a/m.)

Cs-135, half life $2.3 \cdot 10^6$ a

Cs-135 is cationic and has typically considered to be a sorbing species with a Kd value of at least 0.05 m3/kg (used in this work). In the following, the effect of WL/Q has been studied. The effect of increased matrix diffusion depth has not been studied as Cs has so high retardation. With WL/Q values of 200 000 and 500 000 a/m the release rate is dramatically decreased.

Ra-226, half life $1.6 \cdot 10^3$ a

In case where Table 4 data is changed so that the first 1 cm of rock matrix has a 10 times lower De value, the effect is dramatic. The increase of release rate is more than two orders of magnitude. This is due to the fact that Ra-226 has high sorption and short half life.

Pa-231, half life $3.2 \cdot 10^4$ a

If the u parameter has a value of 1030 a½ then Pa-231 release rate decreases dramatically.

2.3 Conclusions

In general, the most important nuclides in most cases are I-129, C-14, Cs-135, Ra-226 and Pa-231. Principal features affecting these nuclides, in addition to their relatively long half lives, are:

- I-129: high IRF and weak retardation
- C-14: large inventory in metal parts is released quickly; weak retardation
- Cs-135: high IRF plus fairly weak retardation
- Ra-226; saline chemistry

- Pa-231: fuel degradation rate and solubility limit – important only in cases with high flow. The in-growth from its parent U-235 is not a major factor in geosphere.

The release rates of radionuclides to the biosphere depend on many parameters. The main factors affecting the results are presented in Table 5 with cases where the relevant factors are pointed out together with explanatory comments.

Table 5 Main factors affecting the release rates.

Factor	Comment
Defect size	In cases where the diameter is 4 mm the release rates for Cs and neutral species are about 10 times higher than in cases where the diameter is 1 mm. The ratio of defect areas is 16 which shows that other barriers have a larger role in the 4 mm case. In cases where the diameter is 100 mm size the release rates are naturally even higher, but the roles of other barriers are also even more significant.
IRF	When the defect is large the IRF is more dominant, see e.g. I-129 compared with long-lasting releases. Uncertainties in IRF are reflected directly proportional to the maximum release rates.
Fuel degradation rate	The release of Pa-231 linearly depends on the fuel degradation rate in the Large defect Lh cases, but not in Sh1case as Pa-231 is solubility limited in this case where the small size of the defect causes accumulation of Pa in the canister water volume.
Salinity and retardation	The largest effect are the higher releases of Ra-226 and Sr-90 because of reduced retardation in saline flow conditions (Table 2). Pu-239 is an example of nuclide which normally is assumed to have a high retardation but is released faster with lower retardation as pentavalent (Table 3).
Near field flow rate	The effect is most notable for short-lived sorbing nuclides in the far-field transport. Cation releases from the near field are typically one order of magnitude higher with high flow rate. The effects for anions is much lower, because the diffusion resistance in the near field dominate overall transport and release.
WL/Q	Shows the effect of far-field parameter variations. The lumped parameter u values can be interpreted as variations of WL/Q.
Matrix diffusion properties	For Ra-226 the importance is most significant.
Onset of solubility limit	Depends on the degradation of fuel element parts, chemical species of radionuclides, potential parent nuclide, other isotopes of the element in question and on equivalent flow from canister interior to bentonite.

3. Review of parameter uncertainty methods, results of test cases and guidance (BEL V)

As mentioned in [PAMINA 2006], RTDC-2 is dedicated to the Management of Uncertainty during Safety Case Development and one of its Work-Packages (WP), the WP 2.2, is focused on Testing and Development of Approaches for Management of Uncertainty. Finally Task 2.2.A is related to Parameter Uncertainty and the particular role of Bel V (formerly AVN) is defined in Milestone 10 as:

Bel V will evaluate PDFs for several key radionuclide transport parameters, focusing on the potential impact on Performance Assessments (PA) results of different PDF shapes. In the first stage, this will involve testing the sensitivity of PA to different PDF shapes, in order to develop guidance on how PDF shape may vary for different radionuclides. In the second stage of work, guidance will be developed on how representative values (mean, mode,...) from different PDFs may be derived for these key parameters.

3.1 Chosen model

In order to achieve this work a 2D model of a repository system, representative of the Belgian case and inspired by [OND01], has been used. In this geological disposal model the host rock has properties of the Boom Clay. The 2D model is an accommodation between the computing time required and the representativeness of physical phenomena.The outputs studied were similar to usual performance indicators (like radionuclide fluxes at the interface geosphere - biosphere).

Hence, the maximum flux and the time of the maximum flux have been chosen. The outputs' evaluation has been achieved with the MELODIE Finite Element code provided by IRSN [BOU03], which solves both flow and migration equations.

Use of PDFs as a way to model parameter uncertainty of flow and migration models is common practice in the field of PA of radioactive waste disposal systems. Because it can be difficult to obtain direct experimental

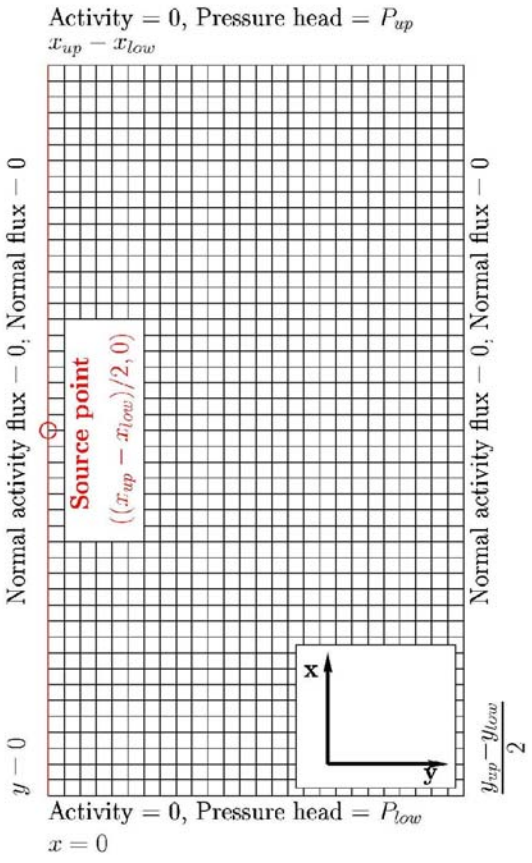


Figure 2 Outline of the 2D model. The red line represents the symmetry BC.

data to quantify the chosen key parameters, PDFs are often arbitrarily chosen by the assessors. The shape and higher moments of these PDFs are usually very different.

The aim of this study is to investigate the potential influence of an arbitrary use of PDFs on the performance indicators commonly used to assess the performance of a disposal system. Assessing this influence may be treated by the methodology of Sensitivity Analysis (SA), where the parameters of this SA are the different parameters characterizing the various PDFs. In these conditions, a global sensitivity analysis, as presented in [Saltelli et al., 2008] would lead to a complex exercise with difficult interpretation. This kind of exercise is indeed more appropriate when the number of parameters is quite important and when the values of the considered parameters for the SA are continuous. Consequently, a rather simple method, especially valuable when the number of relevant parameters remains small is to directly compare the output results of different sets of PDFs instead of testing the sensitivity of PA outputs as in a global SA. This approach does not rely on sophisticated analysis but is consistent enough to make judgement about the impact of PDFs on PA outputs and is much easier for interpretation. Therefore, this latest approach has been chosen by Bel V.

Based on these assumptions, the generic method used for the study of a particular radionuclide consists of:

- Identifying potential parameters for which it makes sense to define a PDF,
- Generating the input samples based on the determined sets of PDF,
- Computing the output PDF,
- Analyzing the results.

3.2 Data used

The study is focused on the radionuclides and parameters for which the amount of experimental data is the most extended.

All the data used come from [SCK ·CEN, 1999] & [ONDRAF/NIRAS, 2001]. Based on the data, only iodine has been used for the study, essentially because of its very high solubility limit. Two key parameters have been selected: the molecular diffusion (d) and the porosity (ω).

Different common PDFs shapes have been used for both key parameters selected: uniform, log-uniform, normal, log-normal, triangular and the so called triangular “right shifted”. Practically, 10.000 LHS have been generated for the different PDFs couples tested. C++ code has been developed to bind the MELODIE code for pre-processing and post-processing.

3.3 Results

Two scenarios have been considered for the study:

- The reference scenario where the host rock (Boom Clay) is considered as a purely diffusive medium;
- An altered scenario where the advection process has been significantly increased through the hydraulic conductivity parameter.

The four first moments and some quantiles have been computed from the outputs distributions (ODs) and have been compared to each others in both scenarios.

Concerning the reference scenario, the characteristics of the ODs are strongly related to the input PDF (IPDF) of the molecular diffusion for both outputs, since there is only a diffusion phenomenon. Results are relatively straight forward and easy to interpret. Conservative assumptions on IPDFs could be made because results can be anticipated. Consequently this scenario is not relevant to highlight the effect of IPDFs on conservatism. The only thing that can be mentioned is that IPDF has a real influence on the output.

As far as the chosen altered scenario is concerned, diffusion and advection are of the same order of magnitude. IPDFs have different impacts on ODs with regard to the output considered. The impact is more important on the time of the maximum flux than on the value of the maximum flux itself. The influence of off-centred IPDFs (non zero skewness) is greater than centred (zero skewness) for ODs. The main conclusion is that a conservative choice for an IPDF can not be made without making such a study.

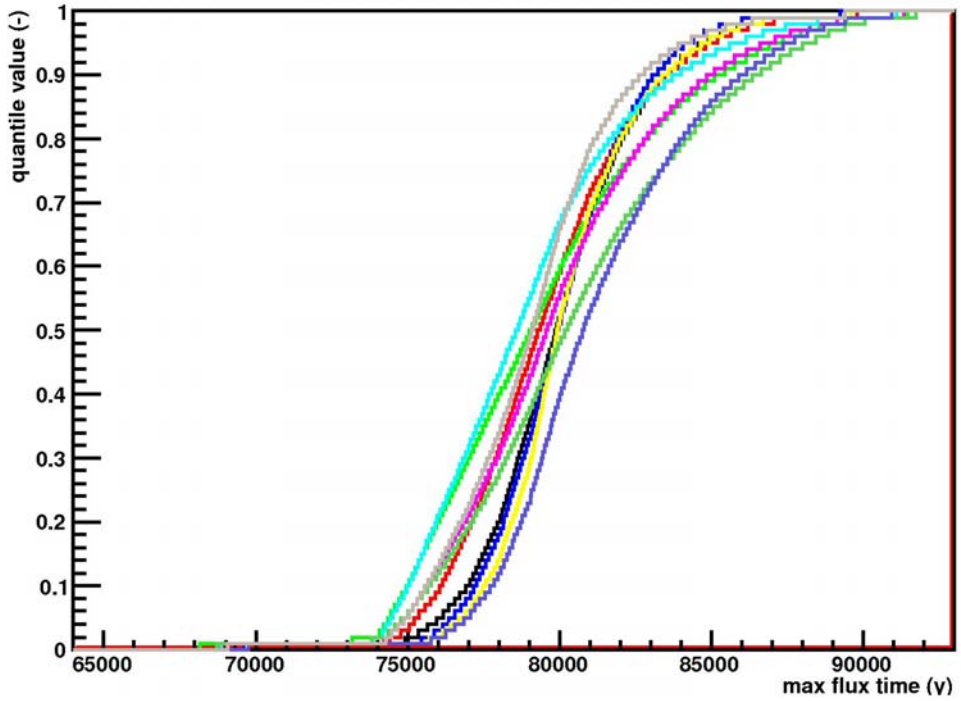


Figure 3 Output cumulative distribution for the time of the maximum iodine flux (see Bel V's contribution for more details).

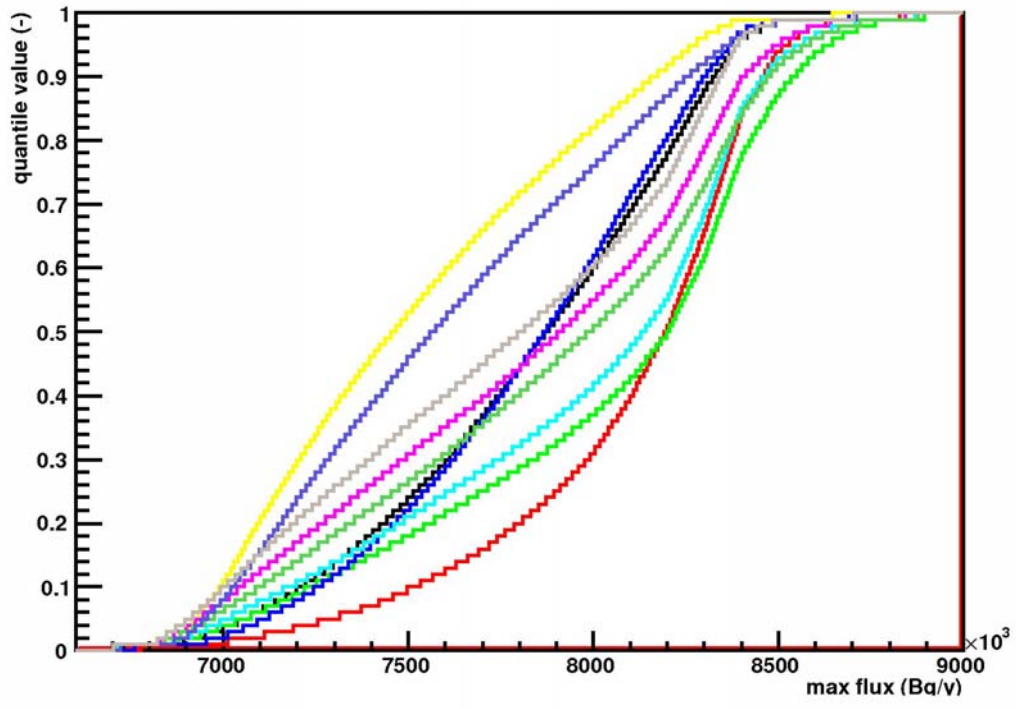


Figure 4 Output cumulative distribution for the maximum iodine flux (see Bel V's contribution for more details).

3.4 Conclusion

In the scope of the present study it has been possible to put in evidence that simplified assessment models taking only into account predominant flow and migration mechanisms provide an easy tool for the treatment of parameter uncertainty. In particular, the interpretation of the behaviours of commonly used PI for different types of IPDFs modelling the uncertainty on the value of the diffusion coefficient is straightforward. It has been observed that for the currently used assumptions for the Boom Clay the impact of the IPDF-type on the interpretation of the PI results is quite limited. These findings would favor the use of simple assessment models (when possible) and put the emphasis on the necessity of thorough justifications of each assumptions used to simplify the model.

This study allowed also to illustrate the more complex situation (for the interpretation of the results and for the treatment of the parameter uncertainty) when simple models cannot be used anymore. In this case the treatment of uncertainty for multiple parameters for which the used IPDFs is not well known is no more straightforward. The sensitivity of the PI results on the IPDF-types seems moreover higher. In this case it would be preferable either to rely on more conservative models (integrating the parameter uncertainty into more conservative BE approach) or to develop a multiple parameter uncertainty assessment with only well-known IPDFs. Moreover, the study highlights that a conservative assumption for IPDFs is not possible in general without a dedicated study and detailed justifications. From a TSO point of view, an arbitrary choice of IPDF without any guarantee of conservatism is not acceptable.

4. Protocol for assessing parameter uncertainty (GRS)

Internationally there is a high degree of consensus on both the nature of uncertainties and how they should be classified in PA [Galson and Khursheed, 2007]. The quantification of parameter uncertainties, however, is not a simple task. Protocols for systematically deriving PDFs are not yet established internationally. PDFs and their characteristics have been sometimes established in the past by modellers rating the general shape of the PDF and its characteristics, such as bandwidth or standard deviation. This approach is highly subjective and unsatisfying since in most cases it is neither traceable nor reproducible, which can be a serious problem in a licensing procedure.

The purpose of the work described here is to propose a protocol for assessing the uncertainty of parameters on the basis of available data. This protocol should urge the modeller to think carefully about the uncertainty of parameters instead of simply defining a PDF out of a rough feeling. Moreover, it should reduce the subjectivity of the uncertainty quantification process and create a traceable scheme for PDF derivation.

The general procedure for describing the parameter uncertainties via PDFs in a PA study is based on three steps:

- identification of the parameters that need to be considered,
- selection of a suitable knowledge base to be used for deriving a PDF for each parameter to be considered,
- derivation of the PDF for each parameter to be considered following the general scheme that is proposed here.

4.1 Identification of uncertain parameters

All input parameters of a model are more or less uncertain. However, there are always parameters for which the values are sufficiently well-known so that their associated uncertainty can be disregarded in the probabilistic model calculations. All parameters for which the uncertainty is disregarded in probabilistic model calculations should be clearly identified. No parameter should implicitly be assumed as fix. This will allow the reviewers of a PA study to follow the line of arguments and make their own decisions.

4.2 Selection and assessment of a knowledge base

It is essential for a traceable process of uncertainty quantification that the knowledge base considered is clearly defined. That does not imply that all available knowledge has to be taken into account, there may be good reasons to exclude specific data sources from the process. It should, however, be clearly stated, on which sources the considerations are based and for which reasons particular sources have been excluded.

Each individual set of information about a parameter value is categorised first in one of four quality levels:

Level 3: Direct measurements from in-situ or laboratory investigations,

Level 2: An accepted model representation, based on qualified concepts and data,

Level 1: Analogy considerations from related parameters, based on plausible assumptions,

Level 0: Plausibility limits, which should be the least that always exists.

The quality levels given here are not intended to classify data or sets of information according to their uncertainty. The purpose of the scheme is rather to create an order of priority for use of the different sources in the uncertainty quantification process.

Often, various sets of information about a parameter value will be available, which may be of different quality level, depending on their origin. Data sets with a quality level lower than the highest available quality level may be used to substantiate the knowledge base. Generally it is recommended to include initially as many different and independent data sets in the assessment as possible.

Sometimes different sources of information are available that belong to the same quality level and should all be taken into account. In this case one has to check first whether these sets are independent – otherwise they have to be considered as being one source – and comparable. If there are more than one independent and comparable sources of information of the same quality level, one has to check whether they confirm each other. If they do, the different data sets can be combined directly into one data set. Otherwise, one should try to judge the reliability of the different sets against each other using expert knowledge. According to the results of this expert judgement one should weight the different data sources when combining the different data sets into one. If it is not possible to perform a qualified expert judgement, all sets of information should be weighted equally.

It is recommended always to use information of the two highest quality levels that are available. Adding the level values uniquely leads to one of the following cases:

5 – A set of measured data supported by models

4 – A set of measured data supported by analogies

3 – A model representation supported by analogies

2 – A model representation supported by plausibility limits

1 – An analogy consideration supported by plausibility limits

0 – A plausibility interval

The general procedure for assessing the quality of information is schematically shown in Figure 5.

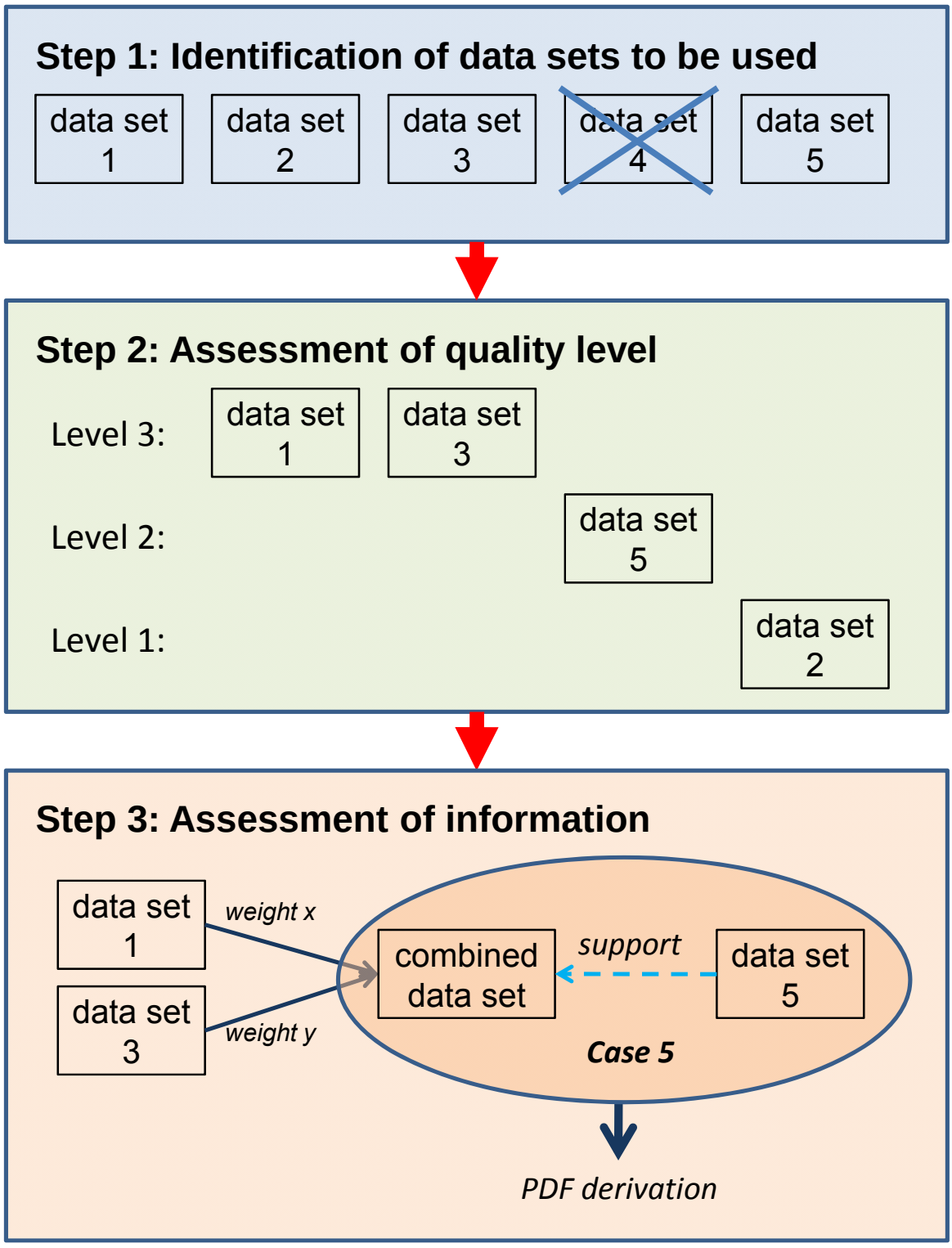


Figure 5 Sketch of general procedure

4.3 Assessment of uncertainty

If it turns out that newly available data or pieces of information do not fit in a previously defined uncertainty interval, this reveals a misassessment. A consistent derivation of a PDF implies that the range and width in the statistical representation of the parameter uncertainty remains constant or decreases when additional information concerning this parameter becomes available at a later stage. In practice, however, consistency of a procedure can never be guaranteed, and proven only with hindsight.

Each of the six cases resulting from the information assessment scheme given above requires its own specific procedure for identifying the most suitable PDF.

Case 0: If a plausibility interval is all that is known, the PDF must be a uniform distribution between the interval bounds.

Case 1: If an analogy is supported by plausibility limits 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: If a model representation for the parameter under consideration exists, it should always allow the derivation of a PDF. For calibration of the PDF only the plausibility limits can be used.

Case 3: This case is very similar to the previous one, as a model representation exists. The calibration, however, can be based on analogies instead of simple plausibility.

Case 4: Though looking like a comfortable situation, this case can become problematic and include considerable subjectivity. The PDF cannot be derived theoretically but has to be derived from the measured data. First, it has to be decided whether the amount of data is sufficient to derive a PDF. If this is the case, the PDF can be derived from the data and calibrated with the analogue. If, however, the measured data do not suffice to derive a PDF one has to follow a different approach. Provided that there is enough information from the analogue to allow a more precise derivation of a PDF, one can transfer this PDF to the actual parameter, calibrating it with the measured data. If even that is not possible, one should perform an expert judgement of the measured data and 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: If measured data are supported by a model representation, 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, one should perform an expert judgement of both and decide which is more reliable. In case of doubt 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.

4.4 Determination of a PDF from given data

In some cases the PDF has to be derived from a set of given data points. If this set is sufficiently large, the PDF can be read off directly. If the number n of data points is five or more, but not enough for the PDF to be obvious, one should perform a statistical analysis. It is recommended that in such cases only (log-)uniform or (log-)normal distributions be used. First it has to be decided whether the PDF should be defined on a linear or a logarithmic scale. This can be done by calculating the arithmetic and the geometric mean of the values and comparing them with the median. If the median is closer to the geometric mean, a logarithmic distribution should be used. The procedure described in the following should then be applied to the logarithms of the values.

If the given data can be represented sufficiently well by a uniform distribution, it should be preferred, even if a normal distribution would also be possible. Therefore, one should apply a statistical test for the null hypothesis that the data follow a uniform distribution. The null hypothesis can be checked using the Kolmogorov-Smirnov-test by determining the maximum absolute deviation of the actual from the theoretical CDF [Conover, 1999]. If the test fails a second test should be performed with a normal distribution. The parameters of the test distribution are taken directly from the mean and standard deviation of the data set under investigation. If, however, also the hypothesis of a normal distribution has to be rejected, it is recommended to use a uniform distribution, regardless of its rejection in the first place.

5. Updating PDFs with Bayesian methods (Facilia)

5.1 Introduction

This section discusses methods to derive probability density functions (PDFs) of a model parameter for which some *a priori* information, as well as experimental data exists. The methods rely on Bayesian inference (formalized by Bayes' theorem). Bayesian inference is the process of fitting a probability model to a set of data and summarizing the result by a probability distribution for the parameters of the probability model. The essential characteristic of Bayesian methods is their explicit use of probability for quantifying uncertainty in model parameters.

Three different Bayesian methods to update PDFs of model parameters are presented here; each having its own field of application.

i) Direct updating – used when the a priori information is a generic PDF or the PDF of an analogue to the parameter of interest.

ii) Hierarchical updating – Used when PDFs of several analogues to the parameter of interest are available.

iii) Updating of model parameters, for example the coefficients of a regression model that relates the parameter of interest to other measurable quantities.

These methods are described in section 5.2 and examples of their application are presented in section 5.3.

5.2 Description of the methods

The main equation used to update a prior PDF with observed data is based on the Bayes' theorem:

$$p(\theta \mid y) = \frac{p(y \mid \theta) \cdot p(\theta)}{\int_{\theta} p(y \mid \theta) \cdot p(\theta) d\theta} \tag{1}$$

where $p(y \mid \theta)$ is the PDF of the observations depending on unknown parameter or parameter vector θ (posterior PDF) and p(θ) is the a priori PDF of the parameter θ (a priori PDF).

The denominator of in equation 1 is equal to p(y) and depends only on known data. It is thus a constant normalization factor which can be removed yielding the proportionality:

$$p(\theta \mid y) \propto p(y \mid \theta) \cdot p(\theta) \quad (2)$$

Below follows a brief description of the principles behind three different Bayesian inference methods that can be used to derive the posterior distribution from the a priori distribution and new observations. These methods were implemented in a specialized tool (Babar) that uses the BUGS language.

5.2.1 Direct updating

In some situations, it is possible to make a relatively good guess of the type of distribution that a parameter should follow. For example, it is commonly assumed that radioecological parameters, such as transfer factors and distribution coefficients, follow a lognormal distribution. This assumption has not only been supported by empirical observations, but also has been justified theoretically by invoking the multiplicative version of the Central Limit Theorem [Ott, 1990]. Once the type of distribution has been established, it is still necessary to estimate the distribution parameters, e.g. the mean and standard deviation for the case of normal distribution or the geometric mean and geometric standard deviation for a log normal distribution. When sufficient empirical data is available, the distribution parameters can be directly estimated using fitting techniques, such as the Maximum Likelihood Estimates. However, often sufficient representative data are lacking, which makes the application of fitting techniques impossible or ambiguous.

In situations when there is no sufficient representative data, Bayesian inference offers a possibility for deriving probability distributions that integrates a priori knowledge from generic data with empirical data obtained for the specific situation of interest. Analytical and numerical methods have been described in the literature [Gelman et al., 2003] for updating probability distributions.

The direct updating method combines a prior PDF with specific empirical observations. Since a normal or lognormal shape will be assumed, sufficient statistics can be used to summarize the prior PDF and the observed data. The sufficient statistics of a normal distribution is the sample size, mean and standard deviation and the corresponding sufficient statistics of a lognormal distribution are the sample size, geometric mean and geometric standard deviation.

The data model

The n empirical observations y_i, i=1,…,n are here assumed to be normally distributed (if instead a lognormal distribution is more appropriate, it is sufficient to apply the model on the logarithm of the observations). The data model is represented as:

$$y_i \sim N(\mu, \sigma^2), \; i = 1,...,n$$

where N denotes the normal distribution with mean μ and variance σ^2. The joint PDF of all n samples, assuming independence, is obtained as multiplication of the individual n normal likelihoods:

$$p(y \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \prod_i^n \exp(-\frac{1}{2\sigma^2}(y_i - \mu)^2) \tag{3}$$

where y (without subscript) denotes the vector of all n samples.

Prior distributions

Given a priori information summarized with mean μ_0, variance σ_0^2 and number of data points n_0 the prior distributions are defined as:

$$\mu \mid \sigma^2 \sim N(\mu_0, \sigma^2)$$
$$\sigma^2 \sim Inv - \chi^2(n_0 - 1, \sigma_0^2)$$

where N denotes the Normal distribution and Inv-χ^2 the Scaled Inverse Chi Square distribution.

The joint prior distribution is given by:

$$p(\mu, \sigma^2) \propto \sigma^{-1}(\sigma^2)^{-((n_0-1)/2+1)} \exp(-\frac{1}{2\sigma^2}[(n_0 - 1)\sigma_0^2 + n_0(\mu - \mu_0)^2] \times$$
$$\times (\sigma^2)^{-n/2} \exp(-\frac{1}{2\sigma^2}[(n-1)s^2 + n(\bar{y} - \mu)^2]) \tag{4}$$

and it is called a Normal Inverse Chi Squared distribution, commonly labelled as

$$Normal - Inv - \chi^2(\mu_0, \frac{\sigma_0^2}{n_0}; n_0 - 1, \sigma_0^2)$$

The posterior distribution

The joint posterior distribution is given as multiplication of the joint prior (equation 4) and the data likelihood (equation 3) and is also Inverse Chi Squared distributed but with updated parameters:

$$Normal - Inv - \chi^2(\mu_n, \frac{\sigma_n^2}{n_n}; n_n - 1, \sigma_n^2) \tag{5}$$

$$\mu_n = \frac{1}{n_0 + n}(n_0\mu_0 + n\bar{y})$$
$$n_n = n_0 + n \qquad (6)$$
$$\sigma_n^2 = \frac{1}{n_0 + n - 1}((n_0 - 1)\sigma_0^2 + (n-1)s^2 + \frac{n_0 n}{n_0 + n}(\bar{y} - \mu_0)^2)$$

The joint two-dimensional posterior distribution (equation 5) is of the same form as the joint prior (equation 4) and can be thus be factored as:

$$\mu \mid \sigma^2 \sim N(\mu_n, \sigma^2) \qquad (7)$$

$$\sigma^2 \sim Inv - \chi^2(n_n - 1, \sigma_n^2) \qquad (8)$$

Samples from the joint posterior distribution are most easily obtained by first generating a sample of σ^2 from (equation 8) and then a sample of μ from (equation 7) using the obtained sample for σ^2.

The posterior distribution of μ can be simplified further by noting that a normal distribution with its variance drawn from a Scaled Inverse Chi Square distribution is equal to a 3 parameter student t distribution μ~t(μ_n,σ_n^2,n_n) with location parameter μ_n, scale σ^2/n_n and degrees of freedom n_n .

Interpretation of the updated parameters

The parameters (equation 6) of the posterior distribution combine the prior information and the information contained in the data. The mean μ_n is a weighted average of the prior mean and the data mean, with weighted equal to the number of prior samples and the number of data samples respectively. The variance σ_n^2 is also a weighted combination of the prior variance and the data average but with the additional term estimating the uncertainty given by the difference between the sample mean and the prior mean.

Posterior predicted distribution

The posterior parameters (equations 7 and 8) characterize the distribution of the considered variable y (μ_n and σ_n^2 being the expected mean and variance respectively), but they ignore the error due to limited number of samples n_0+n. To account for this error, a predicted distribution is calculated as follows:

For k=1,…, N:

1. Obtain a set of posterior samples μ_k,σ_k^2 from (equation 7) and (equation 8)
2. Generate predictions $y_k \sim N(\mu_k, \sigma_k^2)$

The posterior predicted distribution can be summarized by for example calculating the mean, standard deviation or percentiles of the obtained samples y_1, y_2, …,y_N.

5.2.2 Hierarchical models

Many assessment cases involve multiple parameters that can be regarded as related or connected in some way, implying that a join probability model of these parameters should reflect the similarities among them. For example, Transfer Factors (TF) within a group of chemical elements may have similarities and can be assessed with a hierarchical model. By doing this, observations of TF for one element will help improving the distributions not only for this element, but also for other elements included in the model and for the whole population represented by the model. This is achieved by modelling observables outcomes conditionally on certain parameters, which themselves are given a probabilistic specification in terms of further parameters, known as hyper-parameters.

The main application of hierarchical models is when data of a parameter, like the TF, are available for several elements with similar properties (for example actinides), but the amount of data is different for different elements. In this case, the PDF estimates will be improved for the elements with less data. Even for elements with no data, which belong to the hierarchical model, a PDF can be generated from a population distribution. This distribution can be updated as soon as the first data point is obtained for the element in question.

The data model

Assume that there are n_j number of normally distributed empirical observations y_{ij}, i=1,…,n in unit j and that there are J units j = 1,…,J. The data model for observation i in unit j is:

$$y_{i,j} \sim N(\mu_j, \sigma^2) \tag{9}$$

The population structure

A population prior structure couples the observations of the individual and units:

$$\mu_j \sim N(\mu, \tau^2)$$

i.e., the individual means are assumed to be independent outcomes from a common population distribution with two hyper parameters, unknown population mean μ and between unit variance τ^2. The hyper parameters are estimated from the hierarchical model from available data but will at the same time function as a prior distribution for the individual unit means μ_j, j = 1,…,J.

The prior structure

Before Bayes' theorem can be applied to the specified data and population structures, prior distributions need to be set on each of unit variance σ^2 and the population parameters μ and τ^2.

Here only non-informative prior distributions are considered for the three parameters, which allow them to be estimated from available data alone. A theoretical non informative prior distribution [Gelman et al., 2003] for unit level variance σ^2 is defined as a uniform distribution of $\log\sigma$, which can be written as $p(\log\sigma) \propto 1$. On the original scale this is equal to $p(\sigma^2) \propto 1/\sigma^2$. Since the BUGS language does not support prior distributions with infinite integrals (so called improper prior distributions), the implementation in this document instead uses σ^2~ Inverse-Gamma (α,β) with "small" α, β as an approximation[1]. The choice of parameters in the implementation discussed here is α = β=0.000001 (this choice of values is extensively used in the BUGS example collection[2]). A non-informative prior distribution for μ is the uniform distribution with infinite bounds. This prior also has an infinite integral so it is instead approximated by a uniform distribution with large, but finite bounds chosen to include all reasonable and probable values of μ.

A non-informative prior distribution for the between unit variance τ^2 is p (τ^2) ∝ τ, which is equal to a uniform distribution on τ. This prior is recommended in [Gelman et al., 2003] for hierarchical variance parameters.

Separate variances

The structure can be extended to allow separate unit variances σ_j^2, j=1,…,J. In this case, the variance of the data structure then includes an additional subscripted index:

$$y_{i,j} \sim N(\mu_j, \sigma_j^2)$$

Prior distributions (informative or non-informative) must be assigned to each one of the J variances.

The posterior distribution

The data model, population structure and prior distributions are combined using Bayes' theorem, but due to the large amount of involved parameters it is non-trivial to sample from the resulting posterior directly. The samples from the posterior can be obtained by using for

[1] For $\alpha = \beta = 0$ the approximation is exact, although this yields an improper distribution.

[2] Available at http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/examples.shtml.

example the WinBUGS software. Alternatively, a so called Gibbs Sampling algorithm can be implemented (see [Gelman et al., 2003] for details).

Posterior predicted distribution of a specific unit

When at least N samples from the posterior distribution have been obtained, a posterior predicted distribution of a given unit can be derived. This distribution can be interpreted as the distribution of probable future outcomes of the parameter of interest.

N samples of the predicted posterior distribution of unit j are obtained as follows:

For k=1,…, N

1. Obtain a set of samples $\mu_{k,j}$, σ_k^2 from the posterior distribution of the mean of unit j and the common variance σ^2 (if separate variances are used, instead draw a sample $\sigma_{k,j}^2$ from the variance of unit j.
2. Generate predicted outcome of unit j as $\tilde{y}_{k,j} \sim N(\mu_{k,j}, \sigma_k^2)$ (or $\tilde{y}_{k,j} \sim N(\mu_{k,j}, \sigma_{k,j}^2)$ if separate variances are used).

The posterior predicted distribution can be summarized by for example calculating the mean, standard deviation or percentiles of the obtained samples y_1, y_2, …y_N.

Posterior predicted distribution of the population

A second type of predicted distribution concerns prediction of outcomes for new units. These can thus be interpreted as the prediction of probable future outcomes from the population.

The posterior population distribution can be predicted if a common variance is estimated, i.e. the data structure (equation 9) is used.

N samples of the predicted posterior distribution of the population are obtained as follows:

For k=1,…,N

1. Obtain a set of samples μ_k, τ_k^2 from the posterior distribution of the population parameters.
2. Obtain a sample for a “new” unit (indicated with asterisk) as $\mu_{k,*} \sim N(\mu_k, \tau_k^2)$ and a posterior sample σ_k^2 of the variance.
3. Generate a predicted outcome for the new unit as $\tilde{y}_{k,*} \sim N(\mu_{k,*}, \sigma_k^2)$.

The posterior predicted distribution can be summarized by for example calculating the mean, standard deviation or percentiles of the obtained samples y_1, y_2,…y_N.

5.2.3 Updating of regression coefficients

The ordinary linear regression model is a widely used statistical tool that relates a parameter of interest (e.g. the distribution coefficient - Kd) with observations of an independent variable (e.g. the pH). In this section, a Bayesian method is used to account for uncertainties in the parameters of the model that incorporates a priori knowledge of the parameters (obtained from literature or previous studies).

The data structure

Consider the ordinary linear regression model:

$$y = \beta_0 + \beta_1 \cdot x_1 + \cdots + \beta_k \cdot x_k + \varepsilon$$
$$\beta_{0,\ldots,k} = \text{coefficients}$$
$$x_{1,\ldots,k} = \text{explanatory variables}$$
$$\varepsilon = \text{measurements error / model error}$$

The error is assumed to be normally distributed with zero Mean and an unknown Standard Deviation σ, i.e. $\varepsilon \sim N(0,\sigma^2)$. The data model can also be written as:

$$y \sim N(\beta_0 + \beta_1 \cdot x_1 + \cdots + \beta_k \cdot x_k, \sigma^2)$$

or more compactly as:

$$y \sim N(X\beta, \sigma^2 I)$$

Where I is the n×n identity matrix, β is the coefficient vector $\beta = (\beta_0,\ldots,\beta_k)$, X is the vector containing the explanatory observations and a column with unit values representing the intercept:

$$X = \begin{pmatrix} 1 & x_{1,1} & \cdots & x_{1,k} \\ 1 & x_{2,1} & \cdots & x_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \cdots & x_{n,k} \end{pmatrix}$$

Using least square estimation a point estimate as well as confidence bounds can be obtained for the coefficients and the residual variance. This section explains how to instead use Bayesian methods to estimate the coefficients and variance. This will allow prior knowledge of the coefficients to be incorporated in the analysis, when new observations are available for y and x.

The prior structure

The prior information can generally be declared using arbitrary probability distributions, but here we are only address the case when conjugate prior distributions are used (a prior distribution is conjugate to the likelihood, if the resulting posterior takes the same functional form as the prior). The advantage of using conjugate prior distributions is that they often yield posterior distributions that are analytically tractable. A conjugate prior distribution for the coefficients is the Normal distribution and a Scaled Inverse Chi Squared distribution for the squared error. This choice is coherent with classical regression analysis where the coefficients are often assumed normally distributed.

The prior structure is thus defined as:

$$\begin{aligned}
&\beta_0 \sim N(\mu_{\beta_0}, \sigma^2_{\beta_0}) \\
&\vdots \\
&\beta_k \sim N(\mu_{\beta_k}, \sigma^2_{\beta_k}) \\
&\sigma^2 \sim Inv-\chi^2(\upsilon_0, S_0^2)
\end{aligned}$$

where $\mu_{\beta 0,\ldots,k}$, $\sigma_{\beta 0,\ldots,k}^2$, v_0, S_0^2 needs to be fitted to the available prior information.

The prior structure for the coefficients can also be written as:

$$\beta \sim N(b_0, B_0)$$

where

$$b_0 = (\mu_{\beta_0}, \ldots, \mu_{\beta_k})$$

$$B_0 = \begin{pmatrix} \sigma^2_{\beta_0} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\beta_1} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2_{\beta k} \end{pmatrix}$$

The posterior distribution

The joint posterior for all parameters is [Gamerman et al., 2006]:

$$p(\beta,\sigma^2 \mid y) \propto (\sigma^2)^{-\frac{n+\nu_0}{2}+1} \exp\left\{-\frac{1}{2\sigma^2}\left[(\beta-\hat{\beta})'X'X(\beta-\hat{\beta}) + SSE + \nu_0 S_0^2\right]\right\} \times$$
$$\times \exp\left\{-\frac{1}{2}\left[(\beta-b_0)'B_0^{-1}(\beta-b_0)\right]\right\}$$

where $SSE = (y - X\hat{\beta})'(y - X\hat{\beta})$ is the sum of squared errors and $\hat{\beta} = (X'X)^{-1}X'y$ the least square estimates of the coefficients.

The joint posterior distribution is difficult to sample from directly (because it is not recognized as any standard PDF). A Gibbs Sampler can however be implemented using the so called conditional posterior distributions $p(\beta,\sigma^2,y)$ and $p(\sigma^2 \mid \beta,y)$ as explained below (see [Gamerman et al., 2006] for details of the derivation):

The conditional posterior distribution for the coefficient vector is:

$$p(\beta \mid \sigma^2, y) \propto \exp\left\{-\frac{1}{2}\left[\sigma^2(\beta-\hat{\beta})'X'X(\beta-\hat{\beta}) + (\beta-b_0)'B_0^{-1}(\beta-b_0)\right]\right\} =$$
$$= \exp\left\{-\frac{1}{2}\left[(\beta-b_1)'B_1^{-1}(\beta-b_1)\right]\right\}$$

where

$$b_1 = B_1\left(B_0^{-1}b_0 + \frac{X'y}{\sigma^2}\right)$$

$$B_1^{-1} = B_0^{-1} + \frac{X'X}{\sigma^2}$$

The expression for the conditional posterior for β is recognized as a Normal (b_1,B_1) distribution.

The conditional posterior distribution for the residual variance is:

$$p(\sigma^2 \mid \beta, y) \propto (\sigma^2)^{-\frac{n+\nu_0}{2}+1} \exp\left\{-\frac{1}{2\sigma^2}\left[(\beta-\hat{\beta})'X'X(\beta-\hat{\beta}) + SSE + \nu_0 S_0\right]\right\} \propto$$
$$\propto (\sigma^2)^{-\frac{n+\nu_0}{2}+1} \exp\left\{-\frac{\nu_1 S_1}{2\sigma^2}\right\}$$

where

$$\nu_1 = \nu_0 + n$$
$$S_1 = (\beta - \hat{\beta})' X'X (\beta - \hat{\beta}) + SSE + \nu_0 S_0$$

and is recognized a $Inv - \chi^2(n_1, S_1)$ distribution.

A non-informative prior for σ^2 is obtained by setting ν_0 = 0 in the above expressions, i.e. no prior information is included in the posterior expression for the variance. It allows the residual variance to be estimated from the data alone.

A simple Gibbs sampler is set up as follows (see [Gelman et al., 2003] for a detailed description as well as procedures to check the convergence of the iterative algorithm):

Set β^0, $\sigma^{2,0}$ as start values, for example as point estimates from least square estimation.

For iter = 1,...,n

1. Draw a sample β^k from $p(\beta|\sigma^{2,k-1},y)$
2. Draw a sample $\sigma^{2,k}$ from $p(\sigma^2|\beta^k,y)$

Note that each parameter is drawn conditioned on the values of previously obtained parameters. After n iterations, the first part of the obtained values should be discarded to diminish the impact of the start values.

5.3 Examples of application

This section presents examples of application of the three updating methods presented above. The examples are given for estimation of PDFs of radioecological parameters, such as Concentration Ratios (CR) and Distribution Coefficients (Kd), which are commonly used in biosphere assessment models.

5.3.1 Direct updating

This example concerns the estimation of the PDF for the parameter concentration ratio (CR) from water to fish of chlorine (Cl). Nine values have been observed and an additional generic PDF based on 16 values is also available from the literature. Statistics summarizing the generic PDF and the observed values (specific CR) are shown in Table 6.

Table 6 Statistics summarizing the generic PDF and observed values used in the updating of Concentration Ratios (CR) from water to fish of chlorine (Cl).

Observations	n	GM	GSD	μ	σ^2
Generic CR (prior)	16	0.51	2.2	-0.67	0.62
Specific CR (data)	9	0.14	5.9	-1.97	3.15

The generic PDF defined by the statistics in the first row of Table 6 will be updated with the nine observations summarized in the second row, using the direct updating method described in the previous section.

A lognormal distribution is assumed for the concentration ratio (CR). The data model for observed values of the CR is thus:

$$\log CR \sim N(\mu, \sigma^2)$$

The posterior of the distribution parameters

The statistics in Table 6 has to be converted to logarithmic scale, i.e., GM, GSD must be converted to the corresponding log mean, μ and log variance, σ^2, used to define the updated parameters (equation 2) of the posterior distribution (equation 3). The GM and GSD are directly related to μ and σ^2 according to $\mu = \log GM$ and $\sigma^2 = \log GSD^2$. The transformed values are shown in the last two columns of Table 6.

The parameters of the posterior distribution can readily be calculated using the available observation μ=-1.97,σ^2=3.15, n=9 and prior μ_0=-0.67,σ_0^2=0.62, n_0=16 summaries (see Table 6). The updated parameters are:

$$\mu_n = -1.14$$
$$\sigma_n^2 = 1.84$$
$$n_n = 25$$

The posterior distribution of μ can thus explicitly be written as a t distribution t(μ_n,σ_n^2,n_n) and the posterior distribution for σ^2 as a Scaled-Inv-X^2 (n_n,σ_n^2) distribution. With a degrees of freedom as large as 25 the t distribution is well approximated by a normal distribution.

The posterior predicted distribution of CR

The posterior predicted distribution of CR is calculated from a lognormal distribution with parameters sampled from their posterior distributions following the procedure explained in

section 5.2. Table 7 displays statistics of the posterior PDF of the distribution parameters and of the predicted distribution of CR. Figure 6 displays the generic, data and posterior distributions on logarithmic scale.

Table 7 Statistics of the posterior distributions obtained from direct updating

Parameter	Mean	SD	GM	GSD	2.5 %	50 %	97.5 %
μ	-1.1	0.3			-1.7	-1.1	-0.6
σ^2	2.0	0.6	-	-	1.1	1.9	3.5
CR	0.9	3.4	0.3	4.2	0.02	0.3	5.3

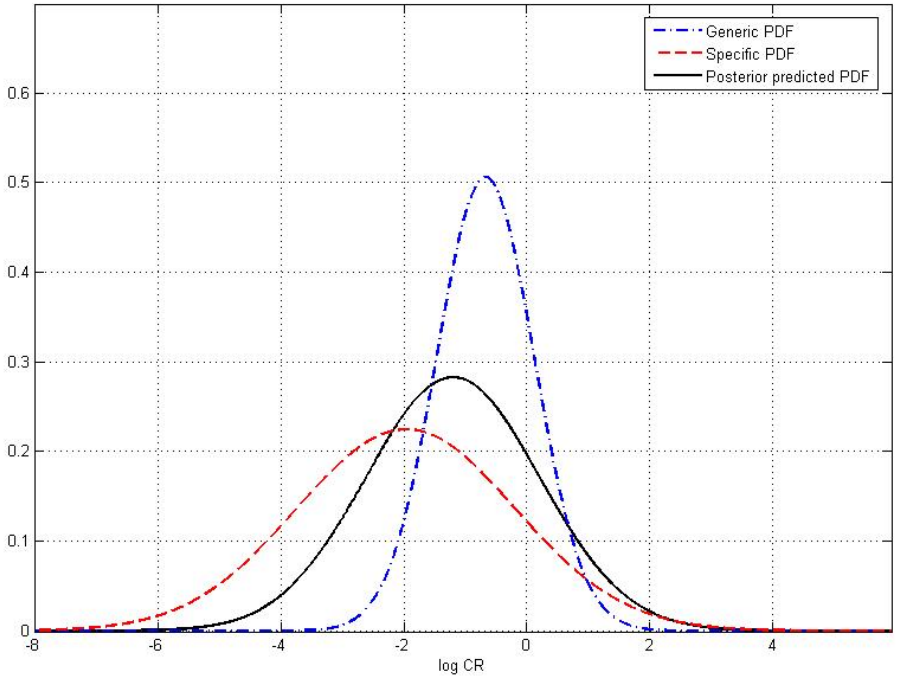


Figure 6 Comparison of the posterior distribution on log scale (normal approximation), observations distribution (a normal PDF fitted to the observed statistics) and the generic distribution (a normal PDF fitted to the generic statistics).

5.3.2 Updating with a hierarchical model

The following example demonstrates how the concentration ratio (CR) of analogue units can be grouped and estimated jointly, giving new estimates for each unit and an estimated distribution for the whole population.

Data of a five species of fish is summarized in Table 8. The species of interest is denoted as reference species and the other species are considered analogue to the reference species.

There are only five data values available for the reference species, which is not sufficient to make a good estimation of the distribution parameters. However, there is good statistics available for the four analogue species. Using a hierarchical model it is possible to derive a distribution for the reference species.

Table 8 Observed statistics of the CR for five species of fish.

Species	n	GM	GSD	μ	σ^2
Reference species	5	100	2.5	4.6	0.92
Analogue 1	308	31	3.1	3.4	1.3
Analogue 2	743	19	4.1	2.9	2.0
Analogue 3	44	90	2.5	4.5	0.84
Analogue 4	22	38	3.2	3.6	1.4

The posterior of the distribution parameters

Table 9 displays summaries of the posterior distribution of the log means (μ) for each of the units. The values of the observed log means (μ) from Table 8 are also included for comparison. Figure 7 displays the mean, median, 95% and 50% posterior probability intervals of the log means (μ). The small vertical line below the bars shows the estimate from observed data. It is evident that the posterior distribution of the reference species is shifted towards lower values of μ compared to the observed value, due the small number of samples (n=5). The large 95% probability interval of the reference unit indicates less precision for units with small number of observations.

Table 9 Mean and SD of the posterior distributions of μ for each of the units. The last row shows the common variance σ^2 estimated from all units. The column "observed" contains the observed values taken from Table 8.

Parameter	Posterior Mean	Posterior SD	Observed μ
μ - reference species	4.3	0.5	4.6
μ -Analogue 1	3.4	0.07	3.4
μ - Analogue 2	2.9	0.05	2.9
μ - Analogue 3	4.5	0.2	4.5
μ - Analogue 4	3.6	0.3	3.6
σ^2	1.7	0.07	0.84-2.0

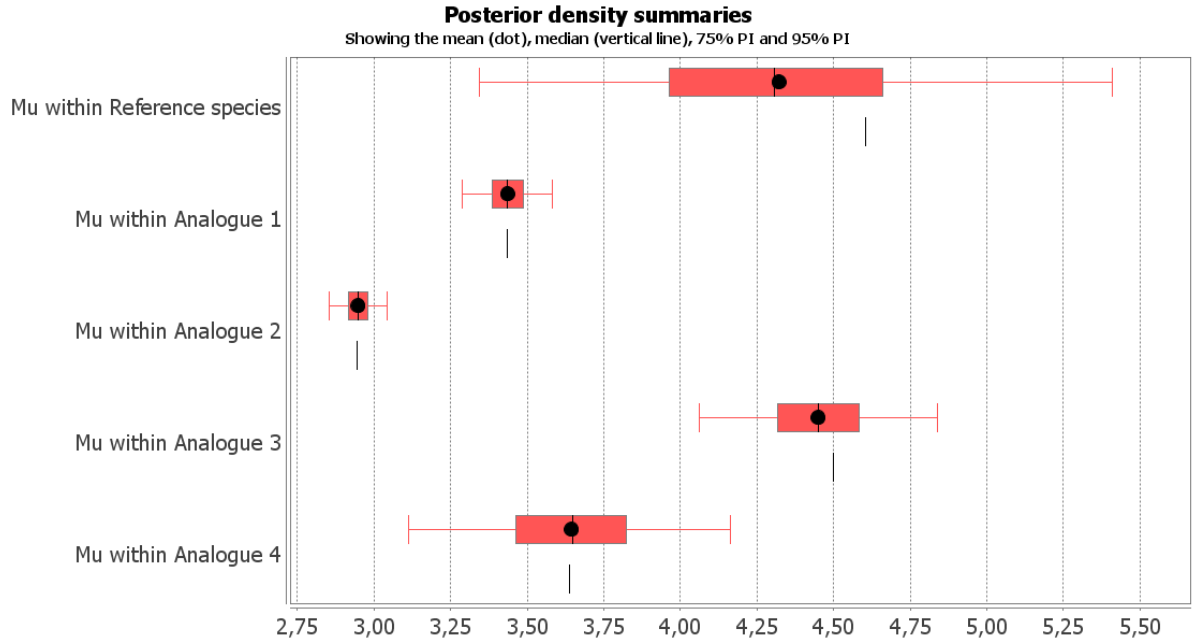


Figure 7 Posterior intervals of the log means (red bars) for different species compared to estimates from individual observations (the separate vertical line below the bars). It is evidient that the estimated log mean of the reference species (the top bar) is lower compared to the observed values. For the other four species, the posteiror estimates are practially the same as the observed values.

The posterior predicted distributions of CR

Given the posterior distributions for all distribution parameters, posterior distributions can be derived for the CR of each individual species, as well as for the whole population. Table 10 shows the GM and GSD of the posterior predicted distributions together with the observed GM and GSD of each unit. The GM of the reference species is estimated as 75 from the hierarchical model, which is clearly lower than the observed GM of 100. For other species the GM are practically the same as the individual observations. Figure 8 shows summarized mean, median, 50% and 95% probability intervals of the logarithm of the posterior predicted distribution of CR for each species; together with the posterior predicted distribution of the whole population.

Table 10 Summaries of the posterior distributions of the CR obtained with the hierarchical model.

Parameter	Posterior GM	Posterior GSD	Observed GM	Observed GSD
CR - population	45	7.1	-	-
CR - reference species	75	4.1	100	2.5
CR -Analogue 1	31	3.7	31	3.1
CR - Analogue 2	19	3.7	19	4.1
CR - Analogue 3	86	3.8	90	2.5
CR - Analogue 4	38	3.8	38	3.2

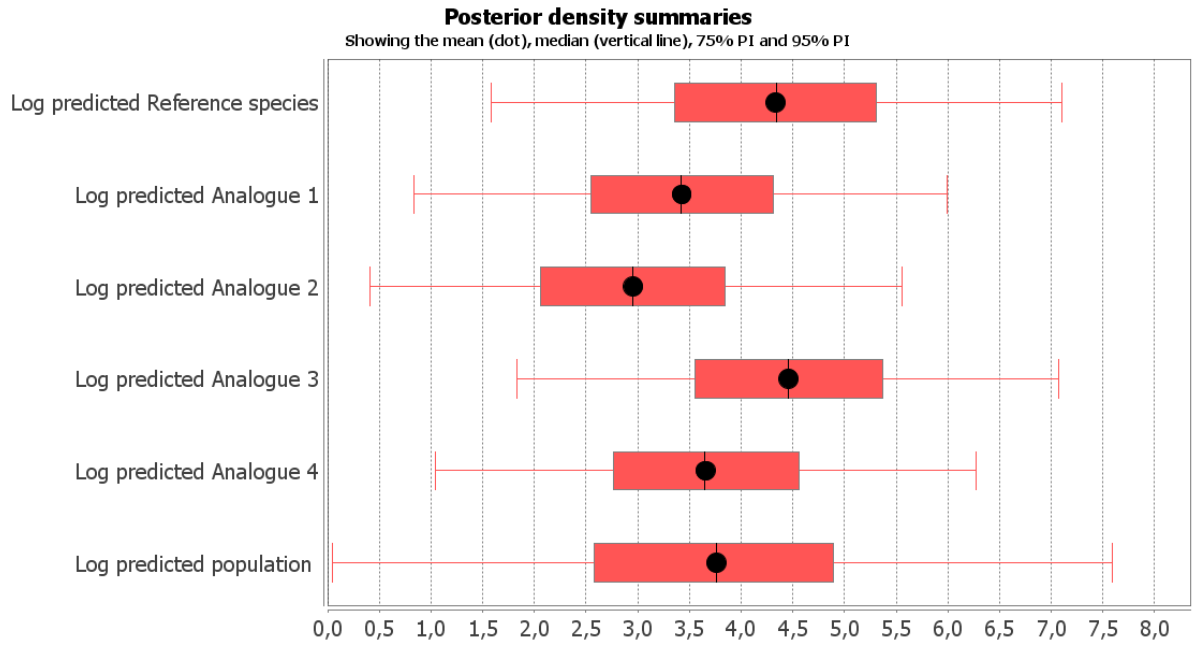


Figure 8 Posterior probability intervals of predictions for each species (first five rows) and the whole population (the last row) – logarithm of the values are presented.

5.3.3 Updating of regression coefficients

The following example demonstrates how the distribution of the parameters of a linear regression model can be updated with new measured values. The model that will be updated is a linear regression model relating the distribution coefficient (Kd) of an element with the pH:

$$\log Kd = a + b \cdot pH + \varepsilon$$

where $\varepsilon \sim N(0,\sigma^2)$ is the random residual error.

Five sets of correlated observations for log Kd and pH shown in Table 11 will be used in the updating procedure.

Table 11 Observations used to update the regression model.

Parameter	Value 1	Value 2	Value 3	Value 4	Value 5
Kd	7.1	8.5	9.2	9.6	7.4
pH	5.2	4.5	5.8	7.0	5.2

The prior distributions

Table 12 contains the mean and standard deviation of the coefficients of the regression model, obtained from previous studies. The available statistics was used to derive the following normal prior distributions:

$$a \sim N(7.3,10.5^2)$$
$$b \sim N(0.21,1.7^2)$$

The prior distribution for the residual error is taken to be non-informative, i.e. by setting $n_0=0$ in the expressions for the posterior distribution.

Table 12 Estimated mean and standard devaition of the regression coefficients used to derive prior distributions.

Parameter	Mean	SD
a	7.3	10.5
b	0.21	1.7

The posterior distributions

The posterior distributions for $\beta = (a\ b)'$ and σ^2 are derived with the Gibbs sampling algorithm, using the conditional posterior distributions (see section 5.2.3) of the regression coefficients. Table 13 displays the posterior summaries after simulating 40 000 values from the posterior distributions. The column SD shows that the posterior uncertainty of the coefficients has decreased significantly compared to the observed SD.

Table 13 Summaries of the posterior distribution of the parameters of the linear regression model.

Parameter	Posterior Mean	Posterior SD	Observed Mean	Observed SD
a	4.8	3.3	7.3	10.5
b	0.6	0.6	0.21	1.7
σ^2	2.0	5.0	-	-

The updated regression model can be used to generate PDFs of the Kd corresponding to specified values or intervals of variation of the explanatory variable (pH). This can be done by performing Monte Carlo simulations with the regression model, where the posterior distributions in Table 13 are assigned to the regression coefficients.

5.3.4 Conclusions

Three different Bayesian methods for updating PDFs of model parameters have been presented and illustrated with examples. These methods can be applied in situations when empirical data is not sufficient to estimate directly the distribution parameters; but there is additional "prior" information available (direct updating). For example there could be information available for a parameter that is analogue to the parameter of interest or for the same parameter measured in an analogue site. The situation when information for several such analogues is available is addressed here using hierarchical models. Such models can be used in a broad sense for pulling together information from several studies of the same parameter or of analogue parameters. The update of the coefficients of a regression model has also been considered. Such models can be used for generating PDFs of a parameter using available information about the range of variation of the explanatory variables.

6. General procedure to treating parameter uncertainty (NRI)

The work of NRI was aimed at devising the general procedure to treating uncertainty in parameters of a computational model used in performance assessments of geological disposal facilities for radioactive waste. The start point was the review of mathematical approaches to model uncertainty [Oberguggenberger, 2005]. The review presented a broad spectrum of approaches which provides the solution for both, the situation when enough data are available to define a statistically representative probability distribution and the situation when only a limited amount of information is available.

A procedure to treating uncertainty consisting of two parallel branches (Figure 9), each running in three steps (1) formalize available information on a parameter uncertainty in an adequate mathematical model, (2) propagate it through model simulations, and (3) express an uncertainty of a computational output was designed. In the spirit of Oberguggenberger's philosophy [Oberguggenberger, 2005], one branch (Figure 9, an orange colored branch) was designed to cover the cases in which the available information on data uncertainty enables to identify a probability distribution, i.e. the frequentist interpretation of probability is applicable. The second branch (Figure 9, a bright green colored branch), on the other hand, covers the cases in which only a sparse set of empirical data is available for the estimation of parameter uncertainty, i.e. the frequentist interpretation of probability is not applicable, parameter uncertainty can be expressed only by subjective measures of uncertainty. In this branch, we decided to model uncertainty with the use of fuzzy set theory. In particular, an uncertain input parameter was described mathematically as a fuzzy real number. This particular choice was motivated by the work of Unit of Engineering Mathematics of University of Innsbruck on the application of fuzzy set theory to modeling the effects of data uncertainties in civil engineering problems [Fetz et al., 2005].

A fuzzy real number is an extension of a real number in the sense that it does not refer to one single value but rather to a connected set of possible values, where each possible value has its own weight between 0 and 1. This weight is called the membership function [Wikipedia, 2009]. A fuzzy real number can be also viewed as a family of nested intervals parameterized by the parameter α [Oberguggenberger, 2005]. The value of this parameter can be interpreted as the degree of possibility that a fuzzy real number takes the value from the corresponding interval A^{α}. The value $\alpha = 1$ represents the highest possibility and the corresponding interval A^{1} contains only an estimated standard value of an input parameter. On the other hand, $\alpha = 0$ represents the lowest possibility, and A^{0} corresponds to the widest interval which covers all possible values of the considered input parameter.

All necessary tools for working with the proposed procedure were built, starting with the assumption that the computational model of interest will be implemented in the GoldSim programming environment (GoldSim Technology Group LLC, USA) which is suitable for the Monte Carlo simulation of the complex system dynamics. Consequently, with respect to the branch based on the frequentist interpretation of probability, it was only necessary to build

the tool for the derivation of probability density function (PDF). The tools for accomplishing the second and the third step of this branch are inherent parts of the GoldSim environment. Contrary, with respect to the second branch based on the fuzzy set theory, it was necessary to implement the tools for performing all three steps.

The developed PDF Derivation Tool is a graphical user interface for fitting PDFs to data which was developed in the MATLAB® programming environment (Version 7.1, the MathWorks, Inc., Natick, Massachusetts, USA) with the help of a collection of statistical tools contained in the Statistics Toolbox (Version 5.0 or higher). The PDF Derivation Tool makes easy to follow a step-wise process for fitting PDF to data which was recommended in [Mishra, 2002]. The process for finding PDF proceeds in five steps:

1. Construct empirical distributions: experiment with multiple bin sizes until a robust indication of PDF shape is obtained.

2. Construct probability plots: select a candidate distribution(s) for fitting.

3. Perform maximum likelihood estimations: estimate parameters of selected candidate distribution(s) and compute goodness-of-fit.

4. Perform nonlinear regression: again, estimate distribution parameters of selected candidate distribution(s) and compute goodness-of-fit.

5. Choose best acceptable fit: if not, then supplement proper PDF.

The transformation method for carrying out fuzzy arithmetic developed by Hanss [Hanss, 2005] was implemented into the computational model of radionuclide transport in the near-field region of a high level nuclear waste repository, which was created in the GoldSim programming environment. For illustration the performance of implementation of the transformation method, we presented the results of the study in which we analysed the influence of uncertainties in two K_d values. These parameters describe the effect of interaction of (i) caesium with bentonite (K_D^{Cs}) and (ii) uranium with bentonite (K_D^{U}), on the uncertainties of ^{135}Cs activity release rate from the near field at different time instants after the time of canister failure (t_K=1000 [yr]). The uncertainties in K_D^{Cs} and K_D^{U} were described, respectively, by the triangular fuzzy number (Figure 10 A) and the log-normal fuzzy number (Figure 10 B), both were estimated using the similar procedure as proposed by Oberguggenberger [Oberguggenberger, 2005]. We performed the three numerical experiments in which the influences of the uncertainties in K_D^{U} and K_D^{Cs} were first considered separately and then jointly. Figure 10 graphically summarizes the results of the latter numerical experiment. Panel C graphically represents the time evaluation of the calculated uncertainties of the release rate of ^{135}Cs activity (R^{Cs}). The width of the surrounding coloured

tail identifies the range of possible values of R^{Cs}, each colour represents a certain degree of possibility.

It was confirmed that it is practically feasible to treat uncertainties in parameters of a complex computational model by the fuzzy set theory.

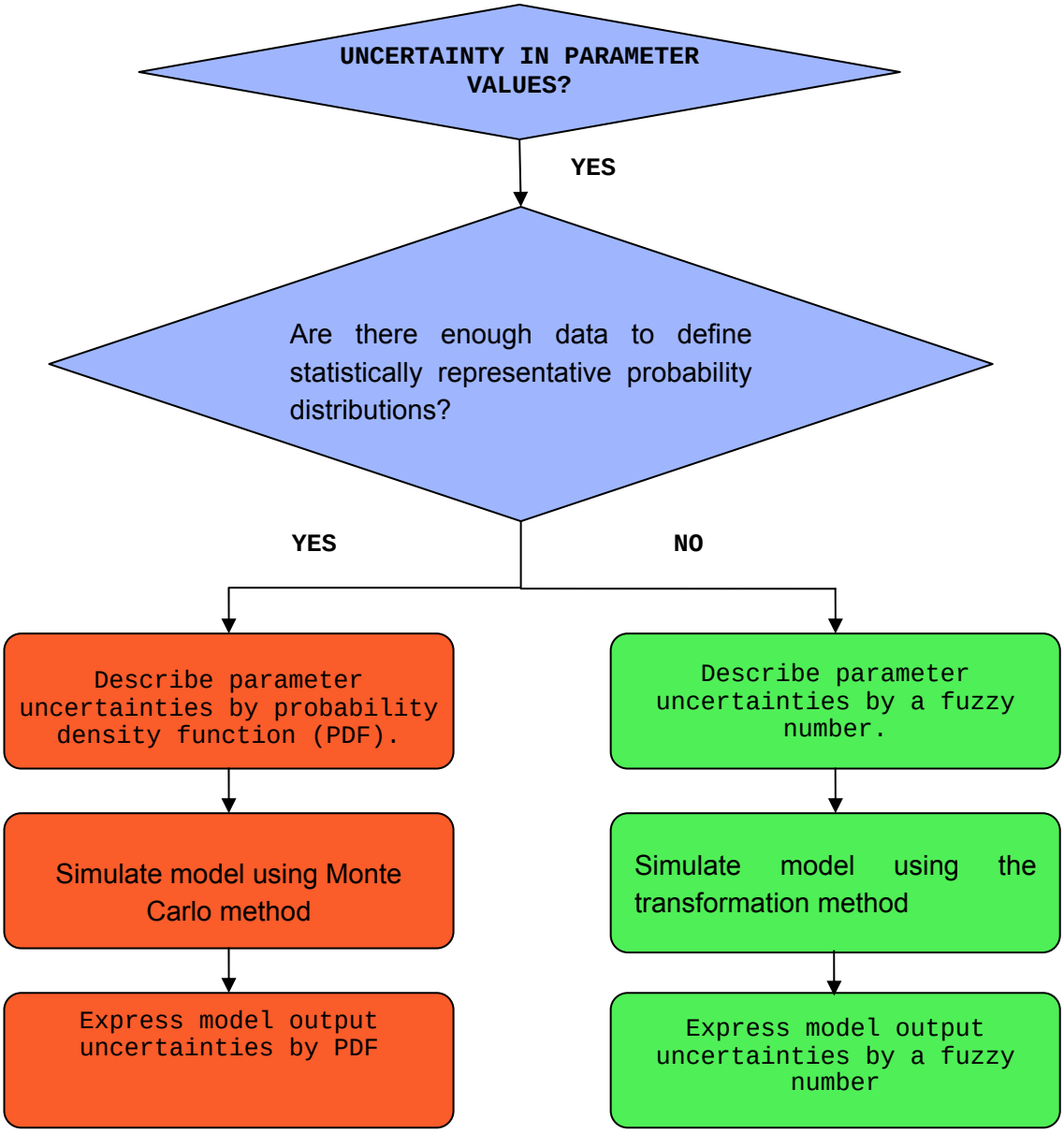


Figure 9 Proposed procedure to treating uncertainty in parameters of a computational model.

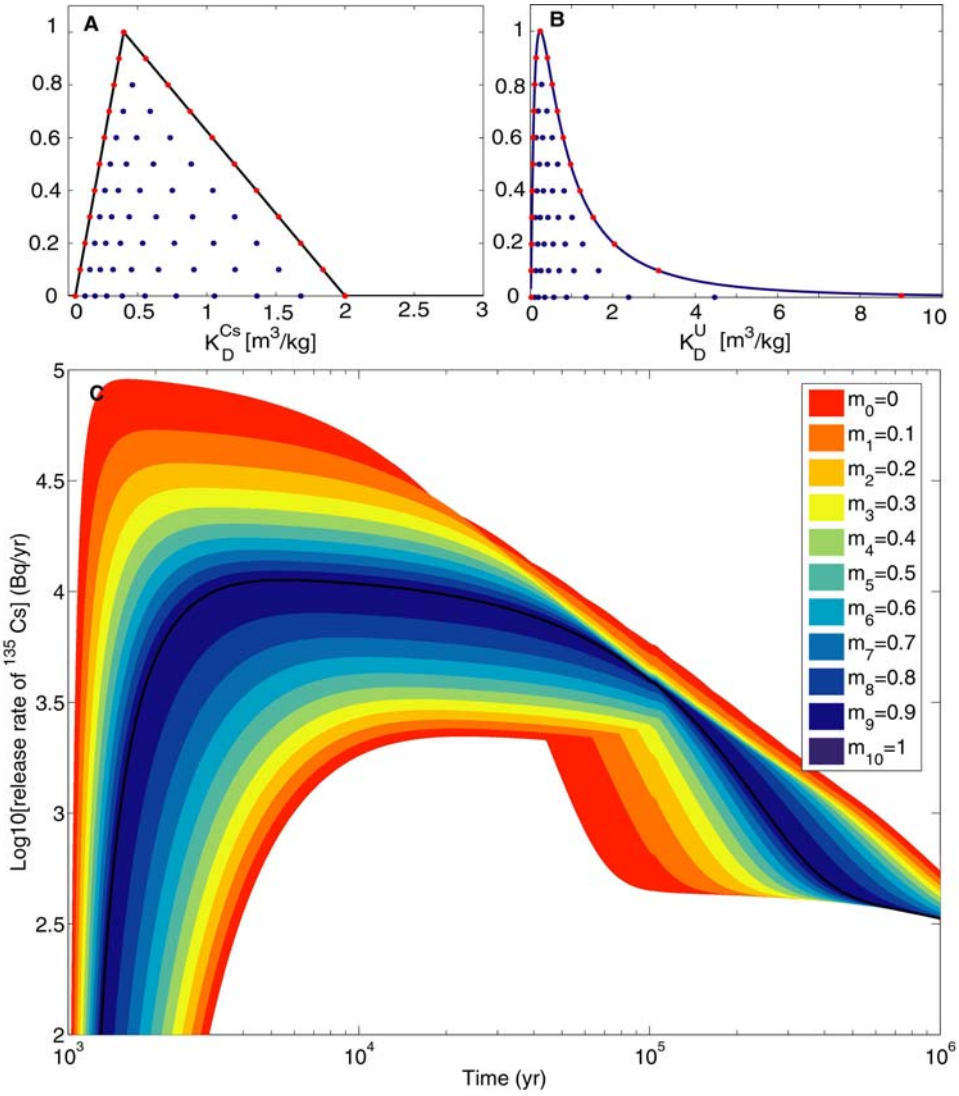


Figure 10 Influence of the uncertainties of the model input parameters K_D^U and K_D^{Cs} on the uncertainties of the model output ^{135}Cs activity release. **A**: K_D^{Cs} was represented by a triangular fuzzy number. **B**: K_D^U was represented by a log-normal fuzzy number. **C**: time evaluation of fuzzy-valued release rate of ^{135}Cs activity, m_i denotes the discrete values of the membership function (see text).

7. Expert judgement for assessing input parameter PDFs (JRC/ENRESA/Amphos21)

Many input parameters used in PA studies are affected by epistemic (lack of knowledge) uncertainty. When confronting this type of uncertainty, experts in the corresponding area are needed to provide their judgements (estimate PDFs for the corresponding uncertain parameters). The key question to answer is: should their opinions/judgements be given in an implicit and informal way or rather in an explicit and formal way?

Scholars in the area of Knowledge Psychology have identified the existence of different inferential mechanisms, typically used under conditions of uncertainty, which can be misused introducing biases in judgements. As a consequence, judgements given by experts can be inconsistent with the information they are based on, not providing an accurate picture of the real expert's uncertainty. The need of feeding computational models used in PA studies with high quality data indicate that some structured procedures have to be developed to minimise the bias in PDFs provided by experts.

A research group within PAMINA has reviewed the vast literature in the area of Expert Judgement (EJ), see [Bolado et al., 2009b], designed a protocol to be used in the assessment of PDFs for uncertain parameters and probability of uncertain events (scenarios), see [Bolado, 2008] and applied it to estimate the solubility limit of five radionuclides (Radium, Tin, Selenium, Uranium and Plutonium) in the near field of the Spanish HLW reference concept in a granite geological formation, see [Bolado et al., 2009b]. Next two subsections describe the protocol developed and the protocol implementation and results obtained for the aforementioned solubility limits. The last subsection is dedicated to lessons learnt.

7.1 An expert judgement protocol

The protocol proposed is based on the SNL/NUREG-1150 protocol, see [Bolado et al., 2009b], adapted to improve expert training and to be applied to small scale applications, i.e. a few variables or events. It has the following phases, which involve three meetings:

1. Selection of the project team

The project team consists of the analyst and the generalists. The analyst is in charge of managing the whole process. He/she should have a deep knowledge in probability, statistics, knowledge psychology and techniques to get opinions from experts. He/she should also be skilful at interacting with people and at moderating group discussions.

The generalists have to provide support to the analyst in all matters related to the questions to be assessed, as for example the decomposition of the question in sub-questions and identifying and obtaining information from different sources. Generalists should have general knowledge about the area under study, not being necessarily top researchers in it.

Generalists are key persons that know the needs of information in the project where the distributions to be obtained will be used. They know well what is needed, why it is needed, the way it will be used and the benefits of getting it in the most accurate possible way. They should also have some management capabilities and be skilful at interacting with experts.

2. Definition of questions to be studied and preparation of support material

The analyst and the generalists must define the questions to be evaluated by the experts. The organisation interested in performing the study does also play an important role in this step; it has to provide the list of questions to be studied. The starting point for any question to be solved is usually vague. It is necessary to arrive at a complete definition of the parameters whose uncertain has to be characterised. Complete definition of a parameter means the full definition of the parameter, including all the problem hypotheses. The final definition should be very clear and accurate, no ambiguity should be allowed. Additionally, if deemed necessary, analysts and generalist should generate any required supporting material.

3. Selection of experts

Qualified Experts are those that: 1)Have the necessary knowledge and experience to perform the assessment, 2) are willing and available to participate in the assessment, and 3) do not have important motivational biases (biases related to ideological and economic attitudes, among others). The selection must be done based on scientific reputation, experimental experience, quantity and quality of publications, among other criteria.

4. Training sessions (first meeting)

The objective of this phase is to show experts normative aspects of EJ elicitation processes. The training sessions are held during the first meeting. The following sub-objectives are pursued, which have associated one training session each:

- Motivate experts to provide rigorous assessments.
- Remember basic concepts of Probability and Statistics.
- Inform experts about basic issues related to cognitive biases.
- Show experts the kind of statistical support they may expect from the project team.
- Training in the assessment of Bayesian probabilities.

5. Refinement of the questions to be studied (first meeting)

The target of this phase is to make all implicit assumptions as explicit as possible, in order to uncover and avoid potential different interpretations of the parameters under study. Generalists present all the supporting material collected together with the questions to be asked. Experts are encouraged to give their own view of the problem and of the definition of the parameters, pointing out, if needed, further information sources, computations to be made, and possible changes in the definition of the questions.

6. First individual work period

After setting the refined questions to be assessed, the experts have roughly three or four weeks to start analyzing the problem individually. They are allowed to use whatever means they consider necessary to address the questions under study. During this period they may contact the project team to get any support in the areas addressed in the training sessions.

7. Presentation of individual approaches adopted by the experts (second meeting)

At the end of the first individual work period, all experts are called to a meeting where each one has to explain to the project team and the rest of the experts the strategy followed to address the questions under study. This is a one-day meeting that starts with a short presentation given by one of the generalists to remind the questions. Afterwards, the floor is given to the experts. Each expert gives a presentation where he/she explains the hypotheses taken into account, the sources of information used, and the decomposition of the problem considered, computer runs performed, and any other information important to understand his/her approach to the problem. The format of this session is the same as a scientific conference with a moderator (the analyst supported by both generalists). Two ideas are behind this session: getting some cross-fertilization among experts and uncover implicit hypotheses that did not come up during the refinement phase.

8. Second individual work period

During a period of roughly one month experts develop their studies taking into account all the information collected along the whole process. As during the first individual work period each expert may use whatever mean he/she wishes and may contact the project team to get support. At the end of this phase each expert has to release a short report summarizing his/her final strategy to solve the problem and whatever information they consider important to understand the way they address the problem.

9. Elicitation sessions (third meeting)

Elicitation sessions are celebrated on individual basis; each expert is elicited separately. All project team members attend each elicitation session. The analyst coordinates the session and the generalists help him providing information when requested. The two main techniques used to assess distributions are the quantile technique and the interval technique. Nevertheless, experts are allowed to give their probability estimates in the format they prefer.

10. Analysis and aggregation of results

After getting the distributions provided by each expert, if no fundamental disagreement exists (acceptable overlap of distributions), distributions are combined mathematically (linear pool). If there are important disagreements, a reconciliation session will be prepared in order to make explicit sources of disagreement and try to reach some convergence.

11. Review

At the end of the process, after getting all individual and aggregated distribution, the project team will generate the report that summarises all the activities developed and results obtained during the process (next step, documentation). This report will be sent to each expert for review. The objective of this phase is to check that experts agree with the way their individual opinions are described in the report.

12. Documentation

Documentation must be as complete as possible, including results and description of the ways to obtain them. The contents of the documentation will follow the order of application of the procedure, recording, in each step, what has been done, why it has been done, how it has been done and who has done it. It ought to be always completely clear to the reader what is a result assessed by an expert and what results are the outcome of an aggregation, sensitivity analysis or any other analysis not provided explicitly by an expert.

7.2 Application case

This protocol was implemented during the second half of 2008 to get pdfs characterising the solubility limit of five radionuclides (Radium, Tin, Selenium, Uranium and Plutonium) in the near field of the Spanish HLW reference concept in a granite geological formation. The project team consisted of one analyst and two generalists. The first generalist was an expert in geochemistry, while the second one was an expert in PA. Four documents were prepared as supporting material to deliver to the experts participating in the process. The first document, [Bolado, 2008], contained the description of the EJ protocol to be applied, the second document, [Bolado et al., 2009b], is a review of EJ techniques and protocols, the third document, [Cormenzana, 2008], is a description of the disposal system considered and the role of solubilities, and the fourth document, [Duro, 2008], is a literature review on radionuclide solubilities. Document 3 contained also a preliminary definition of the questions to be asked to the experts.

Three experts were selected after following the selection procedure: M. Grivé (expert 1), D. Bennett (expert 2) and K. Ollila (expert 3). Unfortunately expert 3 stopped her collaboration before the celebration of the elicitation session; eventually only the distributions provided by experts 1 and 2 were available.

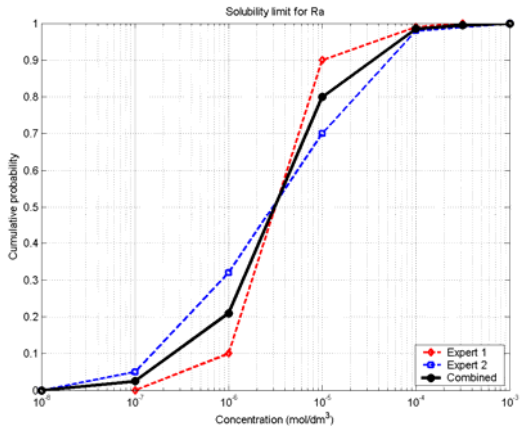
By mid August 2008 experts were invited to participate in the project and got the four support documents. Then the project team convened the first protocol meeting on September 15th and 16th, 2008. Several oral presentations were given about the support material (system description, EJ protocols and compiled information about solubilities). Additionally, experts attended all the training sessions and participated in a calibration session. They could realise the real risk of making overconfident probability statements. Moreover, two sessions were dedicated to debate about the preliminary definition of the questions. At the end of the

meeting an updated definition of the questions was set. The main difference with the original definition was that assumptions/hypotheses were more accurately defined. The project team convened next meeting on October 16th. Between this meeting and the second meeting experts started studying the problem, but they were not asked to solve the problem yet. They were asked to think in depth about the approach that they would adopt, paying special attention to hypotheses that should be taken into account.

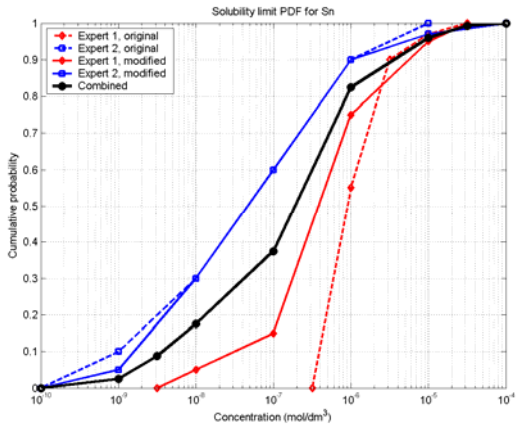
The whole second meeting (a one-day meeting) was dedicated to discuss about the question definition. The outcome of the second meeting was the final question definition: '*Provide probability distribution functions (PDFs) that represent the concentration (solubility) limit of five different chemical elements (Radium, Tin, Selenium, Uranium and Plutonium) under the following near field (iron canister / bentonite buffer interface) conditions* '

- T between 30-40°C, although calculations done at T= 25ºC (justify).
- H2(g) P between 20 and 80 atm, although calculations done at P = 1 atm (justify).
- I corrections: Extended D-H (justify).
- The concentration of carbonate given in table of groundwater composition [Cormenzana, 2008] refers to total carbonate concentration.
- The groundwater used during the assessment should be the one resulting from equilibrating the 100,000 year old bentonite porewater with H2(g) of 20 or 80 atm and magnetite. The Eh, pH and [Fe] from the previous calculations are taken as the reference values for these parameters in the resulting groundwater composition
- Take as central pH value the one for the resulting GW from the previous point and consider an uncertainty of ±1 unit.
- No microbial activity considered, so neither sulphate nor carbonate reduction is considered.
- Calculations will be run with the 5 elements at a time but it will be shown that the inclusion of other elements is (or not) important.
- No degradation of bentonite is considered
- No co-precipitation of a given radionuclide with a major element is considered
- FeSe/FeSe2 might be formed
- Calcite are allowed to form if oversaturated
- No colloidal forms are considered, only true dissolved elements

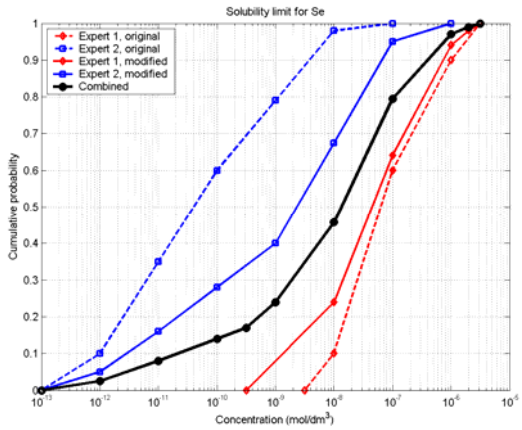
Taking into account this final definition of the problem to solve, experts had roughly six weeks to perform all the study. The elicitation sessions meeting (third and last meeting) was celebrated on December 2nd to 4th. Ten days before the meeting they sent to the project team a short report summarising their respective approaches.



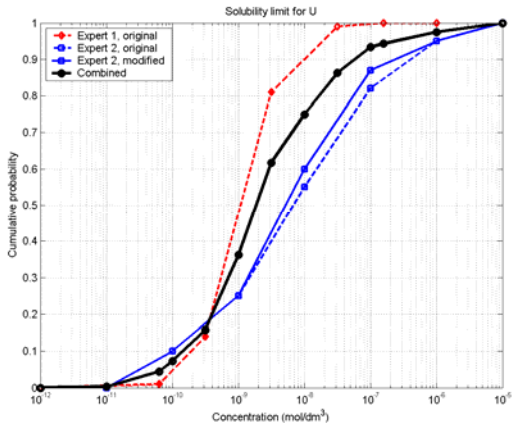
a)



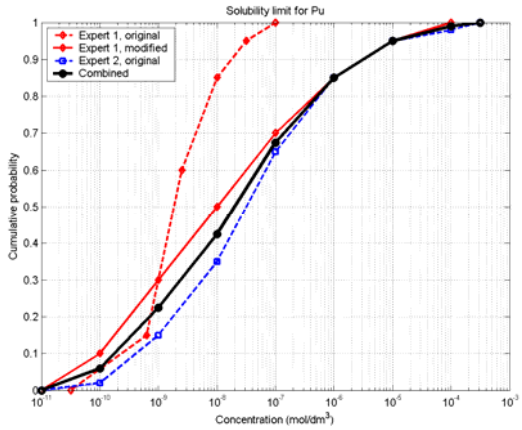
b)



c)



d)



e)

Figure 11 Results of the expert elicitations for the solubility of 5 chemical elements: a) Ra, b) Sn, c) Se, d) U and e) Pu.

Both experts took as a main reference for the study the speciation computations performed with their respective codes (PHREEQC and Medusa in the case of expert 1 and PHREEQC in the case of expert 2) and used different thermodynamic databases, ThermoChimie and Hatches respectively. Both did also take into account experimental data (when available), data from natural waters and data used in other PAs.

The first day of the third meeting was dedicated to elicit expert A's distributions, while the second was dedicated to elicit expert B's distributions. At the end of the second day it was apparent that both experts had some serious disagreements about the distributions assigned to the solubility limit of Sn, and Se, while it was moderate in the case of Pu and even more moderate in the case of U. They had quite a good agreement in the case of Ra. Under these conditions a reconciliation session was mandatory and it was celebrated the last day of the meeting.

The reasons producing the disagreements were uncovered in the reconciliation session (in one case a fundamental disagreement about a problem hypothesis) and this helped getting a remarkable convergence in the distributions provided by both experts for Sn, Se and Pu (specially in this last case). After the reconciliation session (a full working day), the new distributions allowed the mathematical combination in all cases. Figure 11 (a to e) show the original distributions provided by each expert (dashed lines), the distributions after the reconciliation session (solid lines) and the aggregated ones after applying a linear pool (solid thick black lines).

A thorough description of this application may be found in PAMINA milestone M. 2.2.A.12, [Bolado et al., 2009b]

7.3 Lessons learnt

From a methodological point of view the following lessons were learnt in this application case

- Experts found very interesting getting before the actual start of the process the support documents (expert judgement review, protocol description, literature review-datapack and system description).
- Experts quite interesting the training sessions and the calibration session, where they could experience the risk of overconfident statements. Further improvement of these sessions can make more attractive and challenging this type of process to experts.
- The joint refinement (PA owner, project team and experts) of the problem definition was extremely important to avoid implicit hypotheses and misunderstandings. Even being aware of this and dealing with it explicitly in two protocol sessions, a real hypothesis disagreement arose in the reconciliation session.
- Experts found the interval technique most useful. In fact, after the first questions in the elicitation sessions, both experts chose this technique to give their opinions. They found most useful thinking about concentrations in terms of powers of 10.

- The time schedule in the elicitation sessions was really tight. The whole group worked under time pressure, what suggests that we should have been a bit less ambitious, eliciting not more than 4 distributions per day.
- At the beginning of the reconciliation session, each expert had no information about the distributions provided by his/her colleague. This made the schedule of the reconciliation session also very tight. An issue to study in the future is if the reconciliation session (if needed) deserves a specific meeting, which has clear economic effects on the whole process budget.
- Including one more expert in the process would be probably the optimum, in order to get the right balance between diversity of opinions/background and work load.

8. Synthesis

When dealing with parameter uncertainties in probabilistic performance assessment studies for radioactive waste repositories, one is confronted with questions like these:

- Which of the parameter uncertainties are worth of being investigated in a probabilistic study?
- Which influence have the shape and parameters of the selected PDF to the uncertainty of the model results?
- How should one proceed to quantify the existing knowledge and to establish PDFs?
- How should one proceed to improve and assess the knowledge basis by means of expert elicitation?
- How can PDFs be fitted to given data or adapted to improved knowledge?

The different pieces of work presented in this report address these questions individually and complement one another, forming a solid basis for dealing with uncertainties and establishing PDFs.

The work of VTT demonstrates how the influences of parameter uncertainties to the model results can be investigated by means of deterministic parameter variations. These are based on selected values for each of the parameters covering a specific bandwidth but without the need of specifying distribution functions. This kind of local sensitivity analysis is always useful for gaining a deeper system understanding and should be performed as a first step when dealing with parameter uncertainties in a specific study. It is adequate for focussing further attention to those parameters that are really important for the problem under consideration, since it does not make much sense to put big efforts in quantifying the uncertainties of parameters that are actually of low importance for the model results.

As soon as the parameters whose uncertainty can be relevant for the model output are identified, a probabilistic uncertainty and sensitivity analysis using randomly distributed parameter combinations may be taken into account. For this purpose, it is necessary to define PDFs for the uncertain parameters. This, however, can become a difficult task, because it requires the quantification of the degree of nescience. We have to become aware, as exactly as possible, about our knowledge of the parameters, which normally requires literature studies and expert elicitations.

BEL V investigated how the shapes of PDFs influence the results of a probabilistic uncertainty and sensitivity analysis. This is an exemplary study, which, of course, will not be repeated each time a PA is done. It shows, however, that the type of PDF is everything else than unimportant for the results and should therefore not be treated as a secondary issue.

Each type of PDF is characterised by a specific set of statistical parameters, defining, for instance, an interval or the mean and standard deviation. For each uncertain model

parameter, the type of the PDF as well as its statistical parameters should be established in a clear and traceable manner. This is crucial, especially for PDFs used in a safety case or a licensing procedure. This topic might become increasingly important in the future.

GRS proposed a general procedure for establishing PDFs for uncertain parameters. This comprises the type of the PDF as well as its statistical parameters. The procedure is formulated as a protocol that can be worked off step by step, designed to optimise traceability and to minimise subjectivity. It is meant as a guide for the modeller through the process, often characterised by an unclear knowledge situation, urging him to think a bit more about the PDFs and preventing him from simply guessing them. It also provides a help for documentation of the process of PDF derivation.

The proposed procedure needs some input at different stages, which is not defined in detail in the work of GRS. This applies to specific scientific tasks that require some extra effort and should be performed and documented with care. The pieces of work performed by Facilia, NRI and JRC/ENRESA/Amphos21 address such tasks.

Facilia investigated a mathematical method for adapting a PDF based on information from, e.g., a model representation or generic data to an extended set of data, which may result from in-situ measurements or other additional information. The method is based on Bayes' theorem, which generally allows updating a prior distribution to a posterior distribution using additional available information. Several methods for the choice of the prior distribution were tested. The procedure was developed and tested for normal or log-normal distributions, which are often most adequate to describe the uncertainty of physical parameters. It is useful in the process of PDF generation when the decision about the type of the PDF has been made and its statistical parameters have to be established on the basis of different sources of data.

NRI proposed a general mathematical procedure for treating parameter uncertainty on the basis of available data and developed a PDF derivation tool. If there are enough data this tool allows fitting a PDF as well as possible to them, which is an important task in the general procedure of PDF derivation proposed by GRS. If, however, the set of data is too small to allow being processed in this way, a different procedure has to be followed. While for such cases the protocol of GRS proposes a rather simple and stylistic approach to fixing a PDF, NRI considered applying fuzzy arithmetic, which requires a different procedure for model evaluation and analysis of the results. This can be seen as an alternative approach to dealing with uncertain parameters, if the knowledge basis is poor.

An important part of the process of PDF generation is expert elicitation. In order to assess the quality of data, weight different data sets, estimate values or bandwidths, etc. well-funded expert opinions are essential. A good expert judgement, however, can be a very expensive sub-task of the process of PDF generation. Several experts have to be selected and their answers have to be compared and weighted in an adequate manner. This issue was addressed by JRC/ENRESA/Amphos21. A general procedure for expert elicitation was

proposed and tested using the example of solubility limits for different radionuclides. It can be used as an input for the PDF generation procedure proposed by GRS at several stages, although, because of its expensiveness its application will in practice be restricted to really important cases.

The investigations described in this report comprise methodological approaches as well as application studies. Altogether, they show how parameter uncertainty can be handled in probabilistic performance assessment studies and provide a basis for planning and executing a probabilistic uncertainty analysis. In practice, it will neither be possible nor necessary, and not even sensible, to follow the proposed procedures strictly for all parameters, but it is recommended to apply a traceable and well documentable scheme for handling uncertainties instead of simply guessing PDFs.

References

Bolado, R. (2008): An expert judgement protocol to assess the solubility limit distributions for key chemical elements in a generic Spanish repository in granite. PAMINA milestone M2.2.A.4

Bolado, R. Badea, A. And Poole, M. (2009a): Review of expert judgement methods for assigning PDFs. PAMINA milestone M2.2.A.3

Bolado, R., Duro, L., Cormenzana, J.L. and Cuñado, M.A. (2009b): Expert judgement review and exercise. PAMINA milestone M2.2.A.12

Bourgeois, M. (2003): MELODIE, MELO - notice THEorique. IRSN/Département de Protection de l'Environnement, 2003

Conover, W. J. (1999): Practical Nonparametric Statistics. Wiley-VCH 1999

Cormenzana, J.L. (2008): Definition of the repository system and the problem to be solved. Part A: Description of the disposal systemand the role of the solubilities in the performance assessment. PAMINA milestone M2.2.A.5 part A.

Duro, L. (2008): Definition of the repository system and the problem to be solved. Part B: Scientific case and datapack for the expert judgement of the solubilities of Se, U, Sn, Pu and Ra in crystalline media. PAMINA milestone M2.2.A.5 part B

Fetz, T. et al (2005): Fuzzy models in geotechnical engineering and construction management. In Analyzing Uncertainty in Civil Engineering. Springer Berlin Heidelberg. 211-239

Galson, D.A. and Khursheed, A. (2007): The Treatment of Uncertainty in Performance Assessment and Safety Case Development: State-of-the Art Overview. PAMINA WP1.2 Final Report, August 2007.

Gelman et al. (2003): Bayesian Data Analysis, Chapman & Hall/CRC 2003

Gamerman et al. (2006): Markov Chain Monte Carlo Stochastic Simulation for Bayesian Inference, Chapman & Hall/CRC

Hanss, M. (2005): Applied Fuzzy Arithmetic. An Introduction With Engineering Applications. Springer-Verlag Berlin And Heidelberg Gmbh & Co. Kg (Germany).

Mishra, S. (2002): Assigning probability distributions to input parameters of performance assessment models. Technical Report TR-02-11. Swedish Nuclear Fuel and Waste Management Co

Nykyri et al. (2008): Radionuclide release and transport RNT-2008. Posiva Report 2008-06 Posiva Oy, Olkiluoto, Finland

Oberguggenberger, M. (2005): The mathematics of uncertainty: models, methods and interpretations. In Analyzing Uncertainty in Civil Engineering. Springer Berlin Heidelberg, 51-72

ONDRAF/NIRAS (2001): Safety Assessment and Feasibility Interim Report 2, SAFIR 2. Technical report, ONDRAF/NIRAS, 2001.

Ott, R. (1990): A Physical explanation of the log normality of pollutant concentrations, Journal of the Air & Waste Management Association Vol 40, pp 1378-1383

PAMINA (2006): Integrated project Pamina, Performance Assessment Methodologies In Application to guide development of the safety case, annex I description of work. Technical Report FP6-036404, European institutions, 2006.

Saltelli, A. and al. (2008): Global Sensitivity Analysis, The Primer. John Wiley & Sons, Ltd, 2008.

SCK·CEN (1999): Geological disposal of conditioned high-level and long lived radioactive waste, values for the near field and clay parameters used in the performance assessment of the geological disposal of radioactive waste in the Boom clay formation at the Mol site, Data Collection Form. Technical Report R-3344, SCKCEN, 1999

Vieno, T. and Nordman, H. (1999): Safety assessment of spent fuel disposal in Hästholmen, Kivetty, Olkiluoto and Romuvaara, TILA-99. Posiva Oy, Helsinki, Finland. Report POSIVA 99-07

Wikipedia contributors (2009): 'Fuzzy number', Wikipedia, The Free Encyclopedia, 10 May 2009, 16:19 UTC, <http://en.wikipedia.org/w/index.php?title=Fuzzy_number&oldid=289085049> [accessed 10 May 2009]