An Approach of Sensitivity and Uncertainties Analysis Methods Installation in a Safety Calculation

Cédric Sallaberry - *ANDRA*

Service DS/CS – Parc de la Croix Blanche
1/7 rue Jean Monnet – 92298 Châtenay-Malabry Cedex (France)
cedric.sallaberry@andra.fr

Abstract : *In this paper, the ANDRA approach for an installation of sensitivity and uncertainty analysis is described. In a first step, the choices made by ANDRA foreign counterparts are presented and compared. Then, the most classical methods used in sensitivity and uncertainty analysis are listed with their advantages and inconvenient, related to ANDRA problem. Finally the test case used for comparing the methods each other is quickly explained. The conclusion presents the ANDRA point of view about these methods related to the studied problem.*

1. CONTEXT

1.1 Needs of ANDRA

Within the framework of safety assessment of a possible storage of radioactive waste with high activity and long life in major geological formation, simulation allows:

- to account for the various physical phenomena determining the evolution of storage in time and space, by studying the migration of the radioactivity in storage and its environment
- to contribute to evaluate the feasibility and the safety of storage; this evaluation is founded on several kinds of indicators.

The study of the waste storage behavior calls upon several types of phenomena and process, whose level of comprehension is represented with the use of models. These models are then used to carry out a set of analysis, and to simulate the behavior of storage over periods of time and space out of the experiment field. Numerical simulations, implementing the various models selected and describing the physical phenomena brought into play, are intended for :

- evaluating the performances of storage and its components
- evaluating the safety of storage and the impacts that it would be likely to produce
- checking the robustness of the demonstrations compared to the state of knowledge available or accessible.

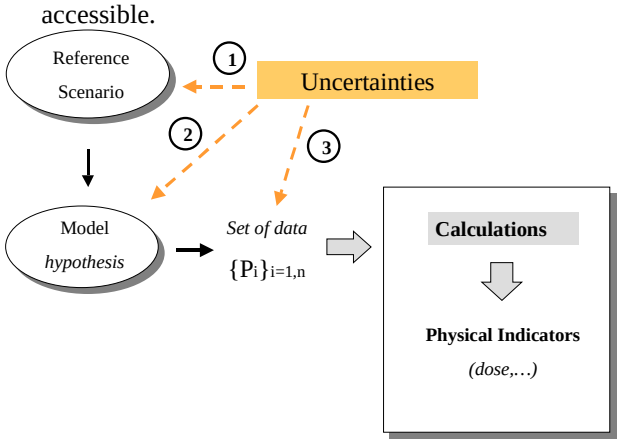

Figure 1 : simplified generic scheme of reference modelization

The figure above synthesizes the generic process of modeling. At the time of the installation of a such methodology, uncertainty can be found on each stage precedent calculation.

Firstable, an uncertainty exists on the evolution scenario : Indeed, one does not know how to describe space and time evolution of the behavior of storage and his environment. It is thus necessary to quantify how these variations can influence storage evolution.

An uncertainty also exists on the models (physical,

mathematical and numerical) selected. The choice carried out on the taking into account of the phenomena, their translation with equations as with algorithm can influence to a significant degree the studied result.

The various models used to describe the relaxation of the radioactivity from the storage to the human comprise many physical parameters, whose value is matched, in the majority of the cases, with an uncertainty being able to be of different nature (uncertainty on a parameter for a material, uncertainty on the choice of a material, uncertainty on the not-catch in account of phenomena, variability). The studies in phenomenologic matter, the experiments which will be carried out within the framework of research in underground laboratory, the bibliographical studies and the experts choices allow to ameliorate the state of knowledge of these parameters (law and field of variation, correlation between parameters.) and if necessary, to reduce uncertainties on the values of parameters or the models used.

This document deals more particularly with uncertainties on the input data of the model. These uncertainties are taken into account in total analysis with the use of uncertainty and sensitivity analysis. ANDRA open a dialog on the treatment of input data uncertainties and their propagation in the models of safety, in order to be able to quantify the influence of input data uncertainties of the models on the various indicators of safety selected.

The objective is to have, in November 2002, suggested the strategy of treatment of uncertainties and to have pre-selected the corresponding computational tools, for tests and implementation made by ANDRA within the framework of the ALLIANCES platform for June 2003. The needs of ANDRA are the definition and the choice of the strategy of the methods and the tools for treatment of uncertainties on the parameters in the safety models.

The uncertainty and sensitivity analyses allow :

- To check the robustness of storage, despite of the lack of knowledge on the value of certain parameters.
- To have a better understanding on the studied system.
- To guide the R&D to improve reliability

1.2 Approach made by ANDRA

The step taken by ANDRA consists initially of two studies undertaken in parallel:

- the first consists of an international review of the choices (and justifications of these choices) retained by ANDRA foreign counterparts to carry out their uncertainty and sensitivity analysis
- the second relates to a review of the various methods being able to be used in sensitivity and uncertainty analysis. Their aptitude to reply to the ANDRA waitings is analyzed.

These studies are supplemented by a comparison of the principal methods on a simplified test case, derived from the conceptual model used for safety assessment calculations in 2001.

2. REVIEW OF THE DIFFERENT CHOICES RETAINED BY OUR FOREIGN COUNTERPARTS

2.1 Countries studied

Our study related to 10 countries: The United States (YMP[1,2] and WIPP[3,4] projects), Canada[5], the United Kingdom[6,7], Belgium[8], Swiss[5], Spain[5], Sweden[9,10], Finland[5], Germany[5] and Japan[11,12]. That represents 11 projects.

Why have these countries been chosen? Three reasons justified our choice:

- These countries undertook a study in the possibility of storage in major geological medium since a certain number of years, and thus have an good idea on the question.
- They lay out a consequent bibliography on the subject, freely available (via the Net or another means)
- Uncertainty and sensitivity analysis are mentioned in part of the bibliographical references.

2.2 Kind of uncertainties taken into account

Three kind of uncertainties are distinguished : uncertainties on the events, uncertainties on the models (conceptual and mathematical) as well as uncertainties on the parameters.

Uncertainties on the events are taken into account by all the studied countries. They use the concept of scenarios in general. The method generally used is deterministic, though certain countries have a total result in probabilistic form by balancing each solution of scenario by a probability occurrence.

Only half of the countries mention to take into account of models uncertainty. These countries use the concept of variation on a scenario. The step used is similar to that of the scenario (deterministic with presence of total result by using weighting).

Parameters uncertainty is considered by all the studied countries. Methodology is very shared, since there are as many country which have chosen a first deterministic approach (Sweden, Switzerland, Finland, Japan) as country which have chose a probabilistic way (the United States, Canada, Germany, United Kingdom). Two countries directly

chose a double deterministic/probabilistic approach (Belgium, Spain). It should however be noted that on the four countries having chosen a deterministic approach, two (Sweden[10] and Japan[12]) are interested with the probabilistic methods (by the use of analytical models or by simplification of the numerical model)

2.3 General Methodology for parameters uncertainty

The required results are always identical: determining which is the probability to reach a certain threshold of amount of radioactivity for human being. The definition of this threshold, as well as the elements taken into account are very variable and specific to each country.

The reasons of the choice between probabilistic and deterministic methods are often identical.

The countries which chose the determinist[9] way underline that :

- In a majority of case, the degree of knowledge on a parameter is too much limited to be able to choose a law. The choice of a law then adds uncertainty on the result.
- The results presented are simple and clear for everyone.
- The number of calculations is restricted

The countries which chose the probabilist[3,4,13] consider the following assertions :

- Determination of the values penalizing for each parameters uses more strongly expert choice, which adds objectivity.
- Nothing indicates, in the deterministic case, that the sum of the values penalizing for each parameters provided the more penalizing case.
- The probabilistic methods allow to obtain a reasonable approximation of the totality of the possible results.
- With probabilistic methods, it is possible to make a sensitivity analysis. This means to analyze the relations between the uncertain input parameters and the output parameter(s) to determine the most influential parameters

2.4 Pre-processing

Deterministic methodology consists of studying two cases : the case considered as the most probable, and the case considered as the most penalizing. The determination of these cases is done by experts opinion and/or experiments return.

The probabilistic case uses the traditional technique of Monte Carlo. Each distribution law associated with a uncertain input parameter is sampled, by using of various methods such as traditional Monte Carlo, Latin Hypercube, Stratified, Importance and Modified Importance Sampling. These samplings lead to a data file, which will be used for each occurrence to simulate a probable evolution of storage.

2.5 Uncertainty Analysis

The various studied countries carry out an uncertainty analysis in order to determine the variation of the output parameter (and the possibility of exceeding a given threshold) by taking account of uncertainties on the input parameters.

As a determinist, the uncertainty analysis is proposed only by one value obtained for the most probable scenario and a value obtained for the more penalizing scenario .

The countries having chosen the probabilistic method use in general the same indicators and curves : mean, standard deviation, median, quantiles (1^{st} and 3^{rd} quartile - 1^{st} and 9^{th} decile or 5^{th} and 95^{th} percentile), cumulative functions of distribution (CFD) or complementary cumulative functions of distribution (CCFD), "hair" or series of curves (one by occurrence)

Certain methods are more rarely used, like the histograms (United Kingdom for example) or box plots (WIPP for example).

2.6 Sensitivity Analysis

Sensitivity is used to quantify which importance input parameters can have on output parameters.

There is no sensitivity analysis carried out in the deterministic case.

In the probabilistic case, the same statistical indicators and charts to carry out a sensitivity analysis can be found. Thus, the majority of the country carrying out a sensitivity analysis using the estimators of 1^{st} (Pearson and Spearman coefficients) and 2^{nd} (PCC/SRC and PRCC/SRRC) order. Certain countries use also 3^{rd} order estimators which are in fact statistical tests (test of Smirnov, Student, Mann-Whitney…) The latter are however regarded as more sensitive and less reliable in their information.

Some countries graphically represent the values of these indicators. The most often found chart in sensitivity analysis is scatter plots. Some other methods, like the histograms or pie chart are frequently found.

2.7 Conclusion

In majority of case (i.e. even if the first approach were deterministic), the countries chose a probabilistic approach in order to realize:

- a more complete uncertainty analysis (sweeping of the spectrum of the data file)
- a sensitivity analysis (hierarchical organization of the most influential input parameters on the output parameters).

On the other hand, the probabilistic method requiring a great number of simulations (what is incompatible with the computing power available and the complexity of the models), all the countries underline the need for bringing simplifications beforehand :

firstly by limiting the number of uncertain parameters taken into account in the treatment of uncertainties : this is done by experts opinion and/or experiments return

secondly by simplifying the model himself: reduction of space dimension, physical or numerical simplification, use of an analytical model

3. STUDY OF VARIOUS METHODS FOR UNCERTAINTY AND SENSITIVITY ANALYSIS

3.1 Deterministic method

3.1.1 Presentation

The deterministic method is the most direct method and the one which requests less simulations. It consists in general of determining for each uncertain parameter a value considered as reasonable or the most probable value (best estimate.) and another value considered as pessimist (weak probability/strong consequence in our case) and to carry out simulations starting from these values.

3.1.2 Advantages

The principal advantages of this method is its lower cost, since it requests very few simulations, and its apparent simplicity of using.

Moreover, the post-processing study is immediate, since related to the choices carried out at the time of pre-processing..

The method avoids the defintion of a probability density for each parameter. Indeed, you can thus determine with enough precision the most probable value and a penalizing value, but it is sometimes difficult to choose a law of probability for a variable as well as the parameters associated with this law (mean, standard deviation, maximum.) .

It should also be noted that this method and the results obtained are easily presentable and much more comprehensible (only one value) for nonscientific or non specialists listeners than more complex methods.

3.1.3 Disadvantages

The problem of this method will be to choose the values of the parameters to obtain the anticipated result. Thus, with which value do you have to fix a parameter according to a uniform law to determine the most probable scenario?

It is difficult to be sure to have represented the extreme values of the output variable by a deterministic method: one does neither know for which value a parameter will be more penalizing, nor which combination of parameters will be more penalizing.

This approach does not make it possible by herself to have a clear decision rule . At the end of the analysis, the used is left with a list of undifferentiated results. So certain results violate the performance objectives, there is no clear numerical criteria which permits to have a rule. Only the judgment can be used.

The sensitivity analysis cannot be made with this method.

3.2 Monte Carlo Probabilistic method

3.2.1 Presentation

The probabilistic method of Monte Carlo[13] type, sometimes called empirical method or stochastic method, is the most employed method for sensitivity analysis by the various foreign counterparts of the ANDRA. Starting with the parameters whose level of knowledge is matched with an uncertainty interval, this method consists of realizing a certain number of simulations (while varying these parameters). So, a statistical analysis of the results obtained can be carried out , in order to determine the importance of various input parameters and their correlation.

3.2.2 Advantages

The probabilistic methods allow :

- to use all the traditional tools of statistical analysis.
- to obtain the probability of an event related to an output parameter
- to have an error estimate which converges towards 0 when the number of tests is increased.
- Moreover, they can put forward various correlation (or non-correlation) between the parameters which could have escaped a preliminary analysis.

3.2.3 Disadvantages

On the other hand, These methods can require a great number of simulations. When each simulation is expensive in time, this default quickly becomes crippling.

They do not make it possible to have complete security on the results, in a qualitative point of view. Indeed, the laws of probability used to represent the uncertainty of the parameters are determined by judgements. It results from it that the law from probability associated to the output parameter is biased by this preliminary judgement.

The probabilistic results can be sometimes difficult to interpret, and will be difficult to present at a non-expert

audience.

The probabilistic methods require the installation of functions of distribution for the various dubious parameters. This stage can sometimes be very difficult and vague: this inaccuracy will be reflected on statistical analysis.

3.3 FORM and SORM Methods

3.3.1 Presentation

FORM and SORM methods [14] are used to determine the failure function. Begin with on output parameter, which is associated to a breaking value or threshold that we wish not to exceed, these methods will determine a function border or failure function, noted H which will determine the frontier between the output values to the lower part and upper part of the threshold. In this case, the output parameter studied is the amount of radioactivity, while the breaking value corresponds to the maximum authorized value. These methods are in general catalogued among the probabilistic methods, and can be used jointly with the methods of the Monte Carlo type. Indeed, they make it possible to determine the most influential areas on the probability of failure. It permits to know, for example, where to apply an importance sampling, to minimize the number of tests by maximizing the quality of the result.

3.3.2 Advantages

One of the advantage of these methods is their lower cost. You can theoretically determine for which values of the input parameters the output can be found in a critical situation. These methods provide relevant indicators :

- probability of failure
- index of reliability,
- importance factor
- sensitivity factor (specific of the method)

They make it possible moreover to give proposition to improve the reliability of the system

3.3.3 Disadvantages

On the other hand, the whole methodology is based and dependent on the determination of the design point P*

- This method can present limit, in particular when there is no field of failure (for certain scenarios of safety) and when the surface of failure is discontinuous
- to be sure of the validity of the point of design P *, optimality tests must be set up and can in certain cases be expensive
- The sensitivity analysis is not global, but localized around the design point.

It can be difficult to explain the method and results for a non-specialist audience.

3.4 Directional method

3.4.1 Presentation

The directional method [14] , like FORM and SORM methods, is based on the determination of a failure surface, noted H . The principal difference is that we do not seek to represent H, but to determine the probability of exceeding this border, by traversing a broad panel of directions.

3.4.2 Advantages

The directional method is robust, whatever the form and the position of the surface of failure.

This method is more effective when the surface of failure is almost hyperspheric in standard space

3.4.3 Disadvantage

The principal problem of the directional method is that it is really effective when you know areas where the probability of failure is concentrated, which is seldom the case.

3.5 Response Surfaces

3.5.1 Presentation

The method of response surfaces [14] consists of obtaining a function which simulates the phenomenon behavior in the field of variation of the influential parameters. This function will be obtained by a certain number of tests (the term of experiment plan is used). Various types of response surfaces can be built : polynomials 1st degree, polyharmonic splines, neurons networks, generalized linear models, PLS regression

3.5.2 Advantages

These methods allow to “replace” (to approach) the numerical code by an analytical function.

The advantage of these methods is their lower cost compared to a probabilistic method. This cost is however not negligible.

These methods provide a surface solution of the problem, which makes it possible to carry out uncertainty and sensitivity analysis without carrying out new simulations.

The step and the type of result obtained are rather simple to present for a neophyte public.

The error can be formally estimated

3.5.3 Disadvantages

On the other hand, they require to take into account only a restricted number of uncertain parameters. Each uncertain parameter adds a dimension to the response surface.

It is necessary to know as a preliminary the influential parameters. The choice of the number of parameter can become a problem in the sense that this number is limited in a lower position by the necessary precision for a good approximation of the solution and upper by the growing of simulation number necessary and the complexity of the model.

The valid surface of answer is made locally, and it is difficult to estimate and to quantify the global approximation

3.6 Fuzzy logic

3.6.1 Presentation

The theory of the fuzzy subsets[15] consists of reasoning not in term of probability but in term of membership function. The methodology of operation is as follows. First you have to carry out an analysis of the physical problem by breaking up the equations into a whole of representative processes of the elementary phenomena.

Then each one of these phenomenon is represented in a simplified way by an analytical function. By sensitivity studies, the parameters whose uncertainties are most influential are selected, and their uncertainties are determined. You can then carry out fuzzy calculation, independently to the numerical code to determine the uncertainty on the final result

3.6.2 Advantages

This method has lower cost. You can determine theoretically for which values of the input parameters, you are in a critical situation. .

When the method was gauged and checked, it is possible to obtain a multitude of information without additional simulation.

It makes it possible moreover to have actions proposals to improve the reliability of the system

3.6.3 Disadvantages

This method requires a preliminary and thorough study of equations system. This system is sometimes relatively complicated (coupling of various models, not linearity), and we precisely request to the method more information.

It implies creation and use of an analytical model. It cannot be used directly on the complex numerical model.

Sensitivity study must be made in a preliminary way by traditional probabilistic methods.

It seems rather difficult to explain the method for a non –specialist audience.

4. INSTALLATION OF A TEST CASE

The study of the various methods presented before is initially realized in a theoretical way. In order to better quantify certain parameters (cost of the method, clarity of the results.) ANDRA sets up a simplified test case on which each method will be tested. This test case gathers all the specific constraints (physical, numerical and data-processing) of the problem studied by ANDRA.

The comparison of the different methods is done on the basis of following criterion:

- Use of the method (pre and post)
- relevance of the sampling and number of simulations necessary
- statistical processing and relevance of the different indicators
- integration facilities of the tools in the ALLIANCES simulation platform

These tests were carried out between July and September 2002

5. CONCLUSION

By the means of these three analyses (international review, methods studying and test case), ANDRA will determine in December 2002 a choice of one or several methods and tools for tests and implementations in the ALLIANCES platform.

Outside with the choice of the method, ANDRA engaged, like its foreign counterparts, a step of simplification of the numerical model used, in order to decrease the computing time of a simulation. It is the same for the most influential parameters pre-selection. This step is regarded as essential within the framework of using sensitivity and uncertainty analysis method.

A first assessment shows that several families can be distinguished from the previous presented methods.

The deterministic method only consists of carrying out a very restricted number of simulation for precise values of the parameters. Simple to use and quick, it provides however few information.

The probabilistic methods consist of traversing a broad spectrum of parameters variation to have a global vision of the system behavior. These methods bring much more information that the determinist, in particular because of the use of statistical tools on the variation of the output parameter, on the hierarchical organization of the input parameters etc. However these methods can appear very expensive in number of simulation . In order to mitigate this problem, various methods are available (FORM/SORM, response surface, directional.) which formally consist of locally studying part of the field of solution considered as criticism.

A third alternative can be found in the fuzzy logic which amounts replacing the probability distributions by membership functions. However it requires a strong scientific preliminary investment, not so easy because of the physical complexity of our models, to set up an analytical model associated to the studied numerical model.

6. REFERENCE

1 : **S. MISHRA & AL.** *" Sensitivity Analysis of total system performance assessment results for Yucca Mountain"* International High-Level Radioactive Waster Management Conference (2001)

2 : "Total System Performance Assessment Sensitivity Analyses for Final Nuclear Regulatory Commission Regulations" Bechtel SAIC Company n° SACFR160 REV 00 (2001)

3 : **J. HELTON & AL.** "Uncertainty and Sensitivity Analysis. Results obtained in the 1996 Performance Assessment for the Waste Isolation Pilot Plant" Sandia Report n° SAND98-0365 (1998)

4 : **J. HELTON & AL.** "Uncertainty and sensitivity analysis. Results obtained in a Preliminary Performance assessment for the Waste Isolation Pilot Plant" Nuclear Science and Engineering n°114 (1993) p 286-331

5 : **J. BRUNO & AL.** *"Projet HAVL – Revue internationale sur la modélisation et les codes de calcul de sûreté pour les stockages géologiques de déchets radioactifs"* Enviros Quantisci n° C RP 0ENQ 01-004/A (2001)

6 : *"Post-closure performance assessment – probabilistic safety assessment : overview"* Nirex Science Report n° S/94/001 (1994)

7 : *"Post-closure performance assessment –modelling of groundwater flow and radionuclide transport"* Nirex Science Report n° S/94/004 (1995)

8 : **B. CORNELIS & AL.** *"Aperçu technique du rapport SAFIR 2"* ONDRAF NIROND n° 2001-05 F (2001)

9 : **T. PAPP & AL.** *"SR97 – Post Closure safety – main report vol I"* Svensk Kärnbränslehantering AB n° TR-99-06 (1999)

10 : **S. MISHRA** *"Assigning Probability Distributions to Input Parameters of Performance Assessment Models"* Svensk Kärnbränslehanterig AB n° TR-02-11 (2002)

11 : *"H 12 : Supporting report 3 : safety assessment of the geological disposal system"* Japan Nuclear Cycle Development Institute (2000)

12 : **H. MAKINO & AL.** *"Sensitivity Analysis of Monte Carlo Simulation Results Using the Kolmogorov-Smirnov d-Statistic"* Int. Association of Mathematical Geology Annual Meeting – Cancun, Mexico (2001)

13 : **J. HELTON & F. J. DAVIS** "Sampling based Methods for Uncertainty and Sensitivity Analysis" Sandia Report n° SAND99-2240 (2000)

14 : **N. DEVICTOR** *"Fiabilité et mécanique : méthodes FORM/SORM et couplages avec des codes d'éléments finis par des surfaces de réponse adaptatives"* Ph-d thesis University Blaise Pascal of Clermond 2 n° 863 (1996)

15 : **C. FREISSINET** *"Estimation des imprécisions dans la modélisation du devenir des produits phytosanitaires dans les sols : une méthode fondée sur la logique floue"* ph-d thesis Université Joseph Fourier – Grenoble I (1997).