

# COMBINING PROBABILITY AND POSSIBILITY TO RESPECT THE REAL STATE OF KNOWLEDGE ON UNCERTAINTIES IN THE EVALUATION OF SAFETY MARGINS

J. Baccou and E. Chojnacki  
*Institut de Radioprotection et de Sûreté Nucléaire,  
BP3, 13115, Saint Paul-les-Durance, France*

S. Destercke  
*CIRAD,  
Campus INRA / Montpellier SupAgro 2,  
place Pierre Viala 34060 Montpellier Cedex 2, France*

Corresponding author *J. Baccou*, [jean.baccou@irsn.fr](mailto:jean.baccou@irsn.fr)

## ABSTRACT

This paper is devoted to some recent developments in uncertainty analysis methods of computer codes used for accident management procedures in nuclear industry. A quick overview on classical probabilistic methods for uncertainty methodology is first given. It turns out that despite its attractiveness relying on a simple implementation and convenient available tools to study the statistics of the code response, probability theory does not provide satisfactory results for uncertainty quantification when the uncertainty is not only of aleatory nature. Therefore, a new approach, called the RaFu method, is introduced to avoid choosing a probability distribution when it is not justified. It is coupled with an efficient numerical treatment that increases its efficiency for industrial studies. Finally, an application of the RaFu method to the evaluation of uncertainty margins for a nuclear reactor is given and a comparison with probability-based methods is provided as well.

## 1 INTRODUCTION

Best estimate computer codes are increasingly used in nuclear industry for the accident management procedures and have been planned to be used for the licensing procedures. Unlike conservative codes, they attempt to calculate accidental transients in a more realistic way. Therefore, it becomes of prime importance, in particular for the French Institut de Radioprotection et de Sûreté Nucléaire (IRSN) in charge of safety assessment, to know the uncertainty on the results of such best estimate codes.

A large majority of uncertainty analysts uses uncertainty methodologies based on a probabilistic modelling and Monte-Carlo simulations to propagate the uncertainties through their computer codes. However, working within the probability theory framework implicitly assumes that all uncertainties are aleatory (i.e. due to the natural variability of an observed phenomenon). In practice, uncertainties can arise from imprecision (a variable has a fixed value which is badly known due to the lack of data, knowledge or experiment).

Therefore, in order to faithfully exploit the state of knowledge, recent works have proposed new methodologies leading to a more faithful quantification of uncertainty with respect to the available knowledge ([1], [2], [3], [4], [5]). However, many existing methods are often computationally costly and are thus applicable to relatively simple models, which limits the efficiency of such approaches in fields (such as nuclear safety) where models can be very complex and where computational costs have to be taken into account.

We propose in this work a new numerical treatment of such methods based on Monte-Carlo sampling techniques which reduces the computational cost and can be applied to complex models. Moreover, by using notions of order statistics, our method proposes a way to estimate the numerical accuracy of the results. The key point of our work mainly consists in setting

some decision step before the uncertainty propagation, whereas usual methods postpone this step after the propagation.

Section 2 gives a quick overview on classical probabilistic methods for uncertainty analysis. Since it will become clear that this approach does not provide satisfactory results in many cases, we introduce, in Section 3, our new method, called the RaFu method. It allows to work within an unified framework to take into account the nature of uncertainty sources and to properly represent the real state of knowledge on uncertainties. It leads also to a numerical implementation ensuring a minimal computational cost. Finally, an application of the RaFu method to uncertainty analysis for a nuclear reactor is given in Section 4 and a comparison with probability-based methods is provided as well.

## 2 PROBABILISTIC UNCERTAINTY ANALYSIS

We recall in what follows the four steps of the uncertainty analysis methodology within the probabilistic framework:

### Step 1: Identification of uncertain parameters

All important factors affecting the model results must be identified. These factors are generally referred to as the “uncertainty sources” or as the “uncertain parameters”.

### Step 2: Quantification of the knowledge about uncertain parameters

The available information about uncertain parameters is formalized. The uncertainty of each uncertain parameter is quantified by a probability density function (pdf). If dependencies are known between uncertain parameters (or classes of uncertain parameters) and judged to be potentially important, they are also specified by correlation coefficients.

### Step 3: Propagation of uncertainties through the computer code

The propagation requires, except for very simple computer codes, a coupling between the code and a statistical software. The numerical estimation is obtained thanks to Monte-Carlo simulations taking into account the pdf and dependencies chosen in the previous step. It leads to a sample of the same size for each output quantity.

### Step 4: Treatment and interpretation of the code responses

The code responses are used to get quantitative insights regarding the output variable. More precisely, the output sample is used to get any typical statistics of the code response such as mean or variance and to determine the cumulative distribution function (CDF). The CDF allows to derive the percentiles of the distribution. Its estimation is crucial for safety assessment since the CDF provides an estimation of whether or not the code response is likely to exceed a critical value. A simple and robust way to get information on the CDF is to use order statistics ([6]). The principle of order statistics is to derive statistical results from the ranked values of a sample. If  $X=(X^{(1)}, \dots, X^{(L)})$  denotes the output sample, the key idea is that the cumulative distribution of  $X^{(k)}$ ,  $F_X(X^{(k)})$ , follows the Beta law  $\beta(k, L-k+1)$  which does not depend on the distribution of  $X$ . Therefore, it is possible to derive confidence intervals for any percentiles directly from the sample values without having to determine the probability distribution of the random variable. This relevant result is very popular in the safety assessment community. It is often used in two ways: when the sample size is fixed, it provides the numerical accuracy (due to the finite sample size) associated to the estimation. It also gives for a fixed numerical accuracy the minimal sample size (and therefore the minimal number of computer runs) to perform in order to reach this accuracy.

The probabilistic model is simple to implement thanks to the uncertainty propagation by Monte-Carlo simulations. Moreover, the use of order statistics provides both simple and

robust estimators of percentiles for any output quantities. However, the probabilistic approach is tailored to handle “aleatory” uncertainty i.e. due to the natural variability or randomness of an observed phenomenon. It cannot be reduced by the arrival of new information. It turns out that in many applications, a second kind of uncertainty often arises. It results from a lack of knowledge or information and one speaks about “epistemic uncertainty”. It can come from systematical error (like a measurement which is not fully reliable), from a poor quantity of data or from subjective uncertainty (an expert providing imprecise valued quantities). In principle, this type of uncertainty can be reduced by increasing the state of knowledge (i.e. use of devices giving more precise measurements, expert providing a more informative opinion,...).

Recent works ([1]) have shown that classical probabilities tend to confuse the two kinds of uncertainty and are not tailored to handle both of them. In practice, it may lead to an unjustified reduction of the final uncertainty of the model response and affect the decision-making process in risk studies. Indeed, in the worst case, because of such an artificial reduction, the decision maker could underestimate the risk and accept a too high level of risk but a more relevant quantification of uncertainties would have shown that the code response is likely to exceed the critical value. For safety reason, it becomes of prime importance to provide a new methodology that gives the engineer a tool to measure the impact of a misleading modelization of uncertainty due to poor knowledge.

Therefore, we propose in the rest of this paper a new method for uncertainty evaluation (called the RaFu method). It allows us to mix different kinds of knowledge representation in order to respect the available information about uncertain parameters and about the nature of their uncertainty. It also integrates an efficient numerical strategy to reduce the computational cost to its minimum.

### 3 THE RAFU METHOD

We describe in the sequel Steps 2, 3 and 4 within the RaFu modelling.

#### 3.1 Quantification of the knowledge about uncertain parameters

The RaFu ([7]) method allows to handle two kinds of uncertainties: aleatory and epistemic uncertainties. As mentioned previously, the probabilistic framework is well appropriate to represent aleatory uncertainty. As for epistemic one, possibility theory ([8]) provides an attractive framework to quantify this second type of uncertainty. In particular, possibility distributions (Figure 1) are well fitted to the situation where a given variable is described by nested confidence intervals (a natural way to express uncertainty about variables). Moreover, a possibility distribution  $\pi$  induces the set of probabilities  $P_\pi = \{P / \forall A, P(A) \leq \sup_{x \in A} \pi(x)\}$ . In this sense, a possibility distribution can be seen as a model of partial probabilistic information. For example, it can be proved that the probability set induced by a trapezoidal possibility distribution (Figure 1, right) contains all the probabilities with the same core (i.e the most likely values are located in the same interval, [2;4] in our example) and the same support ([1;7]). In other words, if the uncertainty attached to a parameter P is summarized by its range of variation ([1;7]) and an interval of values within this range that P is more likely to take ([2;4]), then the trapezoidal possibility distribution of Figure 1, right, can be chosen for uncertainty quantification. Similarly, if the information related to P is its range of variation ([2;4]) and its nominal value (3), a triangular possibility distribution (Figure 1, left) is appropriate.

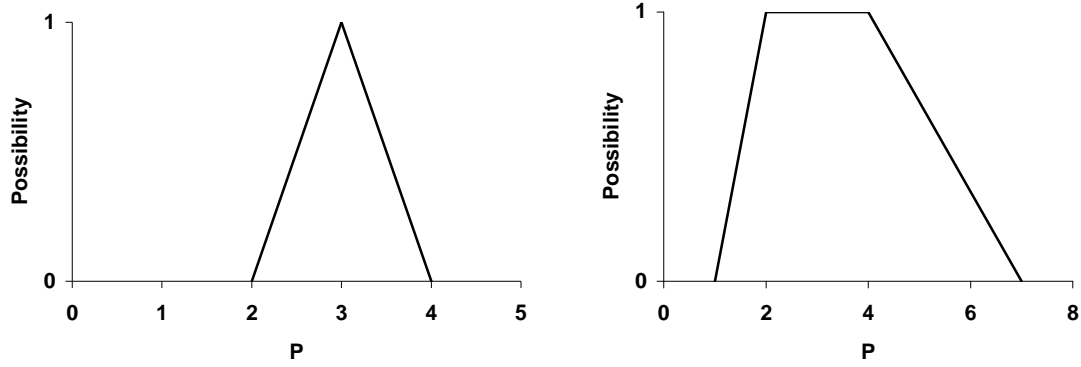


Figure 1: Example of possibility distributions. Left, triangular possibility, right, trapezoidal possibility.

It turns out that the possibility theory is a convenient way to quantify epistemic uncertainty but it can lead also to unrealistic uncertainty margins when, in the case of aleatory uncertainty, a unique pdf is justified. Therefore, our method allows the analyst to select a probability distribution or a possibility one with respect to the nature of uncertainty. This is achieved by working in an unified framework for probability and possibility called the theory of evidence ([9]). In the same way that the probability theory assigns weights to the different values taken by each uncertain parameter (for example all the points within the core of the trapezoidal distribution), the idea of the theory of evidence is to put weights on subset of values (and not necessarily on single values) such as intervals.

### 3.2 Propagation of uncertainties through the computer code and treatment of the responses

The propagation is based on an extension of Monte-Carlo simulations and therefore first requires a sampling of random (aleatory uncertainty)/badly-known (epistemic uncertainty) variables. Note that sampling a random variable gives a value (Figure 2, left) (as in classical Monte-Carlo simulations). As for badly-known variables, the sampling (Figure 2, right) is performed on the possibility distributions associated to each variable. We focus in this work on convex possibility distributions which are the most encountered in practical studies. Therefore, sampling badly-known variables leads to a set of nested intervals called  $\alpha$ -cuts (a set of nested intervals  $\{I_\alpha, 0 \leq \alpha \leq 1\}$  satisfying  $\forall \alpha \in ]0;1[, I_1 \subset I_\alpha \subset I_0$ ).

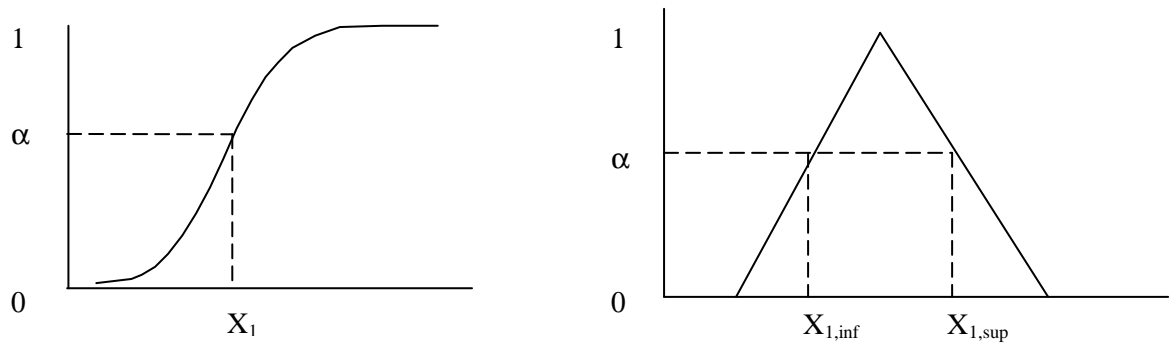


Figure 2: Sampling probability (left) and possibility (right) distributions.

It is therefore similar to performing Monte-Carlo simulations on intervals (i.e. calculations are performed with values at the extreme of each sampled interval). Since both probability and possibility are used to model parameters, the output of the computer code is no more a random variable but a random fuzzy variable. This also implies that the uncertainty derived by this methodology cannot be summarized by a pdf (or a CDF) but by a pair of lower and upper CDFs,  $[\underline{F}, \overline{F}]$ , called probability boxes ([10]). The difference between these CDFs comes from the lack of knowledge modeled by possibility distributions.

There exist many recent works that handle, like the RaFu method, both aleatory and epistemic uncertainties and derive a pair of CDFs. Among them, one can mention the work of Ferson and Ginzburg ([1]) and Baudrit *et al* ([2]) that propose a post-processing technique to extract the relevant information from the resulting random fuzzy variable. Up to now, they concern very simple models or are computationally costly, which limits their efficiency in fields such as nuclear safety where computational cost has to be taken into account.

On the contrary, the RaFu methodology integrates a computational cost reduction strategy. The underlying idea is that in many studies, the analyst is interested in some particular statistical summary (such as  $\alpha$ -percentiles, probability of exceeding a given threshold) which can be evaluated without building the whole random fuzzy variable. Since the RaFu propagation can be seen as an extension of Monte-Carlo simulation to the theory of evidence framework, each statistical quantity of interest is directly estimated using standard results coming from the probabilistic modelling. Moreover, one can exploit convergence theorems to derive the numerical accuracy associated to the limited sample size. The computational cost reduction strategy of the RaFu is then to set a decision step before propagating uncertainties and leading to an optimal (in term of number of code runs) sampling. More precisely, the RaFu method is pre-defined by a triplet of parameters  $(\gamma_S, \gamma_E, \gamma_A)$  specified by the analyst:

- Parameter  $\gamma_S$  is related to the aleatory uncertainty. It provides the statistical quantity the analyst is interested in (usually  $\alpha$ -percentiles in safety studies).
- Parameter  $\gamma_E$  is related to the epistemic uncertainty. It determines how  $\alpha$ -cuts are drawn from possibility distributions according to the behaviour the analyst wants to adopt: for example she/he can adopt a pessimistic/optimistic behaviour and set  $\alpha=0/\alpha=1$  or a balanced attitude (i.e. random  $\alpha$ ) in ambiguous situations.
- Finally, parameter  $\gamma_A$  measures the desired numerical accuracy on the final result. In the case of  $\alpha$ -percentile estimation,  $\gamma_A$  comes from the use of order statistics.

According to the analyst, the RaFu method then determines the minimal sample size and the nature of the required sampling to build the wished response. Number of calculations is thus reduced to its minimal number, in accordance with the analyst's choice. Moreover, computational cost can be easily evaluated, allowing the analyst to eventually revise her/his choices before uncertainty propagation. It is also possible for her/him to provide the maximal number of code runs that can be made and the RaFu method will derive the numerical accuracy that can be reached. For example assume that the analyst wants to have an upper limit of the response 95%-percentile, to be hyper-cautious about epistemic uncertainty i.e. concentrate on  $\alpha=0$ -cut and to have a numerical certainty of 99% to cover the true value, he or she chooses the triplet  $(\gamma_S, \gamma_E, \gamma_A)=(0.95,0,0.99)$ . The RaFu method derives the minimal sampling size, according to a classical result of order statistics, to satisfy the analyst's choice, here 90 and the nature of the sampling. If 90 is too costly, the analyst can choose to lower the numerical accuracy to 95% thus reducing the number of required computations to 59.

Figure 3 displays a flowchart of the RaFu method.

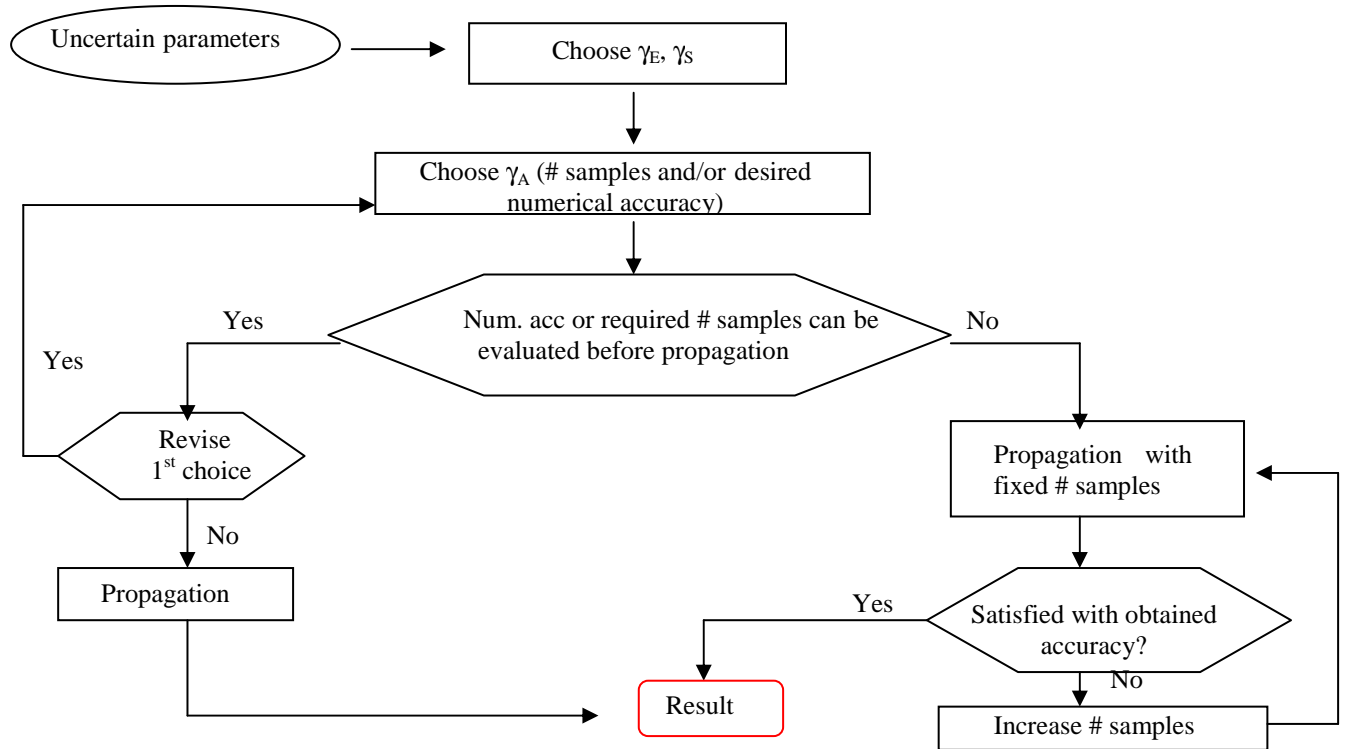


Figure 3: Flowchart of the RaFu method.

Note that even if the RaFu method deals with aleatory and epistemic uncertainty and is in this sense close to Ferson's and Baudrit *et al*'s approaches ([1] and [2]), the required number of samples leading to the same results can be very different. For example, let us consider that  $N$  uncertain parameters have been identified with the  $k$  first ones of aleatory type and  $N-k$  remaining ones of epistemic type. Assume that 100 samplings are done on the  $k$  first aleatory parameters and that for each of them one needs 21  $\alpha$ -cuts ( $\alpha=(0,0.05,\dots,1)$ ) to approximate the whole random fuzzy variable. Then, 2100 interval calculations are required to build the final result when using Ferson's and Baudrit *et al*'s approaches. With the RaFu method (because the post-processing step has been replaced by a pre-processing one), the number of interval calculations is reduced to 200 (resp. 100) to get Ferson's (resp. Baudrit *et al*'s) results.

#### 4 APPLICATION OF THE RAFU METHOD TO THE EVALUATION OF UNCERTAINTY MARGINS FOR ZION REACTOR

The numerical test deals with the evaluation of uncertainty margins for a Westinghouse nuclear reactor (called Zion reactor) in the frame of the international BEMUSE program ([11]). This reactor was shut down in 1998 following 25 years in service. The goal here is to simulate a hypothetical loss-of-coolant accident in order to study the impact on the peak cladding temperature (PCT) of a hot rod in a hot channel which is one of the most important quantity involved in safety analyses. The simulation is performed with the computer code CATHARE V2.5 mod 6.1 ([12]). It turns out that many input parameters such as correlation factor of empirical models, initial and boundary conditions, material properties, etc, are uncertain. Therefore, it is crucial for safety issues to analyse the influence of uncertainty sources on the uncertainty related to PCT.

## 4.1 Uncertain parameters

After sensitivity analysis, a list of 10 uncertain parameters (Table 1) has been proposed by IRSN. Table 1 provides also the available information related to the range of variation and the nominal value associated to each parameter.

Table 1: the 10 most influential uncertain parameters identified by IRSN.

N°	Phenomenon	Nom. value	Range
1	Liquid-wall friction	1	[0.8;1.9]
2	Fuel conductivity( $T_{fuel}<2000K$ )	1	[0.9;1.1]
3	Vapour-wall heat transfer (forced convection regime)	1	[0.5;2]
4	Peaking factor hot rod	1	[0.95;1.05]
5	Heat transfer “flashing”	1	[0.05;1]
6	Initial Upper header mean temperature +10°K	1	[1;4]
7	Initial loop mass flow rate +/-4% (head pump)	1010	[810;1210]
8	Friction form loss in the Pressurizer line	1	[0.5;2]
9	Hot gap size hot rod	1	[0.8;1.2]
10	Initial Power +/-2% (power before scram)	1	[0.98;1.02]

Among these 10 uncertain parameters, IRSN has identified 2 parameters whose uncertainty is aleatory (Parameters N°4 and 9). The 8 remaining ones are assumed to be of epistemic nature.

## 4.2 Numerical tests

In order to show the capability of our approach, the uncertainty attached to PCT is estimated using both probabilistic (i.e. the 10 uncertain parameters are assumed to be of aleatory type) and RaFu modellings (i.e. exploiting the knowledge related to the nature of uncertainty). We assume that for computational cost reason, the analyst intends to build a sample of fixed size (200) and we compare both probabilistic and RaFu methods. In every test, the statistical treatment is performed with the software SUNSET ([13]) developed by IRSN.

### Uncertainty quantification

According to Table 1, the support and the mode (i.e. the most likely value or the nominal value in our study) are known. Therefore, we choose a triangular probability (or possibility) distribution to represent the available information within each framework.

### Uncertainty propagation and statistical treatment

The propagation is performed by Monte-Carlo simulations from probability distributions (probabilistic modeling) and also from probability/possibility ones (RaFu). The statistical quantity of interest is the 95%-percentile which is the most relevant quantity to estimate in safety studies. Table 2 summarizes the 3 parameters required for the RaFu propagation. For  $\gamma_E$ , we assume that the analyst adopts a balanced attitude (compromise between pessimistic and optimistic strategies). Since the sample size has been fixed to 200, the RaFu allows us to derive the numerical accuracy (i.e. sample size effect) associated to the estimation of the percentile.  $\gamma_A$  is therefore defined by the sample size and the degree of numerical accuracy we are interested in (95% here).

Table 2: The three parameters of the decision step within the RaFu modeling.

$\gamma_S$	0.95
$\gamma_E$	random $\alpha$ -cut
$\gamma_A$	95%-accuracy, #samples=200

The numerical result is then an estimation of the 95% percentile within the probabilistic modelling and a couple of type  $[PCT_{95\%,\min}, PCT_{95\%,\max}]$  when choosing the RaFu approach (Figure 4).

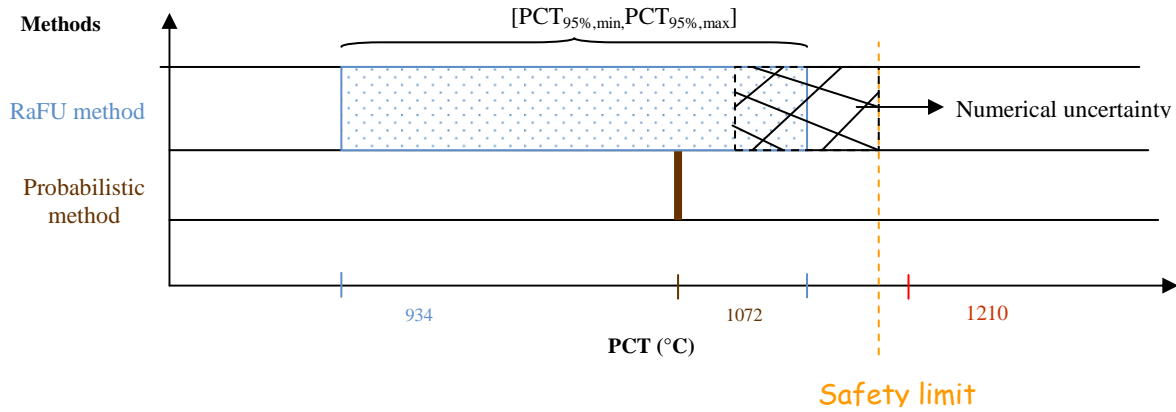


Figure 4: Estimation of the 95%-percentile with the probabilistic and the RaFu modellings.

Note that the 95% percentile obtained with the probabilistic method is lying between the lower and upper percentiles derived from our approach. This is in full agreement with the theoretical framework of the RaFu method since each triangular possibility distribution associated to epistemic uncertainty induces a partial probabilistic modeling encompassing the triangular probability distribution considered in the probabilistic approach.

This figure illustrates two main effects that need to be taken into account if a decision-making process follows the uncertainty analysis:

- Effect of the uncertainty representation with respect to their nature : a difference of  $257^{\circ}\text{C}$  on the 95%-percentile is noticeable when introducing epistemic uncertainty.
- Effect of numerical uncertainty: integrating the sample size effect leads to a difference of  $\sim 20^{\circ}\text{C}$  in the estimation of the 95% percentile. The numerical uncertainty margin is expected to be even larger for smaller sample sizes which is often the case when using complex computer codes ( $\sim 100$  simulations usually performed).

## 5 CONCLUSION

A new approach, called the RaFu method, has been constructed and applied in this paper to uncertainty analysis. Its construction is based on the theory of evidence framework that allows to handle both aleatory and epistemic uncertainties in order to respect the real state of knowledge. It is coupled with an optimal numerical treatment (based on an extension of Monte-Carlo simulations to the theory of evidence framework and on the introduction of a decision step before the propagation) that minimizes the required computation and allows the analyst to possibly revise her/his desires. Moreover, this method offers a way to control the numerical accuracy of the result. The RaFu method has successfully been applied to derive uncertainty margins for a real nuclear power plant. The results are less precise (i.e an interval instead of a value) but are more reliable for safety studies since the nature of uncertainty is taken into account and more faithfully represented than in classical probabilistic modeling. Moreover, thanks to a numerical strategy, this method is well fitted to uncertainty analysis of complex computer codes used in many industrial applications.



## REFERENCES

- [1] S. Ferson and L. Ginzburg, "Different methods are needed to propagate ignorance and variability", *Reliability Engineering and System Safety*, volume 54, pages 133-144 (1996).
- [2] C. Baudrit, D. Guyonnet and D. Dubois, "Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment", *IEEE Trans. Fuzzy Syst.*, volume 14, pp. 593-608 (2006).
- [3] J.C. Helton, J.D. Johnson, W.L. Oberkampf, C.J. Sallaberry, "Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty", *Reliability Engineering & System Safety*, Volume 91 (10-11), pp. 1414-1434 (2006).
- [4] M. Bruns, "Propagation Of Imprecise Probabilities Through Black-Box Models", M.Sc. Thesis. Georgia Institute Of Technology (2006).
- [5] R. Kenett and S. Zacks, "Modern Industrial Statistics", Duxbury (1998).
- [6] W.J. Conover, *Practical Nonparametric Statistics*, Wiley Series in Probability and Statistics, New-York (1999).
- [7] J. Baccou, E. Chojnacki and S. Destercke, "Numerical accuracy and efficiency in the treatment of epistemical and aleatory uncertainty", submitted to *Int. J. of Intelligent Systems* (2008).
- [8] D. Dubois, H.T. Nguyen and H. Prade, *Possibility Theory, Probability and Fuzzy Sets: Misunderstandings, Bridges and Gaps*, pages 343-438, *Fundamentals of Fuzzy Sets*, Kluwer Academic Publishers, Boston (2000).
- [9] G. Shafer, *A Mathematical Theory of Evidence*, University Press, Princeton (1976).
- [10] S. Ferson, L. Ginzburg, V. Kreinovitch, D.M. Myers and K. Sentz, "Constructing probability boxes and Dempster-Shafer structures", Technical report, Sandia National Lab. (2003).
- [11] A. Petruzzi, F. D'Auria, J-C. Micaeli, A. De Crecy and J. Royen, "The BEMUSE programme (Best-Estimate Methods – Uncertainty and Sensitivity Evaluation ", *Proc. of the Int. Meet. On Best-Estimate Methods in Nuclear Installation Safety Analysis (BE-2004)*, IX, volume 1, page 225-235 (2004).
- [12] B. Brun, "Current implementation and future plans on new code architecture, programming language and user interface", *Proc. of OECD/CSNI workshop on transient thermal-hydraulics and neutronic requirements*, Anapolis (1996).
- [13] E. Chojnacki and A. Ounsi, "Description of the IPSN method for the uncertainty and sensitivity analysis and the associated software: SUNSET", *Proc. of ASME/JSME ICONE*, volume 3, page 545-550, Louisiana (1996).