A BAYESIAN ENSEMBLE OF SENSITIVITY MEASURES FOR SEVERE ACCIDENT MODELING

Seyed Mohsen Hoseyni [1], Francesco Di Maio*[2], Matteo Vagnoli[2], Enrico Zio[2,3], Mohammad Pourgol-Mohammad[4]

[1] *Department of Basic Sciences, East Tehran Branch, Islamic Azad University*

Tehran, Iran

[2] *Energy Department, Politecnico di Milano*

Via La Masa 34, 20156 Milano, Italy

francesco.dimaio@polimi.it

[3] *Chair on System Science and Energetic Challenge*

Fondation EDF – Electricite de France

Ecole Centrale, Paris, and Supelec, Paris, France

4 *Department of Mechanical Engineering, Sahand University of Technology*

Tabriz, Iran

ABSTRACT

In this work, a sensitivity analysis framework is presented to identify the relevant input variables of a severe accident code, based on an incremental Bayesian ensemble updating method. The proposed methodology entails: i) the propagation of the uncertainty in the input variables through the severe accident code; ii) the collection of bootstrap replicates of the input and output of limited number of simulations for building a set of Finite Mixture Models (FMMs) for approximating the probability density function (pdf) of the severe accident code output of the replicates; iii) for each FMM, the calculation of an ensemble of sensitivity measures (i.e., input saliency, Hellinger distance and Kullback–Leibler divergence) and the updating when a new piece of evidence arrives, by a Bayesian scheme, based on the Bradley-Terry model for ranking the most relevant input model variables. An application is given with respect to a limited number of simulations of a MELCOR severe accident model describing the fission products release in the LP-FP-2 experiment of the Loss Of Fluid Test (LOFT) facility, which is a scaled-down facility of a pressurized water reactor (PWR).

1. INTRODUCTION

Simulation codes used for thermal hydraulic (TH) or severe accident calculations are very complex and, therefore, computationally burdensome. For this reason, one attempts to perform sensitivity analysis with minimal computational burden, for an effective dimensionality reduction of input variables, which can speed up calculations without diminishing the code capability of representing the real system response. Sensitivity analysis can, indeed, inform on the input variables most affecting the simulation results of the system results [Saltelli, 2002; Pourgol-Mohammad, 2009; Di Maio et al., 2014]. However, most sensitivity analysis methods can turn to be themselves computationally expensive [Secchi et al., 2008; Carlos et al., 2013; Di Maio et al., 2014; Pengfei et al., 2014].

In this work, the focus is on the implementation of a sensitivity analysis method to perform input variables ranking with an affordable computational effort. Different solutions have been proposed in literature. One could make theoretical efforts for developing insightful simple novel sensitivity measures [Liu et al., 2010; Plischke et al., 2013] or resort to traditional sensitivity measures (such as Variance Decomposition method [Helton, 1993; MacKay, 1995; Borgonovo, 2007; Cadini et al., 2007], Sobol Indices [Sobol et al., 1995; Saltelli et al., 1999]) in combination with surrogate models like Artificial Neural Networks (ANNs) [Cadini et al., 2007] and Response Surface Methods (RSMs) [Hoseyni et al., 2014], where the idea is to use the surrogate models as substitutes of the complex simulation codes, in the quantification of the contribution of each parameter to the model output.

Alternatively, a relatively novel methodology for global sensitivity analysis is proposed in [Carlos et al., 2013; Di Maio et al., 2015] that does not resort to any surrogate model. The sensitivity results are drawn directly by analyzing Finite Mixture Models (FMMs) built on the results of an uncertainty propagation of the input variables uncertainties onto the simulation code. Ensemble of sensitivity measures can be used [Di Maio et al., 2014], where a set of alternative input variable rankings are aggregated to overcome possible misjudgments of the single sensitivity measure in case of a limited number of code runs. The aggregation methods commonly used are the "majority voting" and "sum" [Kukkonen et al., 2007], where the former consists in assigning the ranking position to each input variable which has been voted by the majority of the calculated sensitivity measures, whereas the latter aggregates by taking the sum of the ranking positions for each input variable provided by the individual sensitivity measures and, then, sorts them with respect to their scores.

The limitations of these two traditional aggregation strategies are that "majority voting" fails in the case of no agreement among the ranking orders provided by the different sensitivity measures, whereas the "sum" strategy fails when the evidence of the superior capability of one (or more) sensitivity measure(s) cannot be accommodated over the remaining ones, within its rigorous

assignment of equal weights (equal preferences) to the outcomes of the sensitivity measures considered.

In this work, a novel sequential Bayesian approach is embedded into the framework of ensemble-based sensitivity analysis proposed in [Di Maio et al., 2014] for the aggregation of sensitivity measures, so as to overcome the above-mentioned deficiencies. The proposed sensitivity analysis approach is applied to the LP-FP-2 severe accident following a Loss Of Coolant Accident (LOCA) of the Loss of Fluid Test (LOFT) facility, which is a scaled-down facility of a Pressurized Water Reactor (PWR). The most relevant input variables affecting the output of the MELCOR severe accident model are identified.

The novel sequential approach resorts to the FMMs built based on a limited number of code simulations for approximating the probability density function (pdf) of the severe accident model output variables (i.e., the fission products releases during the LOCA) and uses three alternative sensitivity measures, input saliency [Law et al., 2004], Hellinger distance [Gibbs et al., 2002; Diaconis et al., 1982] and Kullback–Leibler divergence [Gibbs et al., 2002; Diaconis et al., 1982], to identify the input variables bearing the most pronounced effects on the output variable, and make their aggregated ranking.

The novelty of the Bayesian aggregation stems from the capability of providing weights to the sensitivity measures and aggregating them according to the degree of belief that one could have on the calculated measures. In other words, the evidence that one sensitivity measure overcomes the others (i.e., the larger degree of belief/weight on one sensitivity measure over the others) is, in principle, obtained by performing a set of sequential experiments with the same Design Of Experiments (DOEs). When the sequential experiments cannot be performed, one could resort to bootstrapped samples of the original outcomes of the experiments to build the FMMs replicates. For each new FMM, the values of sensitivity measures are calculated and treated as the new data (under the same DOE) within a Bayesian updating scheme of the original sensitivity values. A Bradley-Terry model is devised for Bayesian updating [Weng et al., 2011] of the sensitivity measures of the provided input variable ranking orders, that is a full-pair comparison scheme for on-line ranking of multiple players (e.g., input variables) in successive matches (i.e., experiments).

The paper is organized as follows. Section 2 discusses the principles of the ensemble-based sensitivity analysis along with the description of FMMs and the three sensitivity measures used. In Section 3, a definition is given for the Bayesian aggregation approach, the bootstrap technique and the updating rules. Section 4 presents the case of study. Section 5 analyses the results of the implementation of the proposed Bayesian ensemble strategy on the case of study. Conclusions are made in Section 6.

2. Ensemble-based sensitivity analysis

In this Section, a brief summary is provided of the ensemble-based approach for sensitivity analysis presented in [Di Maio et al., 2014]. As shown in Fig. 1, the idea is to rely on a limited number N of code simulations to build a FMM that reproduces the multimodal pdf of the output variable (Section 2.1). In order to identify the input variables with most pronounced effects on the severe accident model output variable, three alternative sensitivity measures are computed (i.e., input saliency [Law et al., 2004], Hellinger distance [Gibbs et al., 2002; Diaconis et al., 1982] and Kullback–Leibler divergence [Gibbs et al., 2002; Diaconis et al., 1982]) (Section 2.2). The reliability of the sensitivity measures is quantified by the sequential Bayesian scheme that aggregates the rankings according to the degree of belief that one could have on the calculated measures (Section 2.3).
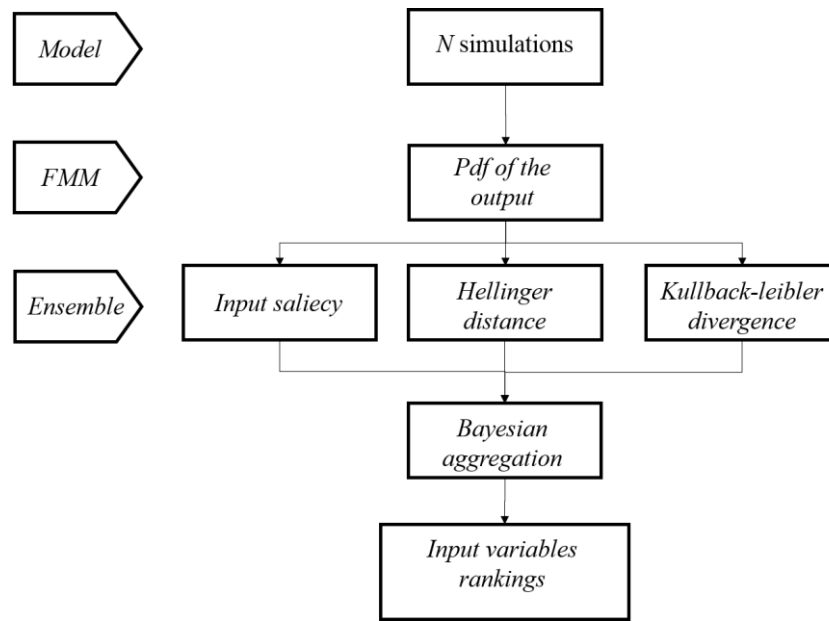

Fig.1 *Flowchart of ensemble-based methodology*

2.1 FMM model

Let y represent the output of a complex code model m, viz:

$$y = m(x_1; x_2; \ldots; x_l; \ldots; x_D) \qquad l = 1, \ldots, D \tag{1}$$

where x_l is the realization of the *l-th* input variable of the model, whose inputs are distributed with pdf $f(x_l)$. The random output variable y follows a finite mixture density $f(y)$ with K models if [Carlos et al., 2013]:

$$f(y) = \sum_{k=1}^{K} \pi_k f_k(y \mid \theta_k) \qquad with \; \forall k, \pi_k \geq 0 \tag{2}$$

where $f_k(y \mid \theta_k)$ are K different probability density functions, θ_k is the set of parameters of the *k-th* model of the mixture and π_k are the mixing probabilities that satisfy:

$$\sum_{k=1}^{K} \pi_k = 1 \tag{3}$$

The Expectation Maximization (EM) algorithm [Mc Lachlan et al., 2000] is used to fit $f(y)$ to the N available data $y = (y_1, \ldots, y_N)$, $i = 1, \ldots, N$ and for the identification of parameters θ and π [Di Maio et al., 2014]. Once the parameters θ and π of the mixture models are known, the best approximation of the pdf of the output of the model is completely characterized with small number of code simulations.

2.2   Sensitivity Measures

2.2.1 Input saliency

In the case of independence of input variables, the FMM of Eq. (2) can be rewritten as a function of the D input variables of the model:

$$f(y) = \sum_{k=1}^{K} \pi_k f_k(y \mid \theta_k) = \sum_{k=1}^{K} \pi_k m(\prod_{l=1}^{D} f(x_l \mid \theta_{kl})) \qquad (4)$$

where m is the severe accident model function and $f(x_l \mid \theta_{kl})$ is the pdf of the *l-th* input x that determines the output y to belong to the *k-th* cluster. The *l-th* input does not affect the output if its distribution is independent from the cluster, i.e., it follows its common density $q(x_l \mid \theta_{kl})$ among all the clusters [Pudil et al., 1995; Vaithyanathan et al., 1999]. The decomposition of the term $f(x_l \mid \theta_{kl})$ in Eq. (4) with weights ρ_l into a distribution accounting for the contribution of the *l-th* input in the *k-th* cluster $f(x_l \mid \theta_{kl})$ and with weight "1- ρ_l" in the common distribution $q(x_l \mid \theta_{kl})$, yields to:

$$f(y) = \sum_{k=1}^{K} \pi_k m \left(\prod_{l=1}^{D} \rho_l f(x_l \mid \theta_{kl}) + (1 - \rho_l) q(x_l \mid \lambda_l) \right) \qquad (5)$$

The saliency ρ_l is the importance of the *l-th* input in affecting the output y. In fact, if ρ_l is large, it means that the input variable distribution varies significantly from one cluster to another of the "natural" clusters corresponding to each model $f_k(y \mid \theta_k)$ and, thus, the input is important in determining the variability of the output. On the other hand, if ρ_l is small, the inputs follow the common distribution in any cluster and, thus, the input is not relevant in shaping the distribution of the output [Di Maio et al., 2014].

The estimation of the input variable importance ρ_l is a model parameter identification problem that does not admit any closed form analytical solution [Figueiredo et al., 2002]; thus, it can be tackled by the Expectation Maximization (EM) algorithm as proposed in [McLachlan et al., 2000; Di Maio et al., 2014].

2.2.2 Hellinger distance

The Hellinger distance measures the difference between the pdf of the *l-th* input contributing to shape the *k-th* cluster $f(x_l \mid \theta_{kl})$ and its common distribution $q(x_l \mid \theta_{kl})$ [Gibbs et al., 2002; Diaconis et al., 1982] by:

$$H_{lk}\left(f(x_l \mid \theta_{kl}), q(x_l \mid \lambda_l)\right) = \left[\frac{1}{2}\int_{-\infty}^{+\infty}\left(\sqrt{f(x_l \mid \theta_{kl})} - \sqrt{q(x_l \mid \lambda_l)}\right)^2 dx\right]^{\frac{1}{2}} = \left[1 - \int_{-\infty}^{+\infty}\left(\sqrt{f(x_l \mid \theta_{kl})q(x_l \mid \lambda_l)}\right)^2 dx\right]^{\frac{1}{2}} \quad (6)$$

satisfying the inequality $0 \leq H_{lk}\left(f(x_l \mid \theta_{kl}), q(x_l \mid \lambda_l)\right) \leq 1$.

The importance of the *l-th* input variable in affecting the output *y* is computed as:

$$HL_l = \sum_{k=1}^{K} H_{lk}\left(f(x_l \mid \theta_{kl}), q(x_l \mid \lambda_l)\right) \quad (7)$$

The *l-th* input variable is important if HL_l is large (in relative terms).

2.2.3 Kullback-Leibler divergence

The Kullback–Leibler divergence measures the different information carried by the pdf of the *l-th* input in the *k-th* cluster $f(x_l \mid \theta_{kl})$ and its common distribution $q(x_l \mid \theta_{kl})$ [Gibbs et al., 2002; Diaconis et al., 1982]:

$$K_{lk}\left(f(x_l \mid \theta_{kl}), q(x_l \mid \lambda_l)\right) = \int_{-\infty}^{+\infty} f(x_l \mid \theta_{kl})\log\left(\frac{f(x_l \mid \theta_{kl})}{q(x_l \mid \lambda_l)}\right) dx \quad (8)$$

with values in $[0, \infty]$.

The importance of the *l-th* input variable in affecting the output *y* is quantified as:

$$KL_l = \sum_{k=1}^{K} K_{lk}\left(f(x_l \mid \theta_{kl}), q(x_l \mid \lambda_l)\right) \quad (9)$$

The *l-th* input variable is important if KL_l is large (in relative terms).

2.3 Aggregation

The idea of the ensemble is particularly useful for sensitivity analysis when the number N of code simulations is deliberately kept low to control computational cost. As a consequence, due to the limited quantity of data, individual sensitivity methods would mislead the individual rankings because, as shown in Section 2.2, input saliency, Hellinger distance and Kullback-Leibler divergence would quantify different aspects of the input variables importance. On the other hand, the ensemble approach is expected to generate reliable rankings by agreement among the rankings produced by the individual methods without requiring any additional simulation. Thus, the advantage of the ensemble approach is twofold: it can overcome possible misjudgments of the individual methods when alternative sensitivity measures disagree, and, furthermore, it is possible to gain more confidence on the ranking when they agree.

The central issue is to decide how to aggregate the sensitivity ranking outcomes provided by the different methods in the ensemble [Baraldi et al., 2011; Di Maio et al., 2012].

In this paper, we propose a novel Bayesian approach for the aggregation of sensitivity measures results. Classically, in the Bayesian process, for a given prior distribution, one can update the state of knowledge once a new piece of data or evidence becomes available. In the proposed framework, the prior is the value of the sensitivity measures obtained for each parameter from the original FMM; the new evidence (for updating the state of knowledge about the three sensitivity measures) consists in the values of the sensitivity measures derived from the bootstrapping procedure (to mimic the repetitions of experiments under the same DOE).

The ranking results obtained by using the proposed Bayesian approach will be compared with the rankings obtained by traditional ensemble methods such as "majority voting" and "sum" [Kukkonen et al., 2007], where the "majority voting" method fails in the case of no agreement among any of the ranking orders provided by the different sensitivity measures, whereas the "sum" strategy can be proven to be ineffective when the evidence of the superior capability of one (or more) sensitivity measure(s) over the remaining ones cannot be accommodated within its rigorous assignment of equal weights to the outcomes of the involved sensitivity measures.

The benefit of the proposed Bayesian aggregation is to provide weights to the sensitivity measures and aggregate them according to the degree of belief that one could have on the calculated measures. The rationale behind the evidence that one sensitivity measure has larger degree of belief than the others can be obtained (in principle) by performing a set of sequential experiments with the same DOEs; otherwise, when the sequential experiments cannot be performed, one could resort to bootstrapped samples of the original outcomes of the experiments to build the FMMs replicates. In other words, a weight (performance) is assigned to each sensitivity measure in order to i) reward the sensitivity measure with high-ranking stability (reproducibility) during the updating process and ii)

penalize the measures with fluctuating orders throughout the successive ranking that treats the successively accumulated evidences (i.e., the repeated experiments under the same DOE/bootstrapped repetitions).

3. The Bayesian aggregation approach

The proposed approach entails: i) the generation of B bootstrapped samples of the original N code outputs dataset to build B new FMMs replicates (Section 3.1); ii) the evaluation of the three sensitivity measures for each new FMM, that is, thus, treated as a new piece of evidence for the Bayesian updating of the original sensitivity values (Section 3.2) by the Bradley-Terry model [Hunter, 2004; Weng et al., 2011]. Finally, the aggregation of the sensitivity measures is obtained (Section 3.3).

3.1 The Bootstrap

The bootstrap is a computer-based method for assessing the accuracy of statistical estimates with very little assumptions or analysis [Efron et al., 1993]. It is used in different frameworks, e.g., in an order statics framework for computing the TH code uncertainties in order to estimate the safety margins using limited samples of Thermal-Hydraulic (TH) code execution [Pourgol-Mohammad et al., 2007; Zio et al., 2008], or in a framework to propagate uncertainties in TH code calculations [Probst et al., 2006]. The main benefit of bootstrapping approach is to avoid additional computational burden and rely on the available data from the results of uncertainty analysis already performed. In general, 100–200 bootstrap replications are enough to obtain a good estimate [Zio et al., 2008].

The basic idea is to generate a number B of bootstrap samples drawn at random with replacement from the original training set of N input/output patterns. The generic bootstrapped sample b-th is constituted by the same number N of input/output, although, due to sampling with replacement, some of the input/output will appear more than once whereas some will not appear at all [Efron et al, 1993]. In the proposed approach, the bootstrap technique is employed for generating B times the N replicates from the original dataset of the code as shown in Fig. 1. In other words, each bootstrap process takes samples of the original output dataset and generates a new dataset. The FMMs are built on these new data and used for the calculation of sensitivity measures. Thus, new data/evidences are available for use in the Bayesian updating process.

3.2 Bayesian updating of the sensitivity measures

A Bayesian updating process is implemented to derive simple analytic rules for updating the ranking of the code input variables based on the outcomes of the sequential experiments [Weng et al., 2011], based on the Bradley-Terry model (typically used for ranking competitive games) [Hunter, 2004]. The proposed process for updating the sensitivity measures of the input variables (that can be seen as players) uses data provided by the bootstrapped FMMs replicates (that can be seen as consecutive games). With this premises, it can be said that the *l-th* player (input variable) beats the *q-th* player (input variable) if the first one has lower rank position (higher strength) than the latter. The resulting sensitivity measures represent the games results and, thus, an input variable is more relevant if its sensitivity measure value is higher.

Given a prior distribution of a random variable, Bayesian process allows updating its knowledge when new data or evidence becomes available. In our case, the prior knowledge is the value of the sensitivity measures obtained for each input variable from the original FMM, whereas the new information used for updating the state of knowledge about the three above mentioned sensitivity measures are the values of the sensitivity measures derived from the B replicates of the original N simulations.

Let η_l be the strength, i.e. the sensitivity measure value of the *l-th* input variable whose value is to be estimated. The Bradley-Terry Bayesian ranking method starts by assuming that η_l has a prior distribution $N(\mu_l, \sigma_l^2)$, with μ_l and σ_l^2 known [Weng et al., 2011]. Upon a game (bootstrap replication) completion, based on its ranking results $T = [T_1, ..., T_D]$, the normal distribution $N(\mu_l, \sigma_l^2)$ is updated to the posterior $N(\mu_l^*, \sigma_l^{2*})$, whose updated mean μ_l^* and variance σ_l^{2*} are considered as prior information for the next game up to the B updates.

The procedure for updating η_l is hereafter described. Let $\bar{Z} = [Z_1, ..., Z_D]'$ be the vector of standard normal variables with probability density function Φ_D [Weng et al., 2011]

$$Z_l = \frac{\eta_l - \mu_l}{\sigma_l}, l = 1, ..., D \tag{10}$$

where D is the number of players (input variables). The posterior density distribution of $\bar{Z}$ given the game outcome $\bar{T} = [T_1, ..., T_D]$, i.e., the ranking results, is [Weng et al., 2011]:

$$P(\overline{Z} \mid \overline{T}) = \frac{\Phi_D(\overline{Z}) f(\overline{Z})}{\int \Phi_D(\overline{Z}) f(\overline{Z}) dz} \tag{11}$$

where $f(\overline{Z})$ is the probability of the game outcome $P(\overline{Z} \mid \overline{T})$:

$$f(\overline{Z}) = \prod_{l=1}^{D} f(Z_l) \tag{12}$$

The posterior mean μ_l^* and variance σ_l^{2*} of η_l for each l-th input variable, where $l = 1,2,\ldots, D$, are related to Z_l [Weng et al., 2011]:

$$\mu_l^* = E[\eta_l] = \mu_l + \sigma_l E[Z_l] = \mu_l + \sigma_l E\left[\frac{\partial f(\overline{Z}) \big/ \partial Z_l}{f(\overline{Z})} \right] \tag{13}$$

and

$$\sigma_l^{2*} = Var[\eta_l] = \sigma_l^2 Var[Z_l] = \sigma_l^2 \left(E[Z_l^2] - E[Z_l]^2 \right)$$
$$= \sigma_l^2 \left(1 + E\left[\frac{\partial^2 f(\overline{Z}) / \partial Z_l^2}{f(\overline{Z})} \right] \Big|_{\overline{Z}=0} - E\left[\frac{\partial f(\overline{Z}) / \partial Z_l}{f(\overline{Z})} \right]^2 \right) \tag{14}$$

The expectation terms in Eqs. (13) and (14) are the average relative changes of f (the probability of the game outcome, that is, the ranking results) with respect to the strength η_l . For example, suppose that input variable 1 beats input variables 2: f is expected to increase for η_1 and the adjustment Ω_1 of η_1 is equal to the relative (i.e., normalized on $f(\overline{Z})$) rate of change of all input variables strengths with respect to η_1 (i.e, $\partial f(\overline{z})/\partial z_1$), whereas, f is expected to decrease for η_2 and the adjustment Ω_2 is negative.

The approximating expectations are, thus, given by:

$$\mu_l^* \leftarrow \mu_l + \Omega_l \tag{15}$$

$$\sigma_l^{2*} \leftarrow \sigma_l^2 \max(1 - \Delta_l, \kappa) \tag{16}$$

where [Weng et al., 2011]:

$$\Omega_l = \sigma_l \frac{\partial f(\overline{Z})/\partial Z_l}{f(\overline{Z})}\Big|_{\overline{Z}=0} \tag{17}$$

$$\Delta_l = -\frac{\partial}{\partial z_l}\left(\frac{\partial f(\overline{Z})/\partial Z_l}{f(\overline{Z})}\right)\Big|_{\overline{Z}=0} \tag{18}$$

$$\frac{\partial f(\overline{Z})/\partial Z_l}{f(\overline{Z})} = \frac{\partial \log \prod_{l=1}^{D} f(Z_l)}{\partial Z_l} \tag{19}$$

and κ is a lower bound to ensure a positive σ_l^2.

Notice that, the posterior density of η is expected to be centered in μ and, thus, the derivative terms are evaluated at $\overline{Z}=0$ so that the η_l is replaced by μ_l. Furthermore, the right terms of Eqs. (15) and (16) are functions of μ_l and σ_l^2, so the updated mean μ_l^* and variance σ_l^{2*} are computed using the current values of μ_l and σ_l^2 [Weng et al., 2011].

Finally, we can calculate the probability (after any of the B updates) that variable l is more relevant than variable q as:

$$P(player\ ``l"\ beats\ ``q") \equiv f_{lq}(Z) = \frac{e^{\eta_l^*/c_{lq}}}{e^{\eta_l^*/c_{lq}} + e^{\eta_q^*/c_{lq}}} \tag{20}$$

where

$$c_{lq} = \sigma_l^{2*} + \sigma_q^{2*} + 2\beta \tag{21}$$

and β is a positive constant [Weng et al., 2011].

The pseudo-code of the Bradley-Terry model for full-pair comparison tailored to the ranking updating problem is as follows:

1. Calculate μ_l and σ_l with $l=1, 2, \ldots, D$.

2. For $l=1, 2, \ldots, D$

      2.1 For $q=1, 2, \ldots, D$, with $q \neq l$,

3. Calculate Ω_l and Δ_l (in Eqs. (17) and (18)) by the following steps:

      3.1 Calculate

$$c_{lq} = \sigma_l^2 + \sigma_q^2 + 2\beta \tag{22}$$

$$p_{lq} = \frac{e^{\mu_l/c_{lq}}}{e^{\mu_l/c_{lq}} + e^{\mu_q/c_{lq}}} \tag{23}$$

$$p_{ql} = \frac{e^{\mu_q/c_{lq}}}{e^{\mu_l/c_{lq}} + e^{\mu_q/c_{lq}}} \tag{24}$$

$$\delta_{ql} = \frac{\sigma_l^2}{c_{lq}}(s - p_{lq}) \tag{25}$$

$$\varepsilon_q = \gamma_q (\frac{\sigma_l^2}{c_{lq}})^2 p_{lq} p_{ql} \tag{26}$$

where s is equal to:

$$s = \begin{cases} 1 & if \ \ rank(q) > rank(l) \\ 1/2 & if \ \ rank(q) = rank(l) \\ 0 & if \ \ rank(q) < rank(l) \end{cases} \tag{27}$$

and γ_q is an arbitrary parameter [Weng et al., 2011].

      3.2. Calculate

$$\Omega_l = \sum_{q \neq l} \delta_q \tag{28}$$

$$\Delta_l = \sum_{q \neq l} \varepsilon_q \tag{29}$$

      3.3. The mean value μ_l and the standard deviation σ_l of the l player are updated using

      Eqs. (15) and (16), respectively.

3.3 Weighting the sensitivity measures within the Bayesian updating procedure

The bootstrapped FMMs replicates and, thus, the input variables sensitivity measures can fluctuate due to the randomness of the sampling (e.g., the uncertainties of the measures in the real experiments with the same DOE) that might lead to wrong ranking results. To counterbalance possible misleading results, we build an ensemble of sensitivity measures introducing a weight (performance) to each sensitivity measure in order to i) reward the sensitivity measure with high-ranking stability (reproducibility) during the updating process and ii) penalize the measures with fluctuating orders throughout the successive ranking that treats the successive accumulated evidences (i.e., the repeated experiments under the same DOE/bootstrapped repetitions). The weight (performance index) of the sensitivity measure, $sm= 1, 2, 3$ for input saliency, Hellinger divergence and Kullback Leibler, respectively, is defined using the Coefficient Of Variation (COV) as:

$$w_{sm} = \frac{1/COV_{sm}}{\sum_{sm=1,2,3} (1/COV_{sm})} \tag{30}$$

The COV of each 'sm' is the mean value of the COV of the sensitivity measures evaluated on all the code input variables at the last update, accounting for the input variables sensitivity measures variability in relation to their mean value, i.e., COV aims to describe the goodness of each sensitivity measure in terms of the relative sizes of the squared residuals and mean values. The higher the COV, the greater the dispersion in the sensitivity measures evaluated on all the code input variables; whereas, the lower the COV, the smaller the residuals relative to the sensitivity measures mean value. COV is defined as [Zio, 2007]:

$$COV_l = \frac{\sigma_l}{\mu_l} \tag{31}$$

where $l=1, 2, ….D$.

Once the weights are calculated for each SA measure, the Bayesian aggregation of their values quantifies for each l-th input variable an ensemble of the three SA measures as the weighted sum of the three η_l :

$$ESA_l = \sum_{sm=1}^{3} w_{sm}\eta_{sm} \tag{32}$$

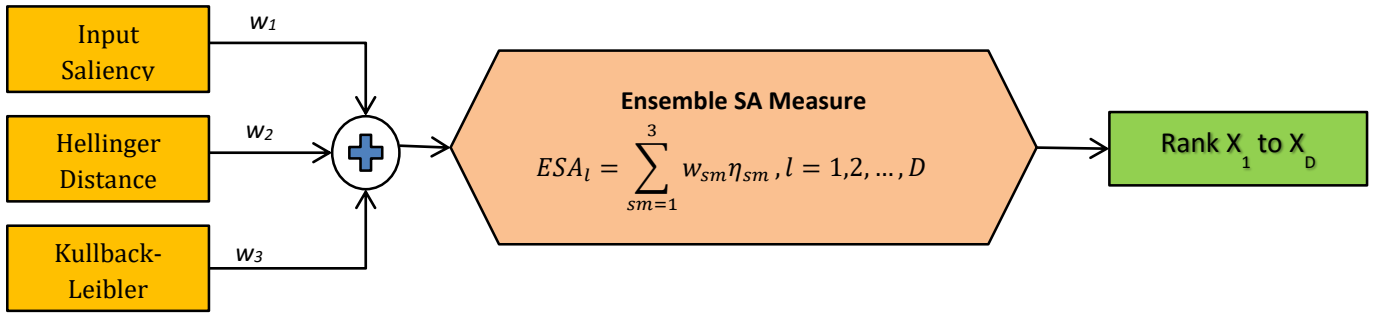Fig. 2 schematically represents the proposed scheme.

Fig. 2: *Bayesian aggregation of SA measures*

4. The LOFT LP-FP-2 experiment

4.1 Description of LOFT LP-FP-2 experiment

The LOFT experimental facility [Kmetyk, 1992] was designed to simulate the major components and system responses of a current-generation Pressurized Water Reactor (PWR) during a LOCA. The experimental subsystems include the reactor vessel, the intact loop, the broken loop, the Blow-down Suppression Tank (BST) system, and the Emergency Core Cooling System (ECCS). The layout of the major LOFT components for test LP-FP-2 is shown in Fig. 3. LOFT is a PWR model with a rated power of 50 MW$_{th}$, including all systems and components needed to simulate a severely damaged core assembly and to determine the fission product release and transport to the piping system [Lewis et al., 2008]. The LOFT-FP project, completed in 1985 [IRSN-200783, 2007], was conducted by the Idaho National Laboratory (INL/INEL) on an assembly of 121 UO$_2$ rods with nuclear heating (in-pile) core. It consisted of tests on rod degradation and fission product release, and involved temperatures up to 2400 K.
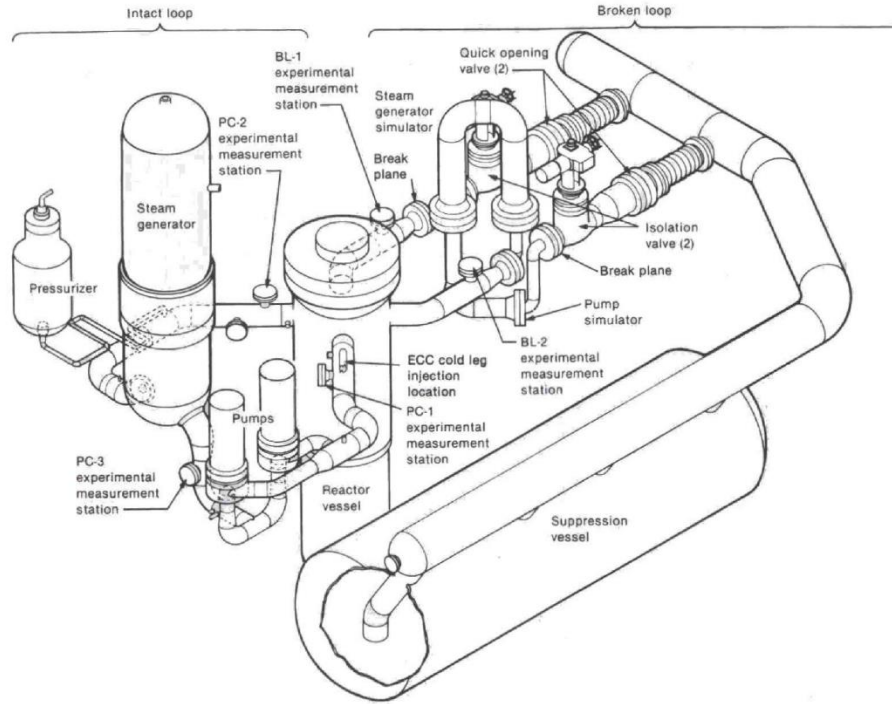
Fig. 3: *Schematic Diagram of the LOFT Test Facility for Experiment LP-FP-2* [Kmetyk, 1992]

The specific objectives of the test are to [OECD/NEA, 2001]: i) simulate a medium break LOCA, with delayed operation of the ECCS; ii) attain substantial fuel damage and maintain maximum fuel temperatures above 2100 K for several minutes; iii) study fission product migration along the leak path, which represents the Low-Pressure coolant Injection System (LPIS) pipe connected to the hot leg.

This test has provided very interesting data on core degradation and fission product release and transport in a V-type LOCA sequence, under low system pressure, in a decay-heated bundle, and at a larger scale compared with other bundle in-reactor tests. The test confirmed the damage zones observed in Three Mile Island-2, showed the inability of blockages to stop fuel oxidation, and showed significant damage and enhanced hydrogen production occurrence during the reflood. Quantitative measurements of fission product deposition have been made. The test has shown a dominance of silver iodide (AgI) as chemical form of Iodine (I).

To show the applicability of the proposed framework of Section 3 for the sensitivity analysis of the MELCOR code simulating the prediction of the fission product release in LOFT LP-FP-2 experiment, the "Xenon release fraction" under CORSOR modeling option is considered as output *y* of interest. MELCOR model nodalization is shown in Fig. 4. A total of $N = 100$ runs of the code have been done, each one with a different batch of input variables values sampled from the distributions listed in Table 1 [Hoseyni et al., 2014]. The very small number of samples has been chosen with the aim of challenging the proposed method: as we shall see in what follows, we will be able, by bootstrapping

B=100 times the N samples, to collect different replicates of the code output under the same DOE and to provide an effective input variables ranking as compared with other aggregation methods of literature that have been proposed for sensitivity analsysis [Di Maio et al., 2014].
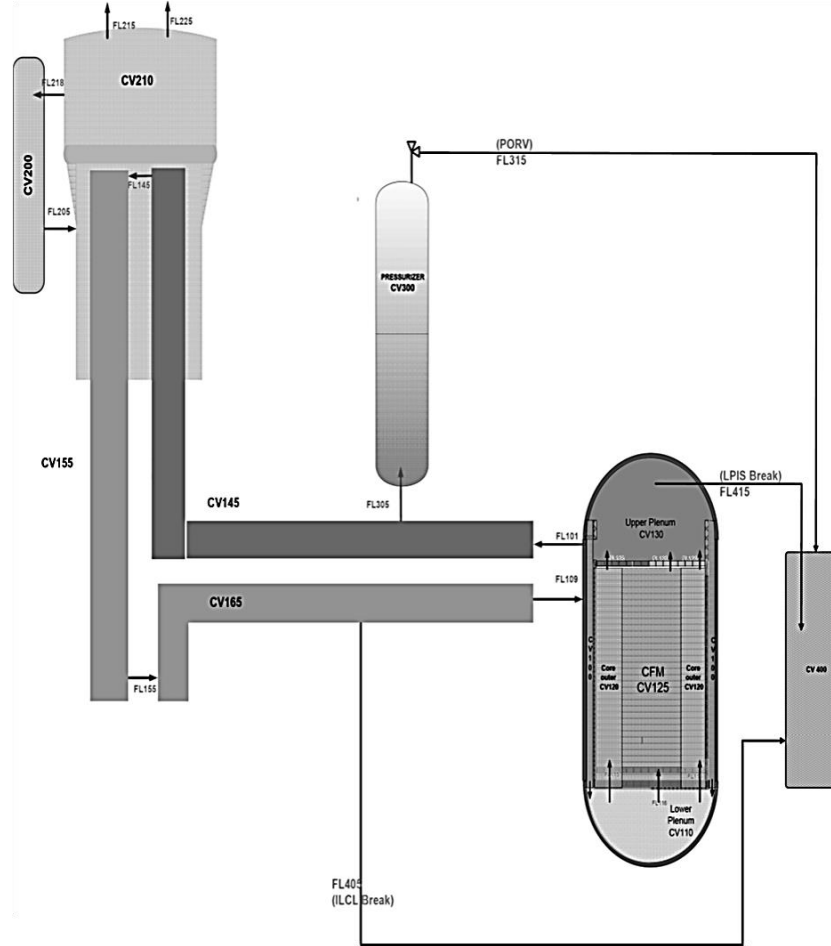

Fig. 4 MELCOR model nodalization for LOFT-LP-FP2 experiment

Table 1: *Uncertain Input parameters of LOFT MELCOR model and their distributions [Hoseyni et al., 2014]*

	Input variables	Description	Type of distribution	Mean value μ	Standard deviation σ
1.	POWER	Fission Power (MW)	normal	28.8E+6	0.467E+6
2.	SAREA-1	ILCL break flow area (FL415)	normal	0.000683	3.48469E-05
3.	SAREA-2	LPIS break flow area (FL405)	normal	0.000683	3.484E-05
4.	BDC-1	ILCL Break Discharge Coefficient (FL415)	normal	1	0.067
5.	BDC-2	LPIS Break Discharge Coefficient (FL405)	normal	1	0.067
6.	BLC-1	ILCL Break Loss Coefficient (FL415)	normal	1.005	0.507
7.	BLC-2	LPIS Break Loss Coefficient (FL405)	normal	1.005	0.507
8.	RVF1	Core radiation view factors	normal	0.03	0.005
9.	RVF2	Core radiation view factor 2	normal	0.22	0.01

	Input variables	Description	Type of distribution		
10.	ZMLT	Zircaloy Melt Temperature	normal	2200	51.02
11.	DSF	Dynamic Shape Factor (RNMS000)	normal	1	0.408
12.	ASF	Agglomeration Shape Factor (RNMS000)	normal	1	0.408
13.	SC7101-1	CORSOR parameter-1 C7101(1,1,1)	normal	900	5.102
14.	SC7101-2	CORSOR parameter-2 C7101(1,1,2)	normal	900	5.102
15.	SC7101-3	CORSOR parameter-3 C7101(1,1,4)	normal	900	5.1026
16.	SC7101-4	CORSOR parameter-4 C7101(2,1,1)	normal	1400	20.408
17.	SC7101-5	CORSOR parameter-5 C7101(2,1,2)	normal	1400	20.408
18.	SC7101-6	CORSOR parameter-6 C7101(2,1,4)	normal	1400	20.408
19.	SC7101-7	CORSOR parameter-7 C7101(3,1,1)	normal	2200	25.510
20.	SC7101-8	CORSOR parameter-8 C7101(3,1,2)	normal	2200	25.510
21.	SC7101-9	CORSOR parameter-9 C7101(3,1,4)	normal	2200	25.510
22.	SC7101-10	CORSOR parameter-10 C7101(1,2,1)	normal	7.02E-09	3.581E-10
23.	SC7101-11	CORSOR parameter-11 C7101(1,3,1)	normal	0.0088	0.00047
24.	SC7101-12	CORSOR parameter-12 C7101(1,2,2)	normal	7.525E-12	3.852E-13
25.	SC7101-13	CORSOR parameter-13 C7101(1,3,2)	normal	0.0142	0.000724
26.	SC7101-14	CORSOR parameter-14 C7101(1,2,4)	normal	7.02E-09	3.581E-10
27.	SC7101-15	CORSOR parameter-15 C7101(1,3,4)	normal	0.00886	0.000452
28.	SC7101-16	CORSOR parameter-16 C7101(2,2,1)	normal	0.000000202	1.03061E-08
29.	SC7101-17	CORSOR parameter-17 C7101(2,3,1)	normal	0.00667	0.000340306
30.	SC7101-18	CORSOR parameter-18 C7101(2,2,2)	normal	1.7019E-07	2.65357E-08
31.	SC7101-19	CORSOR parameter-19 C7101(2,3,2)	normal	0.00667	0.000340
32.	SC7101-20	CORSOR parameter-20 C7101(2,2,4)	normal	1.7019E-07	2.653E-08
33.	SC7101-21	CORSOR parameter-21 C7101(2,3,4)	normal	0.00667	0.00034
34.	SC7101-22	CORSOR parameter-22 C7101(3,2,1)	normal	0.0000174	8.877E-07
35.	SC7101-23	CORSOR parameter-23 C7101(3,3,1)	normal	0.0046	0.000234
36.	SC7101-24	CORSOR parameter-24 C7101(3,2,2)	normal	0.0000174	8.877E-07
37.	SC7101-25	CORSOR parameter-25 C7101(3,3,2)	normal	0.0046	0.000234
38.	SC7101-26	CORSOR parameter-26 C7101(3,2,4)	normal	0.0000174	8.87755E-07
39.	SC7101-27	CORSOR parameter-27 C7101(3,3,4)	normal	0.0046	0.000234
	Input variables	*Description*	*Type of distribution*	*Lower value*	*Upper value*
40.	OPHEIT-1	Core Flow Path Opening Heights (FL115)	uniform	0.01	1
41.	OPHEIT-2	Core Flow Path Opening Heights (FL116)	uniform	0.01	1
42.	SC1214	Turbulent Forced Convective Flow in Tubes	uniform	0.0115	0.0345
43.	SC1221	Laminar Free Convection between Parallel Vertical Surfaces	uniform	0.09	0.27
44.	SC1222	Turbulent Free Convection between Parallel Vertical Surfaces	uniform	0.0325	0.0975
45.	SC1231	Forced Convective Flow over a Spherical Particle	uniform	0.3	0.9
46.	SC1232	Free Convective Flow over a Spherical Particle	uniform	0.3	0.9
47.	SC1001-5	Zircaloy Oxidation Rate Constant Coefficients (T lower)	uniform	1843	1863
48.	SC1001-6	Zircaloy Oxidation Rate Constant Coefficients (T upper)	uniform	1864	1882
49.	OXTHICK	Minimum Oxide Shell Thickness (COR00008)	uniform	1E-6	2E-4

50.	TBLOCK	Time of blockage formation	uniform	1400	1600

5. Sensitivity analysis results for the LOFT LP-FP-2 model

Based on the $N = 100$ original code runs [Hoseyni et al., 2015], the histogram of Xenon release is shown in Fig. 5, where several maxima can be seen to be approximated by a multimodal analytical distribution. Therefore, in Fig. 5 it is also shown the analytical reconstruction (solid line) of the prior distribution of the Xenon release function, $f(y)$, that has been obtained by the FMM method, where the number K of clusters is set equal to 3 and the parameters values of the K Gaussian models have been obtained by Expectation Maximization algorithm [Carlos et al., 2013]. The choice of using three Gaussian distributions for the FMM parameters identification is the result of a trial and error procedure. An automatic optimization of the number K of distributions in the FMM is also possible [Figueiredo et al., 2002]. As a last remark, it is worth pointing out that the common distributions $q(x)$ used in this application are, instead, uniform distributions.
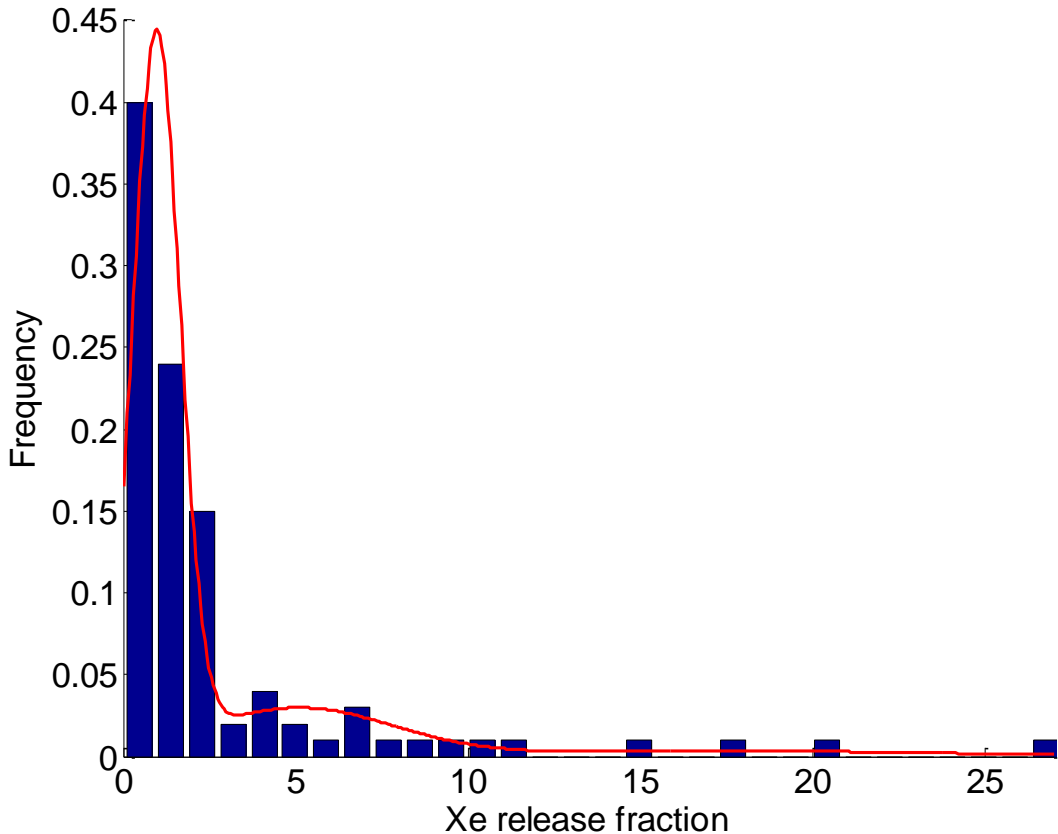

Fig. 5: *reconstructed PDF of the model output based on a three-cluster FMM (solid line) for the original code output dataset.*
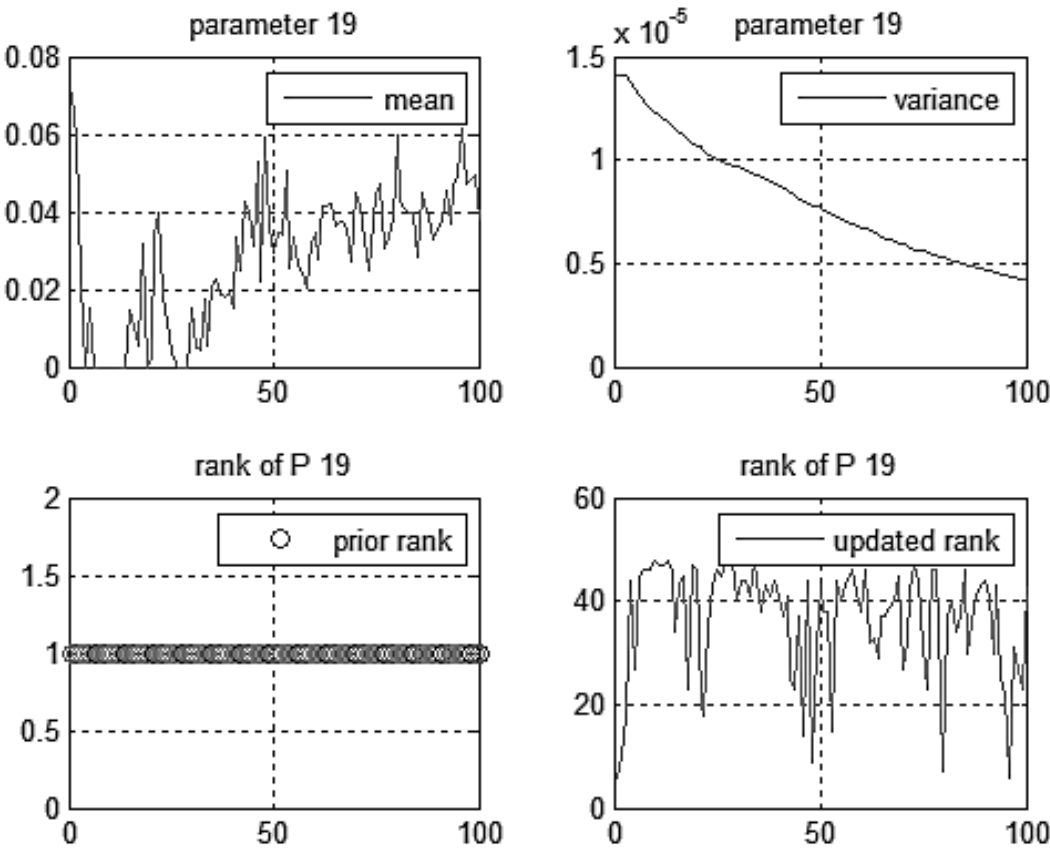
In order to obtain additional information to provide a realistic evaluation of the Xenon release fraction variability and, thus, an increase of the robustness of the sensitivity analysis results, the original Xenon dataset of code simulations has been replicated by Bootstrap to generate B datasets for a total of $B = 100$ bootstrap samples (Section 3.1).

For each b-th bootstrap sample, the analytical reconstruction of the N simulations replicates pdf $f(y)$ is performed by FMMs and, thus, the sensitivity measures are evaluated. The new values of sensitivity measures, derived from the b-th bootstrap, are used as new evidences for updating the state of knowledge about three sensitivity measures, i.e., the mean value μ_l and standard deviation σ_l of the sensitivity measures are updated into the posterior mean μ_l^* and variance σ_l^{2*} by the Bradley-Terry algorithm (Section 3.2).
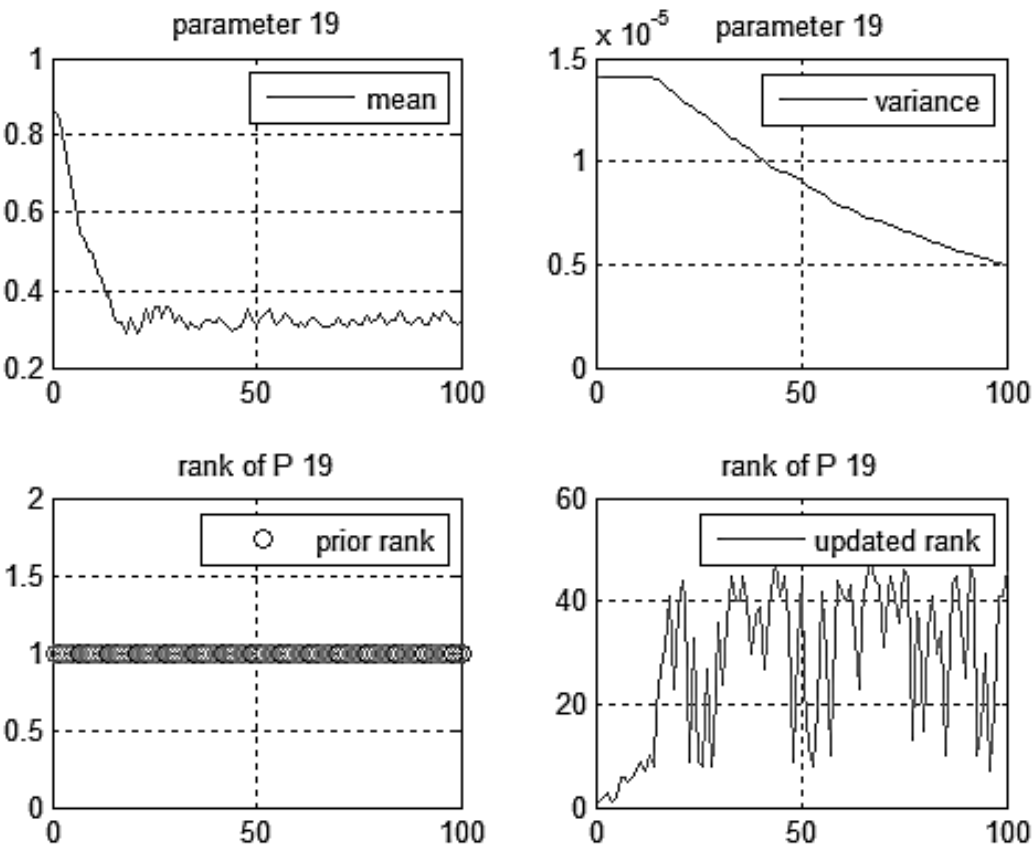
Without loss of generality, the updated mean μ_l^* (Eq. 15), updated variance σ_l^{2*} (Eq. 16) and the updated rank position for input variable $l = 19$ as a function of the b-th updating step are shown in Fig. 6. Whatever the sensitivity measure considered, Fig. 6 shows that:

1) the variance decreases as the updating procedure proceeds;
2) the mean and, consequently, the rank behave differently from the variance, and fluctuate.
3) the Hellinger distance and Kullback–Leibler divergence show larger stability than input saliency in the updating procedure (i.e., smaller ripples in the plot of the mean).

Input
Saliency

parameter 19

parameter 19

rank of P 19

rank of P 19

Hellinger

parameter 19

parameter 19
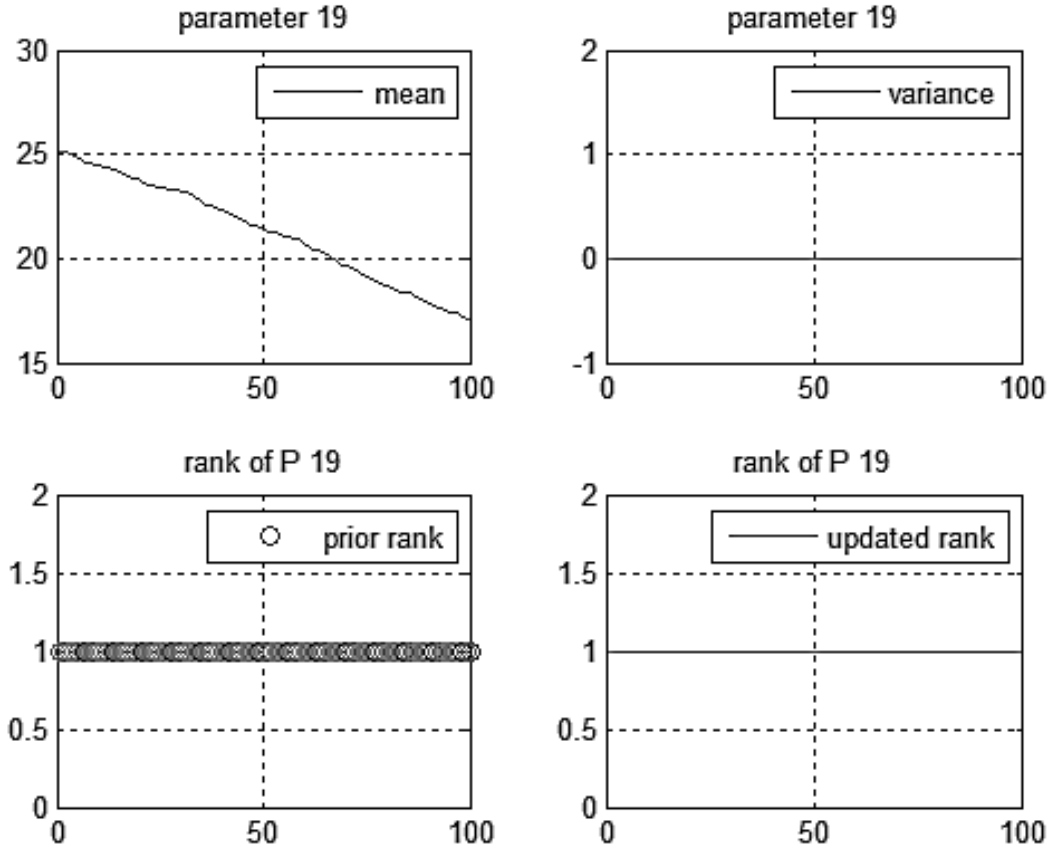
rank of P 19

rank of P 19

Kulback-Leibler

Fig. 6*: Updated mean, variance and rank for input variable 19 of Table 1*

We, then, perform the Bayesian ensemble by calculating the weighted sum of Eq. (32) with the weights listed in Table 2 (Section 3.3), where the acknowledged better stability of Hellinger distance and Kullback–Leibler divergence with respect to input saliency is rewarded with larger weights, that account, indeed, for the high-ranking stability (reproducibility) of the sensitivity measure during the updating process (as pointed out before in Section 3)).

Table 2: *Bayesian weights of SA measures for aggregation*

SA Measure	Assigned weight
ρ	0.0582
HD	0.4595
KL	0.4823

The results of the final Bayesian ensemble of sensitivity measures are reported in Table 3 where the first 20 ranking positions are compared by showing the input variables that would have been assigned

by the traditional sum (R_{sum}) and majority voting (R_{mv}) aggregation approaches [Kukkonen et al., 2007] that are the much easier but approximate methods for sensitivity analysis purposed in [Di Maio et al., 2014].

In fact, it can be seen that, the proposed Bayesian method, R_{sum} and R_{mv} approaches agree only on the two most important input variables (i.e., the Minimum Oxide Shell Thickness (OXTHICK) and the Agglomeration Shape Factor (ASF)), whereas they disagree on the other input variables. One might argue that R_{sum} or R_{mv} approaches still remain the best choices due to the fact that among themselves they agree on the first four positions, and to their lower computational burden. In fact, R_{sum} and R_{mv} are computed using only the *B*=1 original MELCOR output dataset and, thus, any further bootstrap replication is avoided.

However, even if the proposed Bayesian method uses B=100 bootstraps of the original MELCOR output dataset (i.e., the computational time is B times larger than for R_{sum} and R_{mv}), the Bayesian ensemble aggregation is hereafter shown to be preferable because it holds two fundamental properties of Bayesian processes: i) the memory property and ii) the exchangeability property.

The memory property entails a prior belief (e.g., the value of sensitivity measures obtained for each input variable from the original FMM) to be differently updated depending on the quantity and quality of the available evidence (e.g., the new bootstrap replicates) used for computing the posterior [Friedman et al., 1997]. Consequently, different batches of evidences lead to different posterior distributions (e.g., different sensitivity measures values and ranking positions). Figure 7 shows that the memory property holds in our case: two different batches of evidences of t = 20, 40 successive bootstrap replicates, respectively, are used. Dashed line with circles in Fig. 7 shows the updated aggregated sensitivity measures for input variable 19 for t = 20 successive updates, whereas dotted line with squares are those for t = 40 updates: a perfect matching in the updating process of the aggregated sensitivity measures is obtained using the first and the second batch up to the 20-th bootstrap, due to the use of the same cumulated evidence. Contrarily, assuming that the bootstrap replicates are not available up to t = 20 and that we start calculating from t = 21 up to t = 40, the aggregated sensitivity measures at time t = 40 differ from those of the batch that exploits all the 40 bootstrap replicates, due to the lack of the first 20 evidences (continuous line with triangles in Fig. 7).

Table 3: *Comparison of the input variables ranking position provided by different approaches*

Rank	Bayesian	R_{sum}	R_{mv}
1	OXTHICK	OXTHICK	OXTHICK
2	ASF	ASF	ASF
3	SAREA-1	C7101(3,3,2)	C7101(3,3,2)

4	SC1214	BDC-1	BDC-1
5	C7101(3,3,2)	SAREA-1	SC1221
6	SC1221	BDC-2	C7101(1,3,1)
7	DSF	SC1221	C7101(3,1,4)
8	C7101(3,2,2)	C7101(1,3,1)	C7101(1,1,1)
9	C7101(1,1,2)	C7101(3,3,1)	-
10	C7101(1,31)	C7101(1,2,4)	-
11	C7101(2,3,1)	C7101(2,3,2)	-
12	C7101(1,1,4)	C7101(3,1,4)	-
13	RVF1	C7101(2,2,1)	-
14	SAREA-2	C7101(2,2,2)	-
15	C7101(2,2,4)	C7101(1,1,4)	-
16	C7101(1,3,2)	DSF	-
17	C7101(3,1,4)	C7101(1,1,1)	-
18	C7101(2,3,4)	C7101(3,2,1)	-
19	RVF2	RVF2	-
20	BLC-2	7101(1,3,4)	-

The exchangeability property entails that the order of the information fed to the Bayesian updating process is irrelevant to the results (e.g., if the order of $B = 100$ bootstrap replicates is shuffled to updating process leads, independently from the order of the evidences, to the same final ranking). Again for input variable 19, the ensemble of sensitivity measures calculated with the $B = 100$ and the same shuffled replicates, differ for less than 1%, which proves that the proposed process possesses the exchangeability property.

This means that the Bayesian ensemble aggregation has to be preferred to Rsum and Rmv because it is robust to the variability of the experiments results to be used for sensitivity analysis as it allows cumulating knowledge (memory property) irrespectively of the occurrences sequence of the experiment results (exchangeability property).
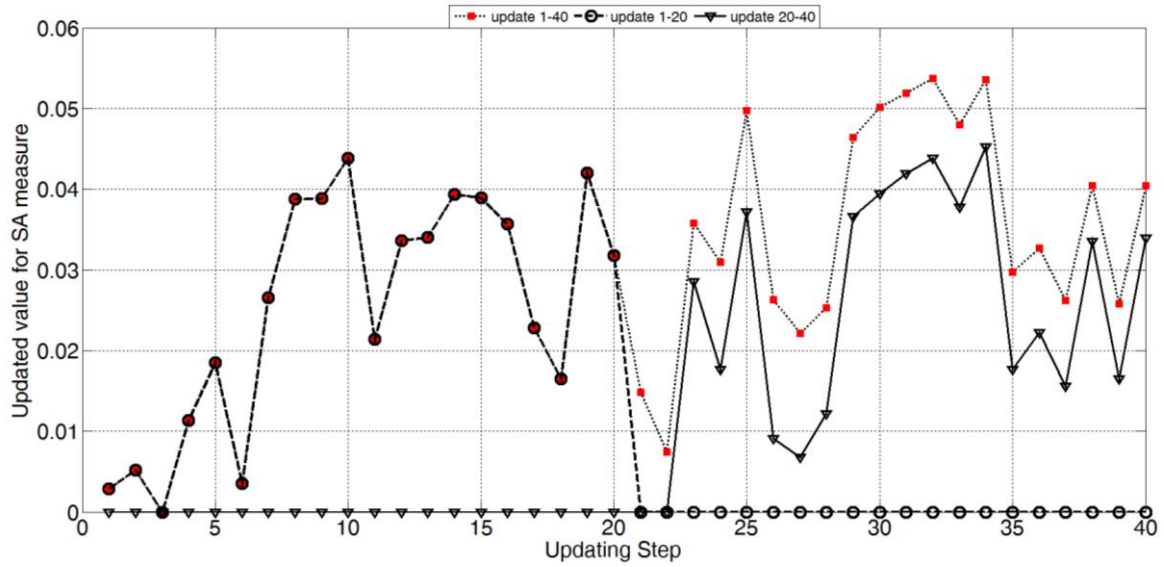
Fig. 7 *Memory property of updating*

6 Conclusion

In this paper, a novel framework is presented for performing an ensemble sensitivity analysis based on the Bayesian updating procedure. The framework is built based on performing an uncertainty analysis on the severe accident model for estimating a Gaussian FMM to retrieve the analytical pdf of the model output with as few simulations as possible. A bootstrap technique is adopted to replicate *B* new model output datasets. An innovative ensemble strategy is designed to aggregate three sensitivity methods, namely input saliency, Hellinger distance and Kullback–Leibler divergence, that exploits a Bayesian updating procedure based on the Bradley-Terry algorithm for an incremental learning and iterative update of the input variables sensitivity measures. An application is shown with regards to a long-running MELCOR code simulating the fission product release during a LOCA in a scaled-down PWR. The results show the capability of the proposed framework in discerning between influent and negligible input variables, with properties of memory and exchangeability inherent in the Bayesian updating process.

Nomenclature

COV	Coefficient Of Variation
DSA	Deterministic Safety Assessment
DOE	Design of Experiment
ECCS	Emergency Core Cooling System
EM	Expectation Maximization
FMM	Finite Mixture Model
FOM	Figure of Merit
FP	Fission Products

INL	Idaho National Laboratory
LOCA	Loss of Coolant Accident
LOFT	Loss of Flow Test
LPIS	Low Pressure Injection System
MV	Majority voting
NPP	Nuclear Power Plant
pdf	Probability density function
PWR	Pressurized Water Reactor
RF	Release fraction
RN	Radio Nuclide
SA	Sensitivity Analysis
SNL	Sandia National Laboratories
TH	Thermal-hydraulics

References

[Baraldi et al., 2011] P. Baraldi., E. Zio, G. Gola, D. Roverso, M. Hoffmann, "*Two novel procedures for aggregating randomized model ensemble outcomes for robust signal reconstruction in nuclear power plants monitoring systems*", Annals of Nuclear Energy 38, 212–220, 2011.

[Borgonovo, 2007] Borgonovo, E., "*A new uncertainty importance measure*", Reliability Engineering and System Safety 92, 771–784, 2007.

[Cadini et al., 2007] F. Cadini, E. Zio, F. Di Maio, V. Kopustinkas, R. Urbonas, "*A Neural-network-based Variance Decomposition Sensitivity Analysis*", International Journal of Nuclear Knowledge Management, Vol. 2, No. 3, pp. 299-312, 2007.

[Carlos et al., 2013] Carlos, S., Sánchez, A., Ginestar, D., Martorell, S., "*Using finite mixture models in thermal hydraulics system code uncertainty analysis*", Nuclear Engineering and Design 262, 306–318, 2013.

[Diaconis et al., 1982] P. Diaconis, Sandy L Zabell, "*Updating subjective probability*", Journal of the American Statistical Association 77 (380), 822–830., 1982.

[Di Maio et al., 2012] F. Di Maio, J. Hu, W. Peter, M. Pecht, K. Tsui, E. Zio, "*Ensemble approaches for clustering health status of oil sand pumps*",Expert Systems with Applications 39, 4847–4859, 2012.

[Di Maio et al., 2014] F. Di Maio, G. Nicola, E. Zio, Y. Yu, "*Ensemble-based sensitivity analysis of a Best Estimate Thermal Hydraulics model: Application to a Passive Containment Cooling System of an AP1000 Nuclear Power Plant*", Annals of Nuclear Energy, vol. 73, p. 200–210, 2014.

[Di Maio et al., 2015] F. Di Maio, G. Nicola, E. Zio, Y.Yu, "*Finite Mixture Models for sensitivity analysis of Thermal Hydraulic Codes for Passive Safety Systems safety Analysis*", Nuclear Engineering and Design, 289, 144–154.

[Efron et al., 1993] B. Efron and T. R.J., "*An Introduction to Bootstrap*", New York: Chapman & Hall/CRC, 1993.

[Figueiredo et al., 2002] M. Figueiredo, A.K. Jain, "*Unsupervised learning of finite mixture models*", IEEE Transactions on Pattern Analysis and Machine Intelligence 42 (3), 1–16, 2002.

[Friedman et al., 1997] Friedman, N., Goldszmidt, M., "*Sequential update of Bayesian network structure*", UAI'97 Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence, 165-174, 1997.

[Gibbs et al., 2002] A. L. Gibbs, Edward Su, "*On choosing and bounding probability metrics*", International Statistical Review 70 (3), 419–435, 2002.

[Helton, 1993] Helton, J.C., "*Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal*", Reliability Engineering and System Safety 42, 327–367, 1993.

[Hoseyni et al., 2014] S. M. Hoseyni, M. Pourgol-Mohammad, A. Abbaspour Tehranifard , F. Yousefpour, "*A Systematic Framework for Effective Uncertainty Assessment of Severe Accident Calculations; Hybrid Qualitative and Quantitative Methodology*," Reliability Engineering and System Safety, vol. 125, pp. 22-35, 2014.

[Hoseyni et al., 2015] S. M. Hoseyni, M. Pourgol-Mohammad, "Model uncertainty in severe accident calculations: a structural methodology with application on LOFT LP-FP-2 experiment," Journal of Nuclear Technology, under review, 2015.

[Hunter, 2004] Hunter, D.R., "*MM algorithms for generalized Bradley-Terry models*", Annals of Statistics, 32 (1), pp. 384-406, 2004.

[Kmetyk, 1992] L. Kmetyk, "*MELCOR 1.8.1 Assessment: LOFT Integral Experiment LP-FP-2*", SAND92-1373, 1992.

[Kukkonen et al., 2007] S. Kukkonen, J. Lampinen, "*Ranking-dominance and many-objective optimization*", Evolutionary Computation, 3983–3990, 2007.

[IRSN-200783, 2007] IRSN, "*Research and development with regard to severe accidents in pressurized water reactors; Summary and outlook*", IRSN-200783, 2007.

[Law et al., 2004] Martin H.C Law, Mario A.T. Figueiredo, Anil K. Jain, "*Simultaneous feature selection and clustering using mixture models*", IEEE Transactions on Patterns Analysis and Machine Intelligence 26 (9), 1154–1166, 2004.

[Lewis et al., 2008] D. R. Lewis, B.J., F. Iglesias, Ducros and T. Kudoe, "*Overview of experimental programs on core melt progression and fission product release behaviour*", Jouranl of Nuclear Material, vol. 2008, pp. 126-143, 2008.

[Liu et al., 2010] Liu, Q., Homma, T., "*A new importance measure for sensitivity analysis*", Journal of Nuclear Science and Technology, 47 (1), pp. 53-61, 2010.

[MacKay, 1995] MacKay, M.D.," "Evaluating Prediction Uncertainty", Los Alamos National Laborator Technical Report, NUREG-GE-6311, 1995.

[McLachlan et al., 2000] G. McLachlan and D. Peel, "*Finite Mixture Models*", New York: John Wiley & Sons Inc., 2000.

[NEA, 2001] OECD/NEA, "*In-Vessel Core Degradation Code Validation Matrix Update 1996-1999*", NEA/CSNI/R(2000)21, 2001.

[Pengfei et al., 2014] W. Pengfei, Lu Zhenzhou and R. Wenbin S. Jingwen, "*Regional Sensitivity Analysis Using Revised Mean and Variance Ratio Functions*", Reliability Engineering & System Safety, Vol. 121, pp. 121-135. Elsevier, 2014.

[Plischke et al., 2013] Plischke E., E. Borgonovo , C. L. Smith, "*Global sensitivity measures from given data*", European Journal of Operational Research, vol. 226, p. 536–50, 2013.

[Pourgol-Mohammad et al., 2007] Pourgol-Mohammad, M., Modarres, M., and Mosleh, A.,"*Integrated Methodology for Thermal-Hydraulic Code Uncertainty Analysis with Application*", Nuclear Technology, vol. 165, pp. 333-359, 2007.

[Pourgol-Mohammad, 2009] Pourgol-Mohammad, M. "*Thermal-hydraulics system codes uncertainty assessment: a review of the methodologies*", Annals of Nuclear Energy 36, 1774–1786, 2009.

[Probst et al., 2006] J. Joucla, P. Probst, "*Rank Statistics and Bootstrap: A More Precise Evaluation of the 95th Percentile in Nuclear Safety LB-LOCA Calculations*", 14th Int. Conf. Nuclear Engineering (ICONE 14), Miami, Florida, 2006.

[Pudil et al., 1995] P. Pudil, J. Novovicova, J. Kittler, "*Feature selection based on the approximation of class densities by finite mixtures of the special type*", Pattern Recognition 28 (9), 1389–1398, 1995.

[Saltelli et al., 1999] Saltelli, A., Tarantola, S., Chan, K.P.-S., "*A quantitative model-independent method for global sensitivity analysis of model output*", Technometrics, 41 (1), pp. 39-56, 1999.

[Saltelli, 2002] Saltelli, A., "*Sensitivity analysis for importance assessment*", Risk Analysis, 22 (3), pp. 579-590, 2002.

[Secchi et al., 2008] P. Secchi, E. Zio and F. Di Maio, "*Quantifying Uncertainties in the Estimation of Safety Parameters by Using Bootstrapped Artificial Neural Networks*", Annals of Nuclear Energy, Vol. 35, pp. 2338-2350. Elsevier, 2008.

[Sobol et al., 1995] Saltelli, A., Sobol, I.M., "About the use of rank transformation in sensitivity analysis of model output", Reliability Engineering and System Safety, 50 (3), pp. 225-239, 1995.

[Vaithyanathan et al., 1999] S. Vaithyanathan, B. Dom, "*Generalized model selection for unsupervised learning in high dimensions*", Advances in Neural Information Processing Systems. MIT Press, Cambridge, Mass., pp. 970–976, 1999.

[Weng et al., 2011] R. C. Weng and C.-J. Lin, "A Bayesian Approximation Method for Online Ranking", Journal of Machine Learning Research, vol. 12, pp. 267-300, 2011.

[Zio et al., 2008] E. Zio and F. Di Maio, "*Bootstrap and Order Statistics for Quantifying Thermal-Hydraulic Code Uncertainties in the Estimation of Safety Margins*", Science and Technology of Nuclear Installations, Volume 2008 (2008), Article ID 340164, 9 pages, doi:10.1155/2008/340164.