Bootstrapped-ensemble-based Sensitivity Analysis of a trace thermal-hydraulic model based on a limited number of PWR large break loca simulations

Francesco Di Maio, Alessandro Bandini, Enrico Zio, Sofia Carlos Alberola, Francisco Sanchez-Saez, Sebastián Martorell

BOOTSTRAPPED-ENSEMBLE-BASED SENSITIVITY ANALYSIS OF A TRACE THERMAL-HYDRAULIC MODEL BASED ON A LIMITED NUMBER OF PWR LARGE BREAK LOCA SIMULATIONS

Francesco Di Maio[1], Alessandro Bandini[1], Enrico Zio[1,2], Sofia Carlos Alberola[3], Francisco Sanchez-Saez[3], Sebastián Martorell[3]

[1]*Energy Department, Politecnico di Milano, Milan, Italy*
[2]*Chair on System Science and Energetic Challenge, European Foundation for New Energy – EDF, École Centrale de Paris and Supélec, Paris, France*
[3]*Universitat Politècnica de València, Valencia, Spain*

Corresponding Author: Dr. Francesco Di Maio
Tel: +390223996372, Fax: +390223996309, Email: francesco.dimaio@polimi.it

Abstract

The safety verification of nuclear systems can be done by analyzing the outputs of Best-Estimate Thermal-Hydraulic (BE-TH) codes, which allow predicting the system response under safe and accidental conditions with greater realism as compared to conservative TH codes. In this case, it is necessary to quantify and control the uncertainties in the analysis, which affect the estimated safety margins. This can be achieved by Sensitivity Analysis (SA) and Uncertainty Analysis (UA) techniques tailored to handle the large computational costs of TH codes. This work presents an Ensemble-Based Sensitivity Analysis (EBSA) based on Finite Mixture Model (FMM) as an effective solution to keep low the code runs and handle the uncertainty in the SA methods. The approach proposed is challenged against a situation of a very low number of code runs: the Bootstrap method is, then, used in support. Three different strategies based on EBSA and Bootstrap are set forth (i.e., bottom-up, all-out and filter strategies). An application is provided with respect to a Large Break Loss of Coolant Accident (LBLOCA) simulated by a TRACE model of the Zion 1 Nuclear Power Plant (NPP).

Keywords: *Safety Margins; Uncertainty and Sensitivity Analysis; Finite Mixture Model; Ensemble of Methods; Bootstrap Method; Large Break Loss Of Coolant Accident.*

1. INTRODUCTION

Safety of Nuclear Power Plants (NPPs) is verified by accurate Thermal-Hydraulic (TH) models that reproduce the system functional response in normal and accidental conditions. Traditionally, conservative calculations of a small set of pre-defined accidental scenarios (i.e., Design Basis Accidents) are made and few safety-significant parameters (i.e., fuel cladding

temperature, containment pressure, etc.) are compared with predetermined safety thresholds. The (positive) differences between the prescribed safety thresholds and the conservatively computed safety parameters values are the so-called safety margins [1].

Recently, a more realistic approach has been advocated, whereby the calculations of the plant safety margins are based on Best-Estimate (BE) TH models, within a probabilistic framework that allows accounting for the uncertainties in the model and its parameters [2]. Indeed, proper quantification and control of the uncertainties affecting the best-estimated safety margins are necessary conditions [3-7].

Several methods have been proposed for the quantification of TH codes uncertainties, e.g., Code Scaling, Applicability and Uncertainty (CSAU) [8-10], the State-of-the-Art Reactor Consequence Analyses (SOARCA) project [11-14], the Automated Statistical Treatment of Uncertainty Method (ASTRUM) and Integrated Methodology for Thermal Hydraulics Uncertainty Analysis (IMTHUA) [15], and for the uncertainty analysis of a reactor physics codes, e.g., by the introduction of surrogate models trained on a limited number of calculations to limit the computational burden, like in [16]. In [17], some of the authors have proposed to combine Order Statistics (OS) [18; 19] with Artificial Neural Networks (ANN) to speed up the calculation for uncertainty analysis by reducing the computational cost associated to the repeated simulations of a TH-BE code.

The methods listed above are limited in that they do not provide the whole distribution of the model response, nor indications on the sensitivity of the model to the inputs variability [20; 21].

With respect to the latter issue, Sensitivity Analysis (SA) is the way for determining how the uncertainty in the model output is apportioned among the model inputs uncertainties. In other words, SA allows finding out which input variables most influence the model output and, thus, in our case, are most relevant for the safety margins quantification.

The manifold of SA techniques presented in the literature can be sorted into three main categories: local, regional and global.

Local and regional sensitivity analyses consider inputs variations on subsets of their overall ranges. Local methods evaluate the effects on the system response of small perturbations of the model input variables in the neighbourhood of some fixed, nominal values, at low computational costs [22; 23]. Thus, local SA provides information about the sensitivity of model output to the inputs variability only at some fixed points. Regional SA methods focus on the contribution of the inputs ranges of variability to the uncertainty of model output [24; 25]. But, they do not give a complete representation of the uncertainty of the model, in terms

of distribution [17].

Global sensitivity analysis (GSA) methods explore the whole distribution range of the model inputs and the effects of their mutual combination. Examples of GSA methods are Response Surface Methodology (RSM), Fourier Amplitude Sensitivity Test (FAST), Delta and Variance Decomposition Method and various concepts based on the functional ANOVA expansion of the input-output mapping [26-30]. Resorting to GSA is most desirable when dealing with complex codes, thanks to its ability to handle non-linear and non-monotone models [31-33]. On the other hand, GSA methods can be computationally expensive if not based on a limited number of TH simulations [25; 34; 35].

In the present work, we consider three GSA methods (Input Saliency (IS) [36], Hellinger Distance (HD) [37; 38], Kullback-Leibler Divergence (KLD) [37; 38]) to form a bootstrapped ensemble for: i) exploiting the capability of global SA methods to provide knowledge on the sensitivity of model output uncertainty to the entire inputs distributions ranges and ii) limiting their potentially large computational burden. Indeed, when the TH calculation data are limited, an Ensemble-Based Sensitivity Analysis (EBSA) has been shown to give satisfactory results with limited computational cost [39].

In this work, we resort, in particular, to a bootstrapped ensemble-based approach. The basic idea herein set forth is to rely on the information available in the multi-modal Probability Density Function (PDF) of the bootstrapped model output in order to perform GSA of a TH code. Then, the amount of code calls that are demanded by the EBSA is reduced as compared to standard GSA techniques, since the amount of simulations needed is simply that required to reconstruct the estimated model output PDF [39]. In this paper, this is achieved by means of a Finite Mixture Model (FMM) [40], whereby the natural clustering generated by the FMM on the TH code output is exploited to estimate global sensitivity measures [38; 41]. The ensemble of three SA indicators is, then, employed for ranking the input factors which most affect the output uncertainty.

The EBSA paradigm enables to combine the output of the three different SA methods and generate reliable and robust rankings, compensating the individual biases of each method, which shows to be particularly effective when the number of TH code runs available is small [42].

The Bootstrap method is used to (artificially) increase the amount of data obtained from the few TH code runs, without altering the original information therein contained [17]. It is a distribution-free inference method, which relies on no prior knowledge about the distribution function of the underlying population [17]. The idea is to generate alternative populations by

sampling with replacement from the original dataset.

Three strategies combining EBSA and the Bootstrap method are here proposed and compared with each other. A bottom-up strategy (ES1) computes a ranking order of the input variables out of each bootstrapped dataset and combines them a posteriori to generate a final aggregated ranking order. An all-out strategy (ES2) merges a priori the information from the bootstrapped datasets into three ranking orders that are, then, combined together. A filter strategy (ES3) computes the expected value and variance of the importance of each input variable over all the bootstrapped datasets.

These strategies have been developed and tested on a real case study that entails a Large Break Loss of Coolant Accident (LBLOCA) [43; 44], which is simulated by the TRACE code [45] of the Zion 1 NPP [43; 44]. The results obtained have been compared with other standard methods of literature.

The rest of the paper is organized as follows. Section 2 briefly recalls the three SA techniques used and explains how the Bootstrap-based information is processed for multiple ranking aggregation. In Section 3, the three ensemble strategies (i.e., bottom-up, all-out and filter) are defined. Section 4 is devoted to the presentation of the case study. Section 5 shows the results of the application of the proposed strategies to the data of the case study. Section 6 presents the comparison of the results obtained to those of other methods of literature. Conclusions are drawn in Section 7.

2. BOOTSTRAPPED EBSA

In the originally proposed EBSA [35], the output values of the original set of TH code runs is processed by a set of SA methods directly and, then, input ranking aggregation is performed. However, when the number of available TH code runs is small, the results of the sensitivity analysis are likely to be biased and the final aggregated ranking not realistic.

Bootstrap (described in Section 2.1) is expected to help improve the reliability of the final aggregated ranking [17], by creating purpose-built bootstrapped datasets out of the initial dataset available.

2.1 Bootstrapped-replications of TH Code Runs

In practice, every bootstrap replication of the available TH code data is generated by sampling with replacement from the original dataset. For the sake of clarity and without loss

of generality, let us assume that the original dataset consists in a bi-dimensional matrix D with N_R rows, where N_R is the number of performed code runs, and $n+o$ columns, where n is the number of code inputs and o is the number of output variables. Every newly generated bootstrapped-dataset is, thus, the assembly of N_R rows, which are randomly sampled with replacement from matrix D [17].

2.2 Ensemble-based Sensitivity Analysis

Let y denote the output of the TH model m:

$$y = m(x_1, \ldots, x_j, \ldots, x_n), \quad j = 1,2,\ldots,n \tag{1}$$

where x_j is the *j*-th model input variable with probability distribution $f(x_j \mid \theta_j)$ charachterised by a set of parameters θ_j. The random output variable y follows (i.e., can be expanded into) a finite number of K PDF models that are merged into one PDF $f(y)$ if [40]:

$$f(y) = \sum_{k=1}^{K} \pi_k f_k(y \mid \theta_k) \tag{2}$$

where $f_k(y \mid \theta_k)$ are K different PDFs (for example, Gaussian distributions), θ_k is the set of parameters of the *k*-th model of the mixture (for example, mean and standard deviation of the Gaussian distribution) and π_k are the mixing probabilities that fulfil:

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

where $\pi_k \geq 0$, $\forall k$.

Mixture models parameters θ and π are computed by Expectation Maximization (EM) algorithm, for optimal approximation of the model output PDF through the (small number) N_R of TH code simulation outcomes (the interested reader may refer to [41; 46]).

"Natural" clusters can be associated with each model PDF $f_k(y \mid \theta_k)$ of the mixture

density $f(y)$ [47]. Such clusters are called "natural" because they reflect the way the TH input variables combine in the model *m* to provide the output *y*, according to the physics (i.e., the "nature") that is reproduced in the code. Some clusters might be representative of input variables values that keep the system in normal operative conditions, whereas some others of values leading the system into accidental conditions. Such clusters are, then, exploited for calculating three SA measures: Input Saliency (IS), Hellinger Divergence (HD), Kullback-Leibler Distance (KLD). For EBSA [39], the results of each SA method are eventually aggregated.

For the sake of completeness, the basics of the three used SA methods are briefly summarized below.

2.2.1 Input Saliency

Once the model output PDF $f(y)$ is reconstructed by means of the K PDF models $f_k(y|\theta_k)$ as in Eq.(2), and assuming that all the x_j model inputs are mutually independent, $f(y)$ can be written as follows:

$$f(y) = \sum_{k=1}^{K} \pi_k f_k(y|\theta_k) = \sum_{k=1}^{K} \pi_k m\left(\prod_{j=1}^{n} (\rho_j f_k(x_j|\theta_j) + (1-\rho_j) q(x_j|\varsigma_j)) \right) \quad (4)$$

where the model *m* is fed with a realization of each *j*-th input variable taken from $f_k(x_j|\theta_j)$ that is the PDF of the *j*-th input variable inside the "natural" *k*-th cluster (i.e., the overlapping area between the the *j*-th input variable and the *k*-th model of the output variable distributions, that is large when the distributions have similar parameters θ_j and θ_k, respectively), to which corresponds $q(x_j|\varsigma_j)$ that is the so-called "common" PDF that the *j*-th input variable follows when it does not affect the distribution of the *k*-th cluster [48; 49] (e.g., when the *j*-th input variable follows an Uniform distribution whereas the output has a Gaussian distribution), properly weighted by ρ_j, that is the saliency index that measures the importance of input x_j in affecting the shape of the distribution of y. The values of ρ_j, *j*=1,2,…, *n*, that solve Eq.(4) can be estimated resorting to an EM algorithm as the one presented in [36].

Basically, if ρ_j is large, then the *j*-th input variable is relevant in affecting the output variability (i.e., the *j*-th input variable has the same distribution of at least one of the *K* models

of the FMM representing the of the output variable distribution), whereas in the opposite case, i.e. ρ_j is small, the input variable follows the so-called common probability distribution throughout all the clusters and it does not contribute in shaping the output y.

2.2.2 Hellinger Distance

Based on the formulation (4) of $f(y)$, the Hellinger Distance $H_{k,j}$ quantifies the gap length (i.e., the functional similarity) between the *j*-th input PDF $f_k(x_j \mid \theta_j)$ and its associated common distribution $q(x_j \mid \varsigma_j)$. One way to define $H_{k,j}$ is the following [37; 38]:

$$H_{k,j} = \left(\frac{1}{2} \int_{-\infty}^{+\infty} \left(\sqrt{f_k(x_j \mid \theta_j)} - \sqrt{q(x_j \mid \varsigma_j)} \right)^2 dx_j \right)^{\frac{1}{2}} \tag{5}$$

For example, if we were to assume $q(x_j \mid \varsigma_j)$ to be an Uniform distribution and $f_k(x_j \mid \theta_j)$ a Gaussian distribution, $H_{k,j}$ would be large because their functional difference is large; on the contrary, if both $q(x_j \mid \varsigma_j)$ and $f(x_j \mid \theta_j)$ were Uniform (due to the fact that the *j*-th input variable does not affect the characteristics of the *k*-th distribution), then $H_{k,j}$ would be small. Overall, the importance index of the *j*-th input variable measuring how much it affects the model output y (i.e., how much its PDF distribution is similar to the output PDF distribution) is given by the sum of the index values $H_{k,j}$ over the whole set of K models:

$$HD_j = \sum_{k=1}^{K} H_{j,k} \tag{6}$$

2.2.3 Kullback-Leibler Divergence

Along the lines of the Hellinger Distance, the Kullback-Leibler SA method strives to identify a disparity in the information carried by the two PDFs of the *j*-th input $f_k(x_j \mid \theta_j)$ and $q(x_j \mid \varsigma_j)$ by calculating:

$$KLD_{j,k} = \int_{-\infty}^{+\infty} f_k(x_j \mid \theta_j) \log\left(\frac{f_k(x_j \mid \theta_j)}{q(x_j \mid \varsigma_j)} \right) dx_j \tag{7}$$

The Kullback-Leibler Divergence is a measure of the different location of the masses of the two probability distributions [37; 38]: the closer to 0, the more similar the information brought by the two PDFs (i.e., the log(·) term approaches 0), that is to say the the *j*-th input follows its "common" PDF rather than the PDF of the *k*-th "natural" cluster.

Similarly to the Hellinger Distance importance index, the *j*-th input relevance in affecting the model output y is described by:

$$KLD_j = \sum_{k=1}^{K} KLD_{j,k} \tag{8}$$

2.3 Multiple Rankings Aggregation

The inputs rankings provided by Input Saliency, Hellinger Distance and Kullback-Leibler Divergence (Sections 2.2.1, 2.2.2 and 2.2.3, respectively) can be aggregated. In this work, two aggregation methods are used, i.e. the Borda and Schulze ranking aggregation methods [50].

The Borda ranking method computes the so-called Borda count of any input variable [48]. The Borda count (BC_j) of input variable x_j is the sum of its positions in all input rankings. Denoting by $p_{j,\vartheta}$ the *j*-th variable order inside the ϑ-th ranking, the Borda count for the input variable x_j is given by:

$$BC_j = \sum_{\vartheta=1}^{n_r} p_{j,\vartheta} \tag{9}$$

where n_r is the total number of rankings; a small value of BC_j means that the *j*-th input variable is among the most important, top-ranked, input variables.

The Schulze method follows the Condorcet criterion [50]: for each pair x_ε and x_ϕ of input variables (where $\varepsilon = 1,2,\ldots,n$, $\phi = 1,2,\ldots,n$, $\varepsilon \neq \phi$), the method counts how many methods rank x_ε above x_ϕ and, conversely, how many methods do the opposite: if the former amount is larger, then x_ε is considered dominant over x_ϕ, and viceversa.

3. BOOTSTRAPPED-EBSA STRATEGIES

As introduced in Section 2.2, ρ_j, HD_j and KLD_j are the SA indexes for ranking the model input variables. Our original idea is to compute them for bootstrapped datasets. Then, $\rho_{i,j}$, $HD_{i,j}$ and $KLD_{i,j}$ are input saliency, Hellinger Distance and Kullback-Leibler Divergence importance indexes of the *j*-th input variable ($j = 1,2,\ldots,n$, where n is the number of model inputs), calculated on the *i*-th bootstrapped dataset BD_i ($i = 1,2,\ldots,n_B$):

$$BD_i \rightarrow \begin{cases} \overline{\rho}_i = \left[\rho_{i,1},\ldots,\rho_{i,n}\right] \\ \overline{HD}_i = \left[HD_{i,1},\ldots,HD_{i,n}\right] \\ \overline{KLD}_i = \left[KLD_{i,1},\ldots,KLD_{i,n}\right] \end{cases} \tag{10}$$

Three strategies are explored to combine EBSA and Bootstrap in order to aggregate the bootstrapped rankings into a final model inputs rank.

3.1 The Bottom-Up Strategy (ES1)

Each *i*-th bootstrapped dataset (BD_i) is treated separately from the others to generate one ensemble-aggregated input ranking and, then, the set of input rankings obtained from all the bootstrapped datasets is processed a posteriori to give the final (bottom-up) aggregated ranking order (Figure 1).

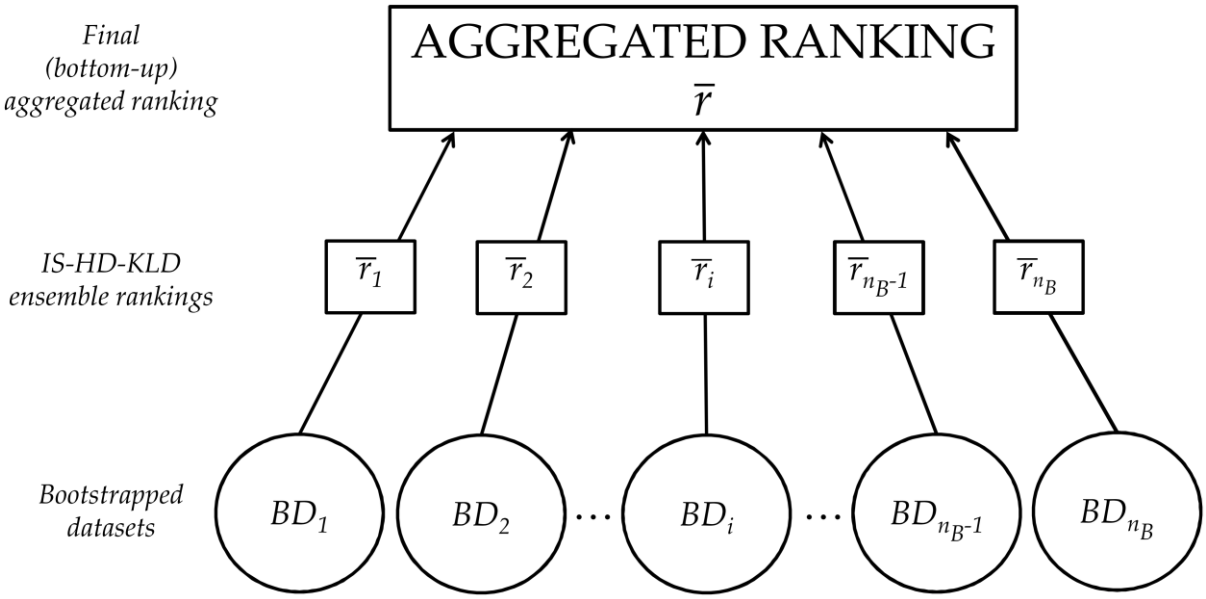

Figure 1: *scheme of the proposed bottom-up strategy (ES1).*

Practically, from the *i*-th bootstrapped-dataset (BD_i) the measures $\bar{\rho}_i$, $\overline{HD}_i$ and $\overline{KLD}_i$, are computed and three input ranking orders are correspondingly obtained by sorting in ascending order. Borda and Schulze methods (Section 2.3) are, then, applied to get the ensemble aggregated input ranking $\bar{r}_i$ for each *i*-th bootstrapped-dataset.

Borda and Schulze methods are, then, applied again to obtain the final (bottom-up) input ranking order $\bar{r}$.

3.2 The All-Out Strategy (ES2)

The idea is to merge a priori the (all-out) information coming from the bootstrapped-datasets (Figure 2).

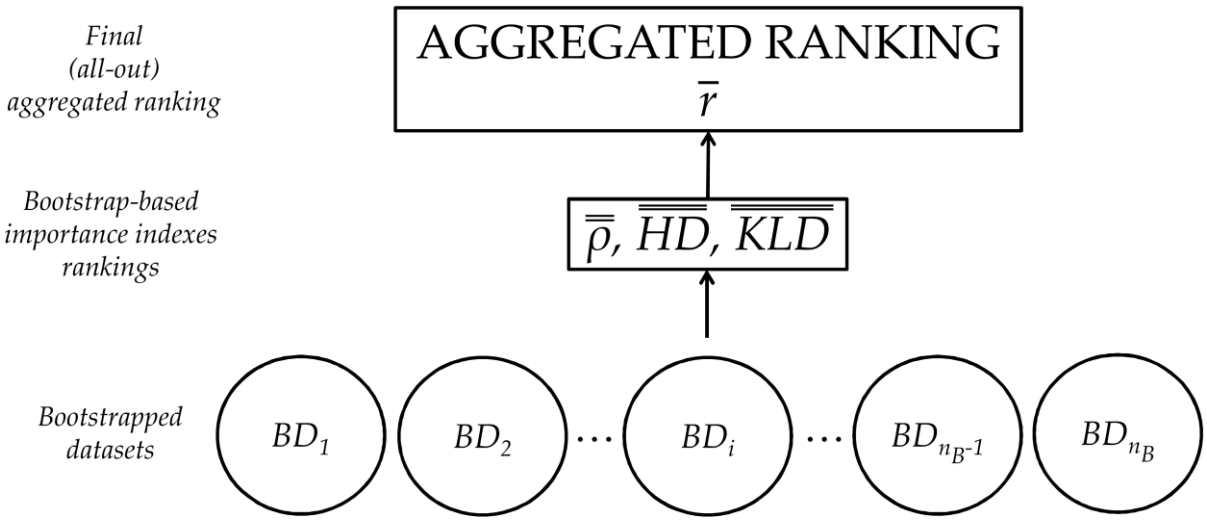

Figure 2*: scheme of the proposed all-out strategy (ES2).*

For each input variable x_j, the expected values of the three measures of importance $\rho_{i,j}$, $HD_{i,j}$, $KLD_{i,j}$ over the population of n_B bootstrapped-datasets are computed:

$$\overline{\overline{\rho}}_j = \mathrm{E}_i\left[\rho_{j,i}\right] = \frac{1}{n_B}\sum_{i=1}^{n_B}\rho_{i,j} \tag{11}$$

$$\overline{\overline{HD}}_j = \mathrm{E}_i\left[HD_{j,i}\right] = \frac{1}{n_B}\sum_{i=1}^{n_B}HD_{i,j} \tag{12}$$

$$\overline{\overline{KLD}}_j = \mathrm{E}_i\left[KLD_{j,i}\right] = \frac{1}{n_B}\sum_{i=1}^{n_B}KLD_{i,j} \tag{13}$$

These values, sorted in ascending order, provide three model input rankings, which are, then, aggregated by the Borda or Schulze methods (Section 2.3).

3.3 The Filter Strategy (ES3)

We compute the mean value μ_{T_j} and the variance $\sigma^2_{T_j}$ of the importance measures $\rho_{i,j}$, $HD_{i,j}$, $KLD_{i,j}$ obtained from the *i*-th bootstrapped dataset for each input variable x_j, viz:

$$\mu_{T_j} = \mathrm{E}\left[T_j\right] \cong \frac{1}{3n_B}\sum_{i=1}^{n_B}\sum_{l} t_{i,j}^{l} \tag{14}$$

$$\sigma_{T_j}^2 = Var\left[T_j\right] \cong \frac{1}{3n_B - 1}\sum_{i=1}^{n_B}\sum_{l} (t_{i,j}^{l} - \mu_{T_j})^2 \tag{15}$$

where $t_{i,j}^{l}$ ($i = 1,2,\ldots,n_B$, $j = 1,2,\ldots,n$, $l = IS, HD, KLD$) is a realization of a random variable T_j such that:

$$t_{i,j}^{IS} = \frac{\rho_{i,j}}{\sum_{j=1}^{n} \rho_{i,j}} \tag{16}$$

$$t_{i,j}^{HD} = \frac{HD_{i,j}}{\sum_{j=1}^{n} HD_{i,j}} \tag{17}$$

$$t_{i,j}^{KLD} = \frac{KLD_{i,j}}{\sum_{j=1}^{n} KLD_{i,j}} \tag{18}$$

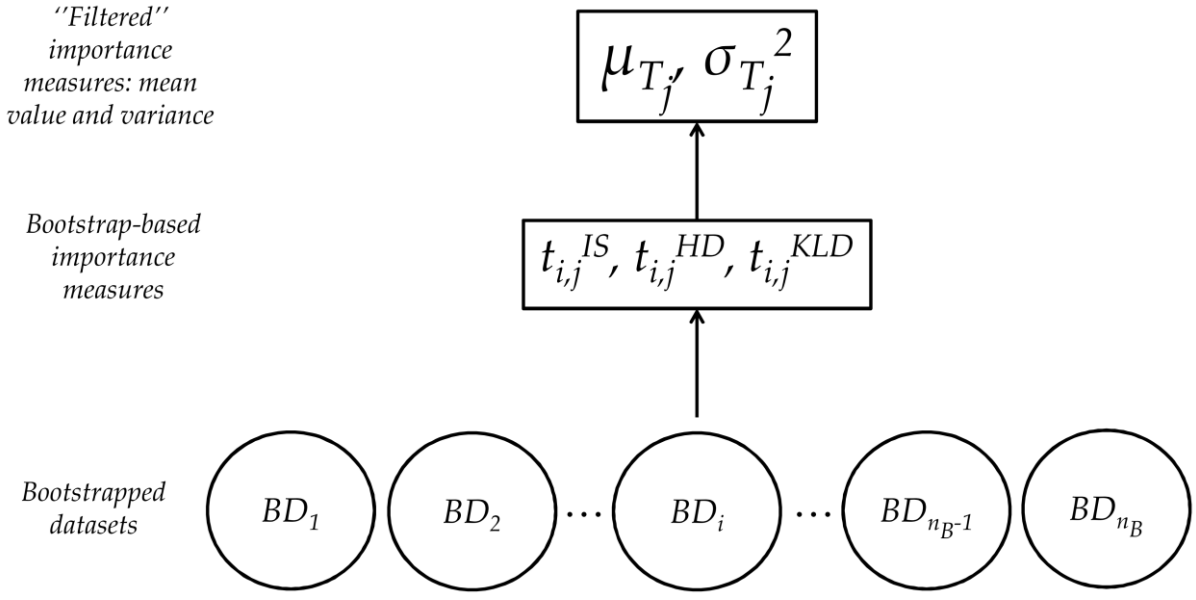

Figure 3: *scheme of the proposed filter strategy (ES3).*

4. CASE STUDY

Among all BE-TH codes employed in nuclear safety, RELAP-5 and TRAC have traditionally been used to reproduce transients of Pressurized Water Reactors (PWRs) and Boiling Water Reactors (BWRs), respectively. Nowadays, the TRACE code (TRAC/RELAP Advanced Computational Engine) is being developed to make use of the most favorable characteristics of the RELAP-5 and TRAC codes to simulate both, PWR and BWR, technologies. In this paper, the TRACE code has been used to simulate a LBLOCA occurring in one of the four cold legs of the Zion 1 NPP [43; 44].

The model of the plant for the TRACE code has been supplied with the TRACE V05 patch 3 version distribution, which has been adapted for the case study. The plant model consists of *97* hydraulic volumes, *36* heat structures and the necessary controls to perform the simulation of a cold leg LBLOCA as shown in Figure 4.

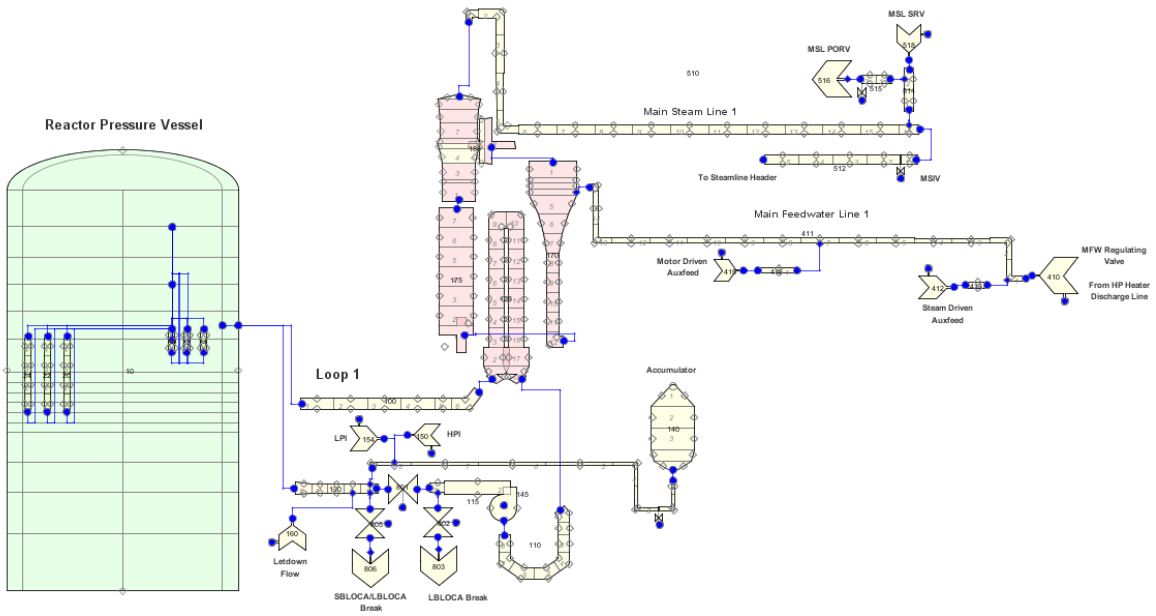

Figure 4*: view of the TRACE model for the Zion 1 NPP LBLOCA simulation.*

The Zion 1 NPP is a four-loop Westinghouse design PWR located in Zion (Illinois, USA). The system parameters values and distibutions are listed in Table 1, taken from [45].

In general, when a LBLOCA occurs the emergency shut-down of the reactor (SCRAM) is triggered and the scenario develops along three phases: i) blowdown, that begins with a cold leg break of size S_{23} (Table 1) and ends when the Emergency Core Cooling System (ECCS)

starts injecting water in the remaining intact loops; ii) refill, that begins when the water injection starts and ends when the mixture level in the lower plenum reaches the core inlet; iii) reflood, that begins when the liquid level in the core increases and ends when the core is totally quenched. In the PWR here considered, the actuation of two ECCS is taken into account: one is a passive safety system consisting of one large accumulator per loop, containing borated water under a blanket of nitrogen at mean pressure $S_{10} = 4.16$ [MPa] (see Table 1), which is poured into the vessel cold legs; the other system is an electrically driven Low-Pressure coolant Injection System (LPIS) that discharges water into the hot legs of the primary loops. In more detail, the accident scenario here analyzed consists in:

1. start of the transient with the break opening followed by a mean blowdown period of *13* [s];
2. injection from the accumulators when the pressure in the primary circuit falls below $S_{10} = 4.16$ [MPa]. In particular, the pressure set point considered in the simulation is *4.14* [MPa] [45];
3. start of LPIS injection when the primary pressure reaches *1.42* [MPa] with a mean delay of $S_{16} = 15$ [s] (Table 1) for the LPIS activation;
4. mean refill phase (started about *13* [s] after the break) period of *26* [s], when the core is filled again with liquid;
5. core quenching completed at about *460* [s] after the break.

Furthermore, as the pressure drop at the start is too sharp, the actuation of the High Pressure Injection System (HPIS) is not considered in the simulations.

The original available dataset consists of $N_R = 96$ runs of the tailored TH code: for each simulation, a batch of the *23* input variables listed in Table 1 is sampled and the maximum Peak Cladding Temperature (PCT) reached during the transient is collected as model output safety parameter.

The first PCT, reached during blowdown, has a value of about *1000* [K], whereas the second larger PCT is reached during the reflood phase, about *170* [s] after the initiation of the transient. The threshold of the PCT not to be reached during the LBLOCA is set equal to *1477* [K].

Input variables	*Description*	*Unit*	*Mean*	*Std. deviation*	*Dist. type*
S_1	Reactor core initial power	MW	3.250E+03	2.167E+01	*Normal*
S_2	UO_2 specific heat capacity	$\frac{J}{kgK}$	3.073E+02	2.049E+00	*Normal*
S_3	UO_2 thermal conductivity	$\frac{W}{mK}$	3.502E+00	1.167E-01	*Normal*
S_4	Reactor core power peaking factor	$-$	1.247E+00	2.078E-02	*Normal*
S_5	Hot gap size of the hot fuel rods	mm	5.400E-02	3.600E-03	*Normal*
S_6	Hot gap size of the reactor core	mm	5.400E-02	3.600E-03	*Normal*
S_7	Pressurizer water initial level	m	8.800E+00	3.333E-02	*Normal*
S_8	Surge line coefficient of friction	$-$	1.000E+00	1.000E+00	*Lognormal*
S_9	Pressurizer initial pressure	Pa	1.550E+07	3.333E+04	*Normal*
S_{10}	Accumulator initial pressure	Pa	4.160E+06	6.667E+04	*Normal*
S_{11}	Accumulator liquid initial temperature	K	3.251E+02	3.333E+00	*Normal*
S_{12}	Accumulator volume	m^3	3.890E+01	3.175E-01	*Uniform*
S_{13}	LPIS water mass flow rate	$\frac{kg}{s}$	8.800E+01	1.467E+00	*Normal*
S_{14}	LPIS water temperature	K	3.026E+02	4.792E+00	*Uniform*
S_{15}	Reactor vessel pressure	Pa	1.344E+07	2.327E+05	*Uniform*
S_{16}	LPIS delay	s	1.500E+01	8.660E+00	*Uniform*
S_{17}	Feed-water initial mass flow rate	$\frac{kg}{s}$	1.736E+04	2.314E+02	*Normal*
S_{18}	Cold leg initial temperature	K	5.650E+02	6.667E-01	*Normal*
S_{19}	Water initial mean temperature in the reactor upper header	K	5.680E+02	1.667E+00	*Normal*
S_{20}	Intact loops primary pumps rotational speed after LOCA	$\frac{rad}{s}$	7.684E+01	5.123E-01	*Normal*
S_{21}	Broken loop primary pump rotational speed after LOCA	$\frac{rad}{s}$	1.851E+02	6.171E+00	*Normal*
S_{22}	Reactor core power after SCRAM multiplier	$-$	1.000E+00	2.667E-02	*Normal*
S_{23}	Break section equivalent diameter	in	3.000E+01	5.774E+00	*Uniform*

Table 1*: model input variables list with their associated distributions [45].*

5. RESULTS

5.1 Application of the Bootstrapped-EBSA to the Zion 1 NPP Case Study

The original dataset of TH simulations has been replicated by Bootstrap to generate $n_B = 1000$ datasets BD_i, $i = 1,2,\ldots,n_B$ (Section 2.1), which ensures a sufficiently large total number of code runs as claimed in [35]. Each BD_i data matrix is comprised of $N_R = 96$ rows and $n + o = 24$ columns, each containing the values of one of the model variables (Table 1).

The analytical reconstruction of the model output PDF $f(y)$ has been performed by Finite Mixture Models (FMMs) for each *i*-th BD_i, where the number K of FMMs clusters is set equal to *2* (i.e., the PDFs $f_i(y)$, $i = 1,2,\ldots,n_B$, is a Gaussian bimodal distribution) as a result of a best-fitting procedure to the experimental model outputs, although an automatic optimization of the number K is also possible [51]. As an example, Figure 5 shows the bimodal output PDF (solid line) of a two-cluster FMM based on the experimental model output

distribution (histogram) of one generic *i*-th dataset.

The bimodal nature of the distribution is explained by the values taken by the ”Break section equivalent diameter” (S_{23}) input variable which turns out to be the input variable that affects the transient evolution the most, as shown quantitatively in what follows: code runs with large break sizes (i.e., break section equivalent diameter between 30 and 40 inches) generate transients with poor core refrigeration capability, which lead the PCT to reach high values, thus giving rise to the high mode of the FMM, whereas code runs fed with small break sizes (i.e., break section equivalent diameter between 20 and 30 inches) originate the low mode of the FMM.

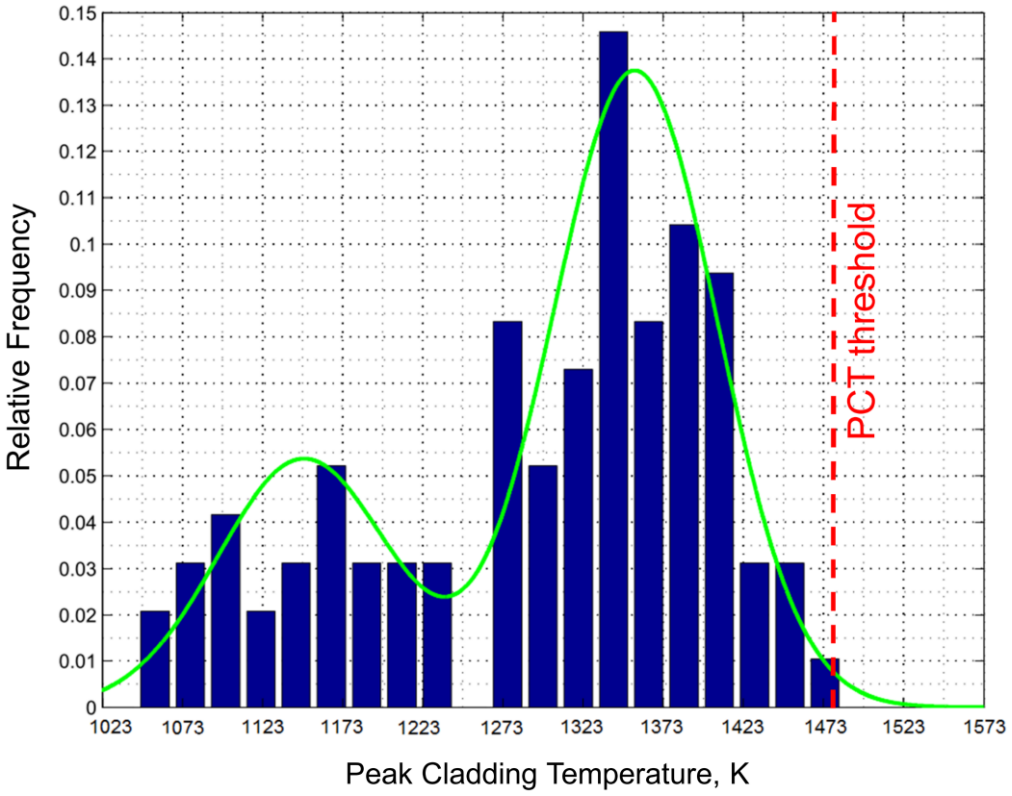

Figure 5*: reconstructed PDF of the model output based on a two-cluster FMM (solid line) for the generic i-th bootstrapped dataset.*

As already explained in Section 2.2, the greater the contribution of an input variable in shaping the K clusters of the reconstructed PDF is, the larger the importance of this input is in affecting the model output variability.

The sensitivity of the PCT output to the $n = 23$ input variables considered is quantified by calculating $\rho_{i,j}$, $HD_{i,j}$ and $KLD_{i,j}$ for each *j*-th input, $j = 1,2,\ldots,n$, on each *i*-th bootstrapped dataset, $i = 1,2,\ldots,n_B$. As an example, Figures 6, 7 and 8 show $\rho_{i,j}$, $HD_{i,j}$ and $KLD_{i,j}$ respectively, for the generic *i*-th bootstrapped dataset of Figure 5.

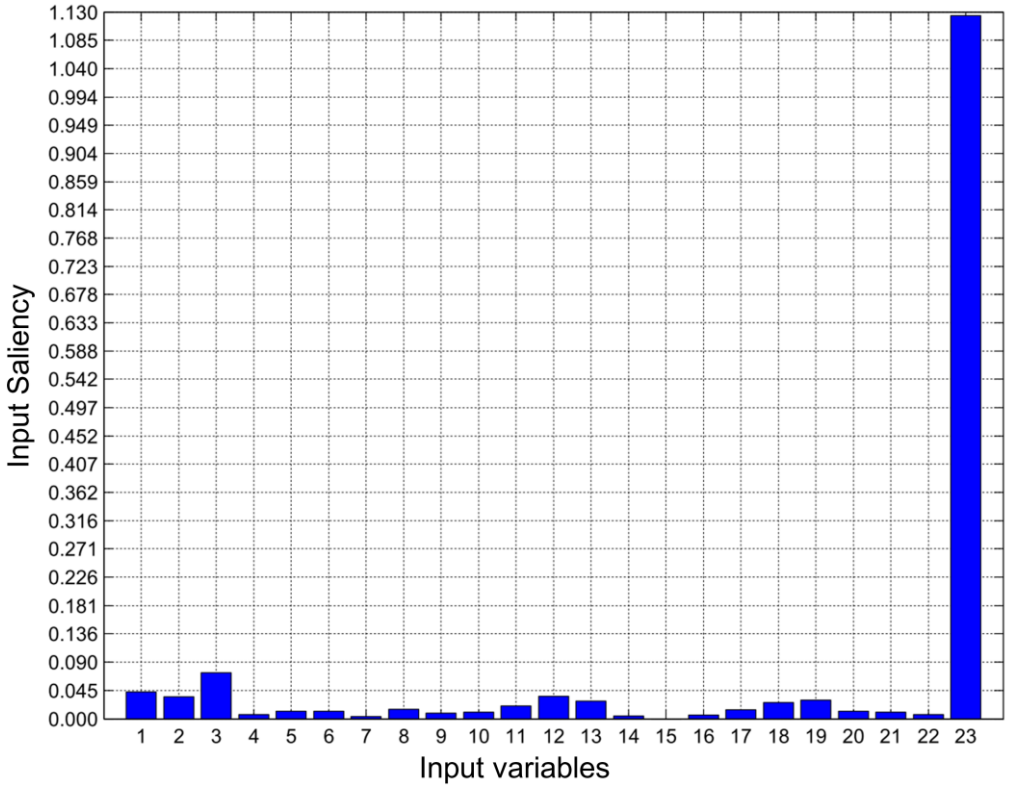

Figure 6*: saliencies (Eq. 4) of all the 23 model input variables for the generic i-th bootstrapped dataset.*

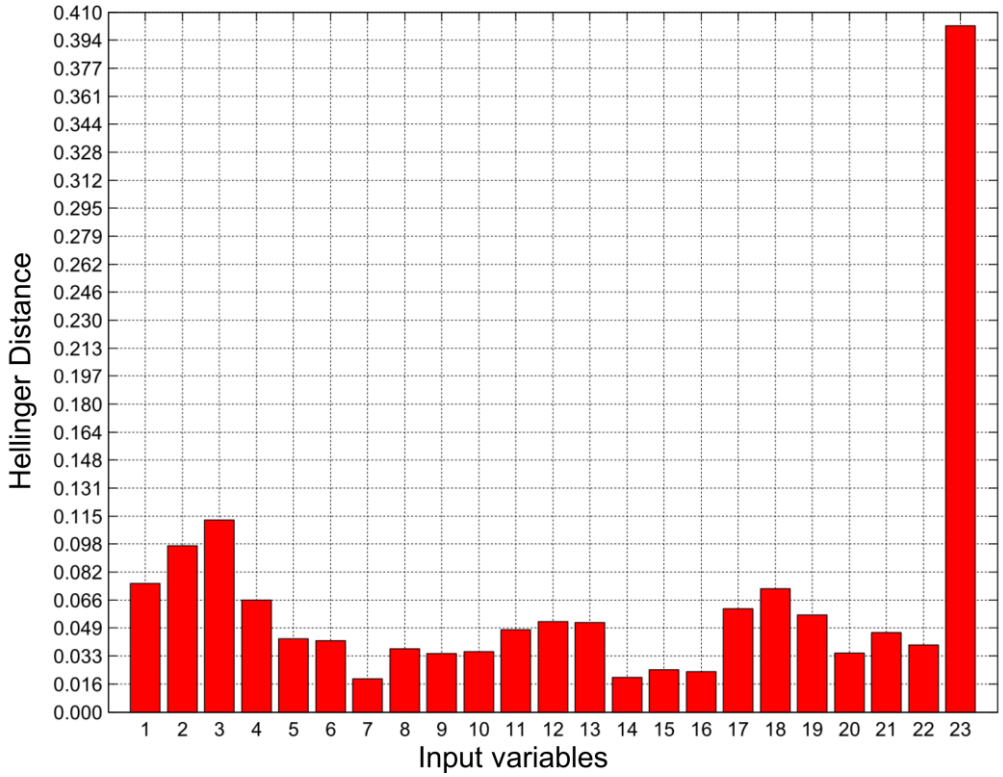

Figure 7*: Hellinger Distances (Eq. 6) of all the 23 model input variables for the generic i-th bootstrapped dataset.*

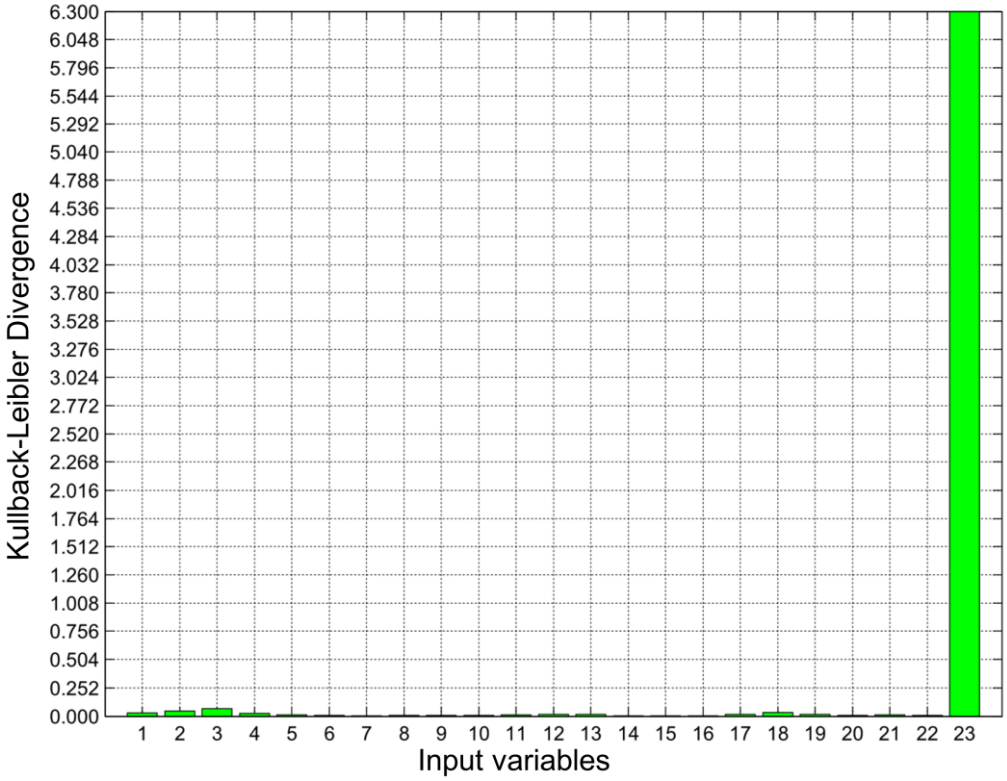

Figure 8: *Kullback-Leibler Divergences (Eq. 8) of all the 23 model input variables for the generic i-th bootstrapped dataset.*

In what follows, the results of the bottom-up (Section 3.1), all-out (Section 3.2) and filter (Section 3.3) strategies for bootstrapped-EBSA are presented and compared in Sections 5.2, 5.3, 5.4 and 5.5, respectively.

For the sake of completeness, in Section 6 the obtained results are compared with those achieved by the EBSA of [39] and the BEMUSE project [43]; benefits and limitations of the proposed bootstrapped-EBSA are highlighted with respect to these other methods.

5.2 Results of the Bottom-Up Strategy (ES1)

The application of the ensemble strategy ES1 (Section 3.1) results in the two final input rankings that are shown in Table 2, obtained by the Borda and Schulze aggregation methods. It can be observed that:

1. *5* input variables (S_{23}, S_{13}, S_2, S_{12}, S_{21}) occupy the same position for both aggregation methods;
2. the first *9* positions are held by the same input variables (S_{23}, S_8, S_4, S_{13}, S_{10}, S_1, S_2, S_{18}, S_{22});

3. input variables S_{23}, S_{13} and S_2 are ranked in the same order by both methods;
4. S_8 and S_4, S_{10} and S_1, S_{18} and S_{22} are switched in position in the two methods;
5. below position *10* the two final rankings disagree: for instance, input variable S_{19} is ranked 10th by the Schulze method and 21st by the Borda method.

ES1 rankings	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°	13°	14°	15°	16°	17°	18°	19°	20°	21°	22°	23°
Borda ranking	S_{23}	S_8	S_4	S_{13}	S_{10}	S_1	S_2	S_{18}	S_{22}	S_{11}	S_9	S_6	S_{12}	S_{15}	S_{14}	S_5	S_{20}	S_{17}	S_7	S_3	S_{19}	S_{16}	S_{21}
Schulze ranking	S_{23}	S_4	S_8	S_{13}	S_1	S_{10}	S_2	S_{22}	S_{18}	S_{19}	S_3/S_{11}	–	S_{12}	S_9	S_{20}	$S_6/S_{15}/S_{17}$	–	–	S_{14}	S_5/S_{16}	–	S_7	S_{21}

Table 2*: Borda and Schulze model input final rankings obtained through ensemble strategy ES1.*

Judging from the Borda and Schuze aggregation results, we can conclude that the bottom-up strategy ES1 identifies the following *9* input variables as the most important in affecting the model output:

1. S_{23}: “Break section equivalent diameter”;
2. S_8: “Surge line coefficient of friction”;
3. S_4: “Rector core power peaking factor”;
4. S_{13}: “LPIS water mass flow rate”;
5. S_{10}: “Accumulator initial pressure”;
6. S_1: “Reactor core initial power”;
7. S_2: “UO_2 specific heat capacity”;
8. S_{18}: “Cold leg initial temperature”;
9. S_{22}: “Reactor core power after SCRAM multiplier”.

5.3 Results of the All-Out Strategy (ES2)

The ensemble strategy ES2 (Section 3.2) yields the results reported in Table 3, with final rankings obtained by the Borda and Schulze methods. As Table 3 shows:

1. *9* input variables (S_{23}, S_8, S_{18}, S_{22}, S_{11}, S_3, S_9, S_{17}, S_{21}) occupy the same position for both the aggregation methods;
2. the first *2* positions are held by the same input variables (S_{23}, S_8);
3. changes in rank positions can be spotted (i.e., position *3* of the Schulze method is occupied by input variable S_4 that is placed in position *7* by the Borda ranking);
4. beyond position *10* the ranking orders disagree.

ES2 rankings	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°	13°	14°	15°	16°	17°	18°	19°	20°	21°	22°	23°
Borda ranking	S_{23}	S_8	S_{13}	S_1	S_2/S_{10}	–	S_4	S_{18}	S_{22}	S_{11}	S_{15}	S_{14}	S_{12}	S_{19}	S_3/S_7	–	S_5/S_{20}	–	S_6/S_9	–	S_{17}	S_{16}	S_{21}
Schulze ranking	S_{23}	S_8	S_4	S_{13}	S_1	S_{10}	S_2	S_{18}	S_{22}	S_{11}	S_{19}	S_{15}	S_{14}	S_{12}	S_3	$S_6/S_7/S_{20}$	–	–	S_9	S_5	S_{17}	S_{16}	S_{21}

Table 3*: Borda and Schulze model input final rankings obtained through ensemble strategy ES2.*

We can conclude that ES2 suggests *10* input variables as relevant for the model output variability:

1. S_{23}: “Break section equivalent diameter”;
2. S_8: “Surge line coefficient of friction”;
3. S_{13}: “LPIS water mass flow rate”;
4. S_1: “Reactor core initial power”;
5. S_2: “UO_2 specific heat capacity”;
6. S_{10}: “Accumulator initial pressure”;
7. S_4: “Rector core power peaking factor”;
8. S_{18}: “Cold leg initial temperature”;
9. S_{22}: “Reactor core power after SCRAM multiplier”;
10. S_{11}: “Accumulator liquid initial temperature”.

5.4 Results of the Filter Strategy (ES3)

The major limitation of ES1 and ES2 is that the model input rankings provided do not tell neither how large the difference in importance is between two consecutive positions nor they allow identifying the “stopping criterion” for deciding the last important input variable to be considered in the ranking.

Indeed, as can be seen in Tables 2 and 3, the ES1 and ES2 rankings are very similar to each other, with negligible disagreement until position *10*. We could be driven by such results to consider all of the first *10* input parameters to be relevant.

The filter strategy (ES3) of Section 3.3 allows computing the mean value μ_{T_j} and variance $\sigma^2_{T_j}$ of the importance metrics (Eq. 14, 15, 16) for each *j*-th input. In Table 4, the values of μ_{T_j}, σ_{T_j} and $\frac{\sigma_{T_j}}{\mu_{T_j}}$ are reported for all the *23* model input variables.

T_j **distribution features**	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}	T_{19}	T_{20}	T_{21}	T_{22}	T_{23}
Mean, μ_{T_j} $(\times 10^2)$	2.01	1.95	1.44	2.15	1.50	1.47	1.49	2.97	1.48	1.99	1.61	1.52	2.27	1.50	1.57	1.40	1.44	1.71	1.48	1.46	1.34	1.73	62.52
Standard deviation, σ_{T_j} $(\times 10^2)$	2.05	1.99	1.63	2.23	1.62	1.58	1.63	2.88	1.59	1.96	1.65	1.66	2.23	1.61	1.67	1.56	1.52	1.77	1.73	1.56	1.52	1.86	27.88
Std.-dev.-to-mean ratio, $\frac{\sigma_{T_j}}{\mu_{T_j}}$	1.02	1.02	1.13	1.04	1.08	1.07	1.09	0.97	1.07	0.98	1.03	1.10	0.98	1.07	1.06	1.11	1.06	1.03	1.17	1.06	1.14	1.08	0.45

Table 4*: values of* μ_{T_j} , σ_{T_j} *and* $\frac{\sigma_{T_j}}{\mu_{T_j}}$ *for* $j = 1,2,\ldots,23$.

In Figure 9 μ_{T_j} and $\frac{\sigma_{T_j}}{\mu_{T_j}}$ are plotted against one another for each *j*-th input variable. The parameter $\frac{\sigma_{T_j}}{\mu_{T_j}}$ is used as a discriminating feature to explain the variability of the input variables importance: a large $\frac{\sigma_{T_j}}{\mu_{T_j}}$ value means that the *j*-th input importance varies a lot in the bootstrapped datasets, whereas a small $\frac{\sigma_{T_j}}{\mu_{T_j}}$ value means that the *j*-th input variable varies little. Judging from the results obtained, variable S_{23} stands out as the most important input since $\mu_{T_{23}} = 0.625$ and $\mu_{T_j} < 0.030$ for $j = 1,2,\ldots,22$. The zoom-in on the bottom of Figure 9 reveals two clusters of model inputs: variables S_8, S_{13}, S_4, S_1, S_{10}, S_2, S_{18} and S_{11} occupy the region where μ_{T_j} and $\frac{\sigma_{T_j}}{\mu_{T_j}}$ are large and small, respectively, whereas variables S_{22}, S_{15}, S_{17}, S_{20}, S_6, S_{14}, S_9, S_5, S_7, S_{12}, S_{16}, S_3, S_{21} and S_{19} exhibit small μ_{T_j} and large $\frac{\sigma_{T_j}}{\mu_{T_j}}$ values.

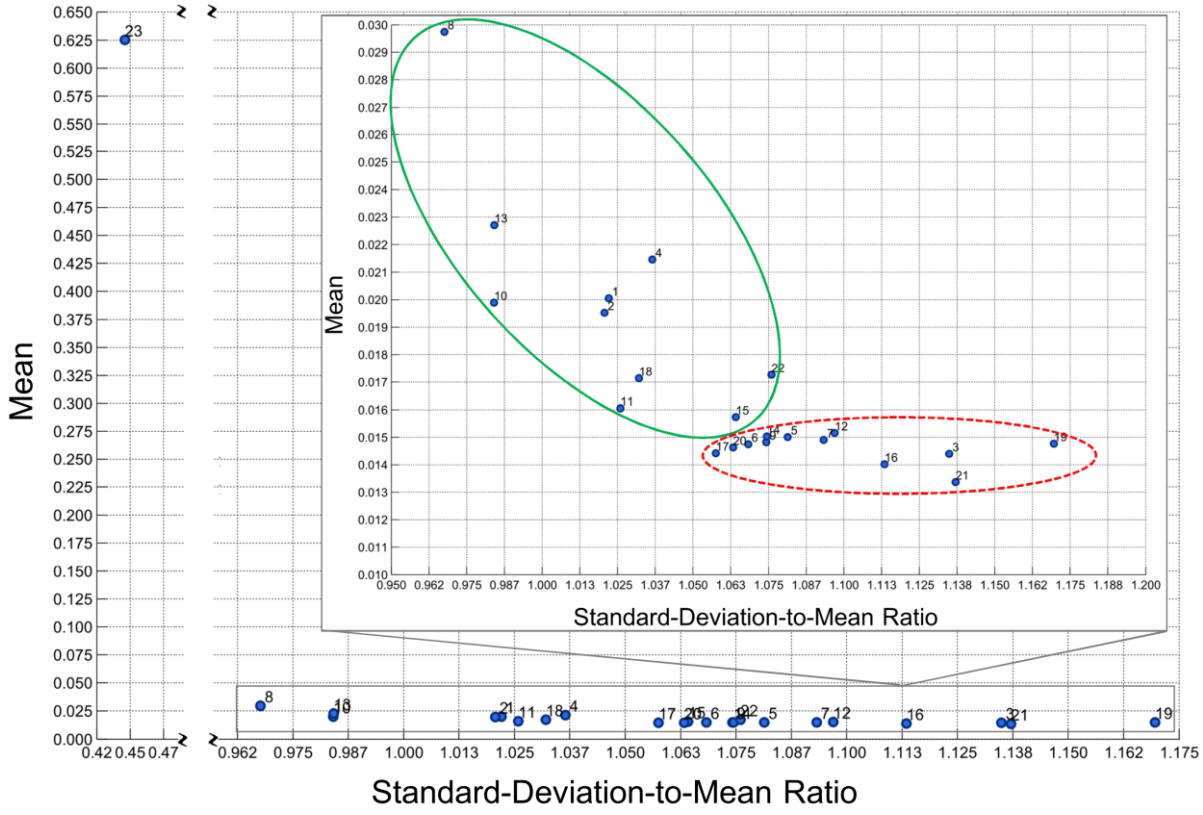

Figure 9: μ_{T_j} *against* $\frac{\sigma_{T_j}}{\mu_{T_j}}$ *plot for the 23 model input variables.*

We can conclude that all input variables contained in the first cluster (solid line) can be judged as relevant, while those belonging to the second cluster (dotted line) shall be neglected. Note that in the case at hand, the identification of the clusters turns out to be quite straightforward, thus allowing a qualitative and visual grouping of the 23 model input variables; in less straighforward cases, one may resort to unsupervised clustering techniques (e.g., Spectral Clustering [52]) for identifying clusters in the structure of the input variables data.

In summary, ensemble strategy ES3 identifies the following variables as important in determining the model output:

1. S_{23}: "Break section equivalent diameter";
2. S_8: "Surge line coefficient of friction";
3. S_{13}: "LPIS water mass flow rate";
4. S_4: "Rector core power peaking factor";
5. S_1: "Reactor core initial power";
6. S_{10}: "Accumulator initial pressure";

7. S_2: "UO_2 specific heat capacity";
8. S_{18}: "Cold leg initial temperature";
9. S_{11}: "Accumulator liquid initial temperature";
10. S_{15}: "Reactor vessel pressure";
11. S_{22}: "Reactor core power after SCRAM multiplier".

Considering all the final aggregated ranking results, one can observe that all the ensemble strategies identify variable S_{23} as the most important among all the model inputs, i.e., the "Break section equivalent diameter". This is physically reasonable for the accidental scenario under analysis: the rupture in question occurs in one of the primary coolant loops and the break cross-sectional area in the cold leg determines the decrease in coolant mass flow through the reactor core that, in turn, greatly affects the increase in the PCT.

Moreover, all the ensemble strategies agree on variable S_8 as the second most relevant input of the model, i.e., the coefficient of friction on the inside wall of the surge line: this affects the pressure losses that are exerted on the cold mass flow on its way from the pressurizer to the reactor. In the event of a LOCA, the unexpected drop in coolant flow results in a quick depressurization of the reactor vessel that is, in part, hindered through the discharge of water from the pressurizer to the reactor hot leg and, eventually, through the fuel rods in the attempt to reduce uncovering and overheating: a larger coefficient of friction may endanger the cooling capability of the system.

Among the other input variables ranked as important by ES1 and ES2 it is worth mentioning S_{22}, i.e., the reactor core power multiplier after the emergency shut-down: this influences the transient evolution as it defines the initial power that the ECCS has to remove from the core to prevent the fuel rods from melting. Finally, ES3 (but not ES1 and ES2) identifies as important S_{11}, i.e., the temperature of the liquid stored in the accumulator tanks that are used as sinks of borated water to be injected in the reactor after the blowdown phase of the LOCA.

6. COMPARISON WITH OTHER METHODS

6.1 Comparison with Input Saliency, Hellinger Distance and Kullback-Leibler Divergence Methods

Figures 10, 11 and 12 show the computed values of ρ_j, HD_j and KLD_j based on

the original dataset for each *j*-th model input. All three of the SA methods assign to variable S_{23} the largest value of importance, but they show lack of agreement on the other inputs ranking: i) the Input Saliency method identifies S_8 and S_1 as the second and third most important variables, respectively, ii) the Hellinger Distance and Kullback-Leibler methods suggest S_4 as the second most important variable and rank inputs S_8 and S_1 both in third position of the ranking. As ρ_j, HD_j and KLD_j are extremely low for the remaining inputs, any other variable could be judged as non-important.

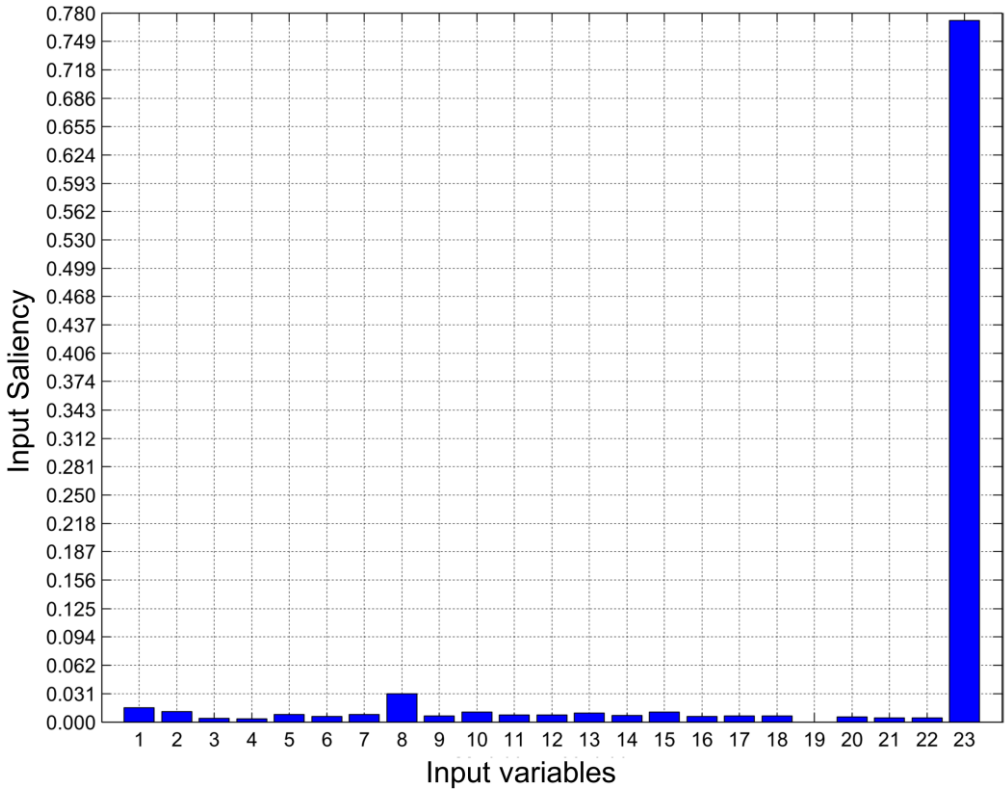

Figure 10: *saliencies (Eq. 4) of all the 23 model input variables for the original set of data.*

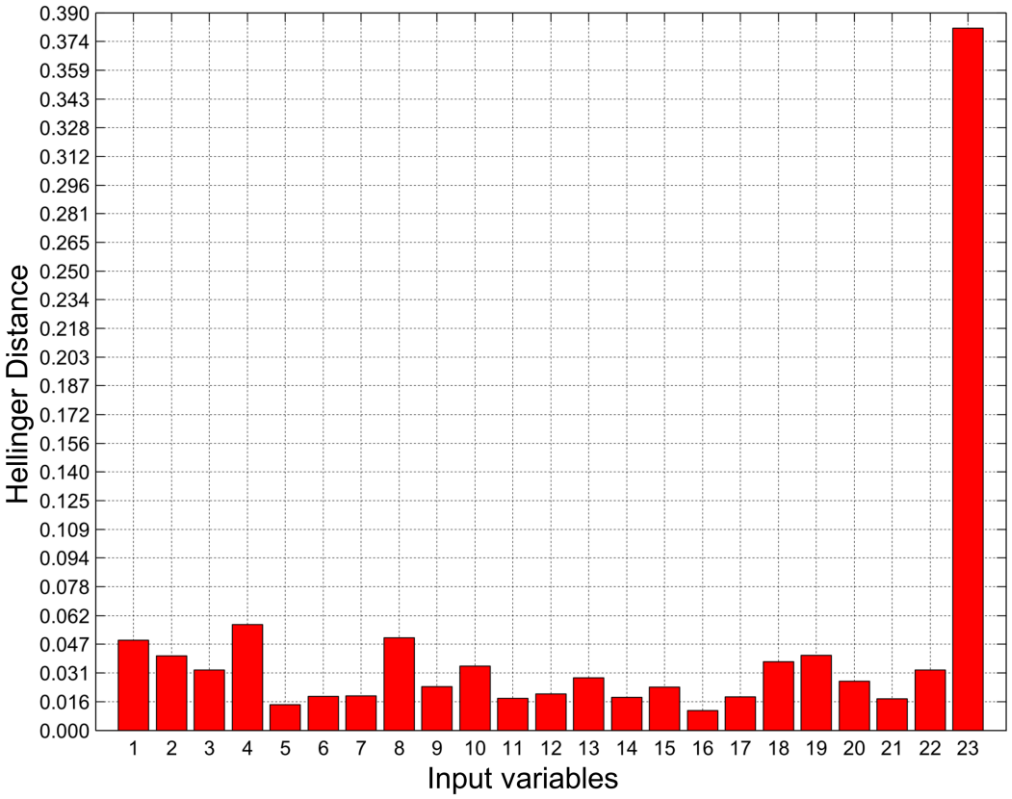

Figure 11*: Hellinger Distances (Eq. 6) of all the 23 model input variables for the original set of data.*

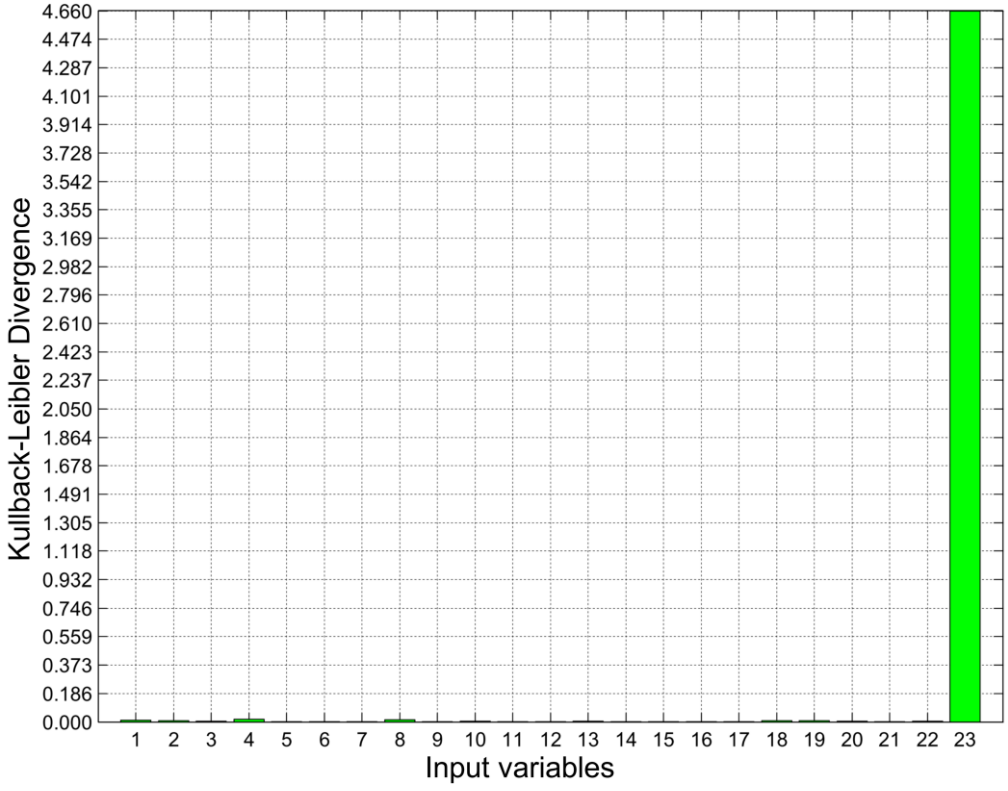

Figure 12*: Kullback-Leibler Divergences (Eq. 8) of all the 23 model input variables for the original set of data.*

Summarizing, the results of the three metrics applied separately to the original set of data agree on identifying only three input variables as relevant (S_{23}, S_8, S_1) and rank them in a different order.

6.2 Comparison with EBSA

If the Input Saliency, Hellinger Distance, Kullback-Leibler Divergence are directly computed on the original set of data as in Section 6.1, without bootstrapping, and the ensemble aggregated rankings by Borda and Schulze methods are applied, the results shown in Table 5 are obtained. It can be seen that these rankings somewhat differ one another and very little can be inferred from them as overall conclusion. One certainty is that input variable S_{23} holds position *1* in both rankings and, thus, is the most relevant. But, after this the two rankings provide inconsistent information, due to the limited data they are built on, only $N_R = 96$ LBLOCA trasients.

Rankings w/o bootstrap	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°	13°	14°	15°	16°	17°	18°	19°	20°	21°	22°	23°
Borda ranking	S_{23}	S_8	S_1	S_2	S_{10}	S_4	S_{18}	S_{13}	S_{15}/S_{19}	–	S_3/S_{22}	–	S_9	S_{12}	S_7/S_{20}	–	S_{17}	$S_6/S_{11}/S_{14}$	–	–	S_5	S_{21}	S_{16}
Schulze ranking	S_{23}	S_4	S_8	S_1	S_2	S_{19}	S_{18}	S_{10}	S_3	S_{22}	S_{13}	S_{20}	S_9	S_{15}	S_{12}	S_7	S_{17}	S_6	S_{14}	S_{11}	S_{21}	S_5	S_{16}

Table 5: *Borda and Schulze model input rankings derived from the original dataset, without bootstrapping.*

6.3 Comparison with Other Methods of Literature

The findings reported in Section 5 have been further compared with SA results of the BEMUSE programme - phase V (Best Estimate Methods - Uncertainty and Sensitivity Evaluation) promoted by the Working Group on Accident Management and Analysis (WGAMA) of OECD [43]. This reference work is considered as a benchmark for validation of the proposed bootstrapped-ensemble-based methodology.

The BEMUSE programme consisted in a comparative exercise of different BEPU (Best Estimate Plus Uncertainty) methods by fourteen different groups of participants. Everyone of them was requested to perform individual GSA and to assign a score equal to *3* to the most important input parameters and *0* to the negligible ones [43]. Table 6 shows the aggregated ranking scores associated to each non-negligible input.

Input variables	*Description*	*Aggregated ranking score*	*ES*1	*ES*2	*ES*3
S_3	UO_2 thermal conductivity	27			
S_{22}	Reactor core power after SCRAM multiplier	22	•	•	•
S_4	Reactor core power peaking factor	14	•	•	•
S_1	Reactor core initial power	10	•	•	•
S_5	Hot gap size of the hot fuel rods	9			
S_6	Hot gap size of the reactor core	5			
S_{11}	Accumulator liquid initial temperature	3		•	•
S_{13}	LPIS water mass flow rate	2	•	•	•
S_9	Pressurizer initial pressure	2			
S_{10}	Accumulator initial pressure	1	•	•	•
S_2	UO_2 specific heat capacity	1	•	•	•
S_{20}	Intact loops primary pump rotational speed after LOCA	1			
S_{21}	Broken loop primary pumps rotational speed after LOCA	1			

Table 6: *aggregated ranking score of the BEMUSE programme [39].*

It can be seen that the set of input variables considered as important (aggregated ranking score > *0*) is somewhat larger than that obtained in Section 5 by the three ensemble strategies: the BEMUSE programme results identify *13* relevant inputs, whereas ES1 recognizes *9* inputs as important, ES2 *10* and ES3 *11*, respectively. Furthermore, the BEMUSE programme agrees with:

1. ES1 on *6* input variables: S_{22}, S_4, S_1, S_{13}, S_{10}, S_2;
2. ES2 on *7* input variables: S_{22}, S_4, S_1, S_{11}, S_{13}, S_{10}, S_2;
3. ES3 on 7 input variables: S_{22}, S_4, S_1, S_{11}, S_{13}, S_{10}, S_2.

The bottom-up strategy (ES1) misses to identify variable S_{11}, i.e., the "Accumulator liquid initial temperature". The results of ES3 enable to also quantify the difference in importance between the inputs and the stability of each input ranking position (Section 3.3), thus, allowing a "stopping criterion" in identifying the relevant parameters of the model. The three major differences between the findings of the filter strategy (ES3) and those of BEMUSE are:

1. "Break section equivalent diameter" (S_{23}) is not included in Table 6, probably due to the straigthforward judgement of S_{23} as an important parameter in BEMUSE, thus deemed as not necessary to be mentioned among the results of data manipulation;
2. "Surge line coefficient of friction" (S_8) is the second most important input variable according to ES1, ES2 and ES3, but it plays a minor role in BEMUSE. It is likely that the LPIS mass flow rate and delay are respectively larger and smaller in the two cases.

This would explain why the "Surge line coefficient of friction" (S_8) is considered to be more important. Indeed, the resulting larger difference in pressure between the pressurizer and the reactor would amplify the influence of the coefficient of friction on the mass flow rate of water that is discharged in the hot legs of the vessel and, thus, on the temperature of the fuel cladding, which is refrigerated by that flow of water;

3. "UO_2 thermal conductivity" (S_3) is neglected in the rankings of ES1, ES2 and ES3, while it is the most relevant for BEMUSE in Table 6. Nevertheless, the fuel thermal characteristics are represented and taken into account by S_2, i.e., "UO_2 specific heat capacity", which is identified as a relevant input by all the ensemble strategies ES1, ES2 and ES3.

Finally, it is worth mentioning that BEMUSE programme made use of a large amount of code runs (*3000*), for the identification of the important model variables, against the $N_R = 96$ code runs of the three ensemble strategies: the proposed bootstrapped-EBSA strategies are, thus, capable to achieve effective results in much less computational time (Table 7).

Steps	*Elapsed time*, $[s]$
Bootstrapped-datasets generation	0.296
Bottom-up strategy (ES1)	2.466E+03
All-out strategy (ES2)	0.031
Filter strategy (ES3)	0.140

Table 7*: computational time employed by the three proposed ensemble strategies (i.e., bottom-up, all-out and filter).*

7. CONCLUSIONS

The identification of the variables that most affect the response of a BE-TH code is an important task in the performance of safety analyses of nuclear systems.

In this paper we have presented an original framework that exploits a combination of EBSA and the Bootstrap method to address this issue when a very limited amount of code runs are available.

This framework consists in bootstrapping the initial available dataset and building an

ensemble of sensitivity analysis metrics to arrive at an aggregated ranking. In this work, three strategies, bottom-up, all-out and filter, have been proposed, developed and compared for the ensemble aggregation of three SA metrics (Input Saliency, Hellinger Distance, and Kullback-Leibler Divergence). The framework has been applied to $N_R = 96$ TRACE TH code simulations reproducing the response of the Zion 1 NPP undergoing a LBLOCA accident.

The results obtained have been compared with other SA methods of literature and the filter strategy (ES3) emerges as most effective. With respect to all the others, ES3 indeed: i) allows defining a "stopping criterion" to distinguish the important variables and those that can be neglected, ii) reproduces similarly satisfactory results as other much more computationally burdensome methods.

ACKNOWLEDGEMENT

The authors would like to thank all the reviewers for their valuable comments to improve the quality of this paper.

REFERENCES

[1] S. Tarantola, V. Kopustinskas, R. Bolado-Lavin, A. Kaliatka, E. Uspuras and M. Vaisnoras. Sensitivity Analysis Using Contribution to Sample Variance Plot: Application to a Water Hammer Model. In Reliability Engineering & System Safety, Vol. 99, pp. 62-73. March 2012.

[2] S. Martorell, Y. Nebot, J. F. Villanueva, S. Carlos, V. Serradell and F. Pelayo. Safety Margins Estimation Method Considering Uncertainties within Risk-Informed Decision Making Framework. In Proceedings of the PHYSOR 2006 Conference. September 1999.

[3] Pilch M, Trucano TG, Helton JC. Ideas Underlying the Quantification of Margins and Uncertainties. Reliability Engineering and System Safety 2011; 96(9): 965-975.

[4] Helton JC. Quantification of Margins and Uncertainties: Conceptual and Computational Basis. Reliability Engineering and System Safety 2011; 96(9): 976-1013.

[5] Helton JC, Johnson JD, Sallaberry CJ. Quantification of Margins and Uncertainties: Example Analyses from Reactor Safety and Radioactive Waste Disposal Involving the Separation of Aleatory and Epistemic Uncertainty. Reliability Engineering and System Safety 2011; 96(9): 1014-1033.

[6] Helton JC, Johnson JD. Quantification of Margins and Uncertainties: Alternative Representations of Epistemic Uncertainty. Reliability Engineering and System Safety 2011; 96(9): 1034-1052.

[7] M. Gavrilas, J. Mayer, B. Youngblood, D. Prelewicz and R. Beaton. A Generalized

Framework for Assessment of Safety Margins in Nuclear Power Plants. In Proceedings to BE 2004: International Meeting on Updates in Best Estimate Methods in Nuclear Installations Safety Analysis, Vol. 36, pp. 28-36. American Nuclear Society, November 2004.

[8] B. E. Boyack, I. Catton, R. B. Duffey, P. Griffith, K. R. Kastma, G. S. Lellouche, S. Levy, U. S. Rohatgi, G. E. Wilson, W. Wulf and N. Zuber. Quantifying Reactor Safety Margins: An Overview of the Code Scaling, Applicability and Uncertainty Evaluation Methodology. In Nuclear Engineering and Design, Vol. 119, pp. 1-15. May 1990.

[9] G. E. Wilson, Boyack, I. Catton, R. B. Duffey, P. Griffith, K. R. Kastma, G. S. Lellouche, S. Levy, U. S. Rohatgi, W. Wulf and N. Zuber. Quantifying Reactor Safety Margins. Part 2: Characterization of Important Contributors to Uncertainty. In Nuclear Engineering and Design, Vol. 119, pp. 17-31. May 1990.

[10] W. Wulf, G. E. Wilson, B. E. Boyack, I. Catton, R. B. Duffey, P. Griffith, K. R. Kastma, G. S. Lellouche, S. Levy, U. S. Rohatgi, U. S. Wilson, and N. Zuber. Quantifying Reactor Safety Margins. Part 3: Assessment and Ranging of Parameters. In Nuclear Engineering and Design, Vol. 119, pp. 33-65. May 1990.

[11] U.S. Nuclear Regulatory Commission (NRC), (2012) State-of-the-Art Reactor Consequence Analyses (SOARCA) Report.NUREG-1935 NRC Washington, DC

[12] Ghosh, S.T., Mattie, P.D., Gauntt, R.O., Bixler, N.E., Ross, K.W., Sallaberry, C.J., Osborn, D.M., SOARCA peach bottom atomic power station long-term station blackout uncertainty analysis: Overviewm (2014) PSAM 2014 - Probabilistic Safety Assessment and Management,

[13] Gauntt, R.O., Ross, K.W., Osborn, D.M., Sallaberry, C.J., Goldmann, A.S., Cardoni, J.N., Mattie, P.D., Ghosh, S.T., Fuller, E., Soarca Peach Bottom Atomic Power Station long-term station blackout uncertainty analysis: MELCOR parameters and probabilistic results, (2013) International Topical Meeting on Probabilistic Safety Assessment and Analysis 2013, PSA 2013, 2, pp. 1310-1326.

[14] Sallaberry, C.J., Mattie, P.D., Kalinich, D.A., Osborn, D.M., Ghosh, S.T., Soarca Peach Bottom Atomic Power Station long-term station blackout uncertainty analysis: Probabilistic methodology and regression technique, (2013) International Topical Meeting on Probabilistic Safety Assessment and Analysis 2013, PSA 2013, 2, pp. 1294-1309.

[15] H. Glaeser, E. Hofer, M. Kloos and T. Skorek. Uncertainty and Sensitivity Analysis of a Post-Experiment Calculation in Thermal Hydraulics. In Reliability Engineering & System Safety, Vol. 45, pp. 19-33. 1994.

[16] Helton JC, Johnson JD. An Uncertainty/Sensitivity Study for the Station Blackout Sequence at a Mark I Boiling Water Reactor. Reliability Engineering and System Safety 1989; 26: 293-328.

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

[18] A. Guba, M. Makai and L. Pal. Statistical Aspects of Best Estimate Method-I. In Reliability Engineering & System Safety, Vol. 80, pp. 217-232. June 2003.

[19] E. Zio and F. Di Maio. Bootstrap and Order Statistics for Quantifying Thermal-Hydraulic Code Uncertainties in the Estimation of Safety Margins. In Science and Technology of Nuclear Installations, Vol. 2008, Article ID 340164. 2008.

[20 D. R. Langewisch and R. Dustin. Uncertainty and Sensitivity Analysis for Long-Running Computer Codes: A Critical Review. Massachusetts Institute of Technology, 2010.

[21] I. S. Honga, D. Y. Oh and I. G. Kim. Generic Application of Wilks Tolerance Limit Evaluation Approach to Nuclear Safety. In Proceedings of the OCDE/CSNI Workshop on Best Estimate Methods and Uncertainty Evaluations, Vol. 45, pp. 48-57. Nuclear Energy Agency, Committee on the Safety of Nuclear Installations, November 2013.

[22] D. M. Hamby. A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models. In Environmental Monitoring and Assessment, Vol. 32, pp. 135-154. September 1994.

[23] M. Ionescu-Bujor and D. G. Cacuci. A Comparative Review of Sensitivity and Uncertainty Analysis of Large-Scale Systems. I: Deterministic methods. In Nuclear Science and Engineering, Vol. 147, pp. 189-203. American Nuclear Society, July 2004.

[24] R. Bolado-Lavin, W. Casaings and S. Tarantola. Contribution to the Sample Mean Plot for Graphical and Numerical Sensitivity Analysis. In Reliability Engineering & System Safety, Vol. 94, pp. 1041-1049. June 2009.

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

[26] J. C. Helton. Uncertainty and Sensitivty Analysis Techniques for Use in Performance Assessment for Radioactive Waste Disposal. In Reliability Engineering & System Safety, Vol. 42, pp. 327-367. 1993.

[27] A. Saltelli, K. Chan and E. Scott. Sensitivity Analysis, 504 pages. John Wiley & Sons Inc., August 2000.

[28] .Helton JC, Johnson JD, Sallaberry CJ, Storlie CB. Survey of Sampling-Based Methods for Uncertainty and Sensitivity Analysis. Reliability Engineering and System Safety 2006; 91(10-11): 1175-1209.

[29] F. Cadini, E. Zio, F. Di Maio, V. Kopustinkas and R. Urbonas. A Neural-Network-Based Variance Decomposition Sensitivity Analysis. In International Journal of Nuclear Knowledge Management, Vol. 2, pp. 299-312. 2007.

[30] E. Borgonovo. A New Uncertainty Importance Measure. In International Journal of Nuclear Knowledge Management, Vol. 92, pp. 771-784. June 2007.

[31] Storlie CB, Helton JC. Multiple Predictor Smoothing Methods for Sensitivity Analysis:

Description of Techniques. Reliability Engineering and System Safety 2008; 93(1): 28-54.

[32] Storlie CB, Helton JC. Multiple Predictor Smoothing Methods for Sensitivity Analysis: Example Results. Reliability Engineering and System Safety 2008; 93(1): 55-77.

[33] E. Zio, F. Di Maio and J. Tong. Safety Margins Confidence Estimation for a Passive Residual Heat Removal System. In Reliability Engineering and System Safety, Vol. 95, pp. 828–836, 2010.

[34] Storlie CB, Swiler LP, Helton JC, Sallaberry CJ. Implementation and Evaluation of Nonparametric Regression Procedures for Sensitivity Analysis of Computationally Demanding Models. Reliability Engineering and System Safety 2009; 94(11): 1735-1763.

[35] Storlie CB, Reich BJ, Helton JC, Swiler LP, Sallaberry CJ. Analysis of Computationally Demanding Models with Continuous and Categorical Inputs. Reliability Engineering and System Safety 2013; 113(1): 30-41.

[36] M. H. C. Law, M. A. T. Figueiredo and A. K. Jain. Simultaneous Feature Selection and Clustering Using Mixture Models. In IEEE Transactions on Patterns Analysis and Machine Intelligence, Vol. 26., n. 9. IEEE, September 2004.

[37] P. Diaconis and S. L. Zabell. Updating Subjective Probability. In Journal of the American Statistical Association, Vol. 77, pp. 822-830. December 1982.

[38] A. L. Gibbs and F. E. Su. On Choosing and Bounding Probability Metrics. In International Statistical Review, Vol. 70, pp. 419-435. December 2002.

[39] F. Di Maio, G. Nicola, E. Zio and Y. Yu. Ensemble-Based Sensitivity Analysis of a Best Estimate Thermal-Hydraulics Model of a Passive Containment Cooling System of an AP1000 Nuclear Power Plant. In Annals of Nuclear Energy, Vol. 73, pp. 200-210. November 2014.

[40] S. Carlos, A. Sánchez, D. Ginestar and S. Martorell. Using Finite Mixture Models in Thermal-Hydraulics System Code Uncertainty Analysis. In Nuclear Engineering and Design, Vol. 262, pp. 306-318. September 2013.

[41] G. McLachlan and D. Peel. Finite Mixture Models, 456 pages. John Wiley & Sons Inc., November 2000.

[42] F. Di Maio, G. Nicola, E. Borgonovo, E. Zio, “Invariant Methods for an ensemble-based Sensitivity Analysis of a Passive Containment Cooling System of an AP1000 Nuclear Power Plant”, Reliability Engineering and System Safety, Volume 151, pp. 12–19, 2016.

[43] M. Perez, F. Reventos, L. Batet, A. Guba, I. Tóth, T. Mieusset, P. Bazin, A. de Crécy, S. Borisov, T. Skorek, H. Glaeser, J. Joucla, P. Probst, A. Ui, B.D. Chung, D.Y. Oh, R. Penica, M. Kyncl, J. Macek, A. Manera, J. Freixa, A. Petruzzi, D. D’Auria and A. Del Nevo. Uncertainty and Sensitivity Analysis of a LBLOCA in a PWR Nuclear Power Plant: Results of the Phase V of the BEMUSE Programme. In Nuclear and Engineering Design, Vol. 241, pp. 4206-4222. October 2011.

[44] M. Perez, F. Reventos, L. Batet, A. Guba, I. Tóth, T. Mieusset, P. Bazin, A. de Crécy, S. Borisov, T. Skorek, H. Glaeser, J. Joucla, P. Probst, A. Ui, B.D. Chung, D.Y. Oh, R. Penica, M. Kyncl,

J. Macek, A. Manera, J. Freixa, A. Petruzzi, D. D'Auria and A. Del Nevo. Main Results of Phase IV BEMUSE Project: Simulation in an NPP. In Science and Technology of Nuclear Installations, Vol. 2010, Article ID 219294. 2010.

[45] S. C. Alberola. Simulation of a LBLOCA Performed for the Zion 1 NPP by the TRACE Code. Private Communication, 2014.

[426] Dempster, A., Laird, N., Rubin, B., 1977. Maximum likelihood estimation from incomplete data via EM algorithm. J. R. Stat. Soc. B 39, 1–38.

[47] F. Di Maio, G. Nicola, E. Zio, Y.Yu. Finite Mixture Models for Sensitivity Analysis of Thermal Hydraulic Codes for Passive Safety Systems Analysis. Nuclear Engineering and Design, 289, 144–154, 2015.

[48] P. Pudil, J. Novovicova and J. Kittler. Feature Selection Based on the Approximation of Class Densities by Finite Mixtures of the Special Type. In Pattern Recognition, Vol. 28, pp. 1389-1398. September 1995.

[49] S. Vaithyanathan and B. Dom. Generalized Model Selection for Unsupervised Learning in High Dimensions. In Advances in Neural Information Processing, pp. 970-976. MIT Press, 1999.

[50] R. C. Prati. Combining Feature Ranking Algorithms Through Rank Aggregation. In The 2012 International Joint Conference on Neural Networks (IJCNN), pp. 1-8. IEEE, June 2012.

[51] M. Figueiredo and A. K. Jain. Unsupervised Learning of Finite Mixture Models. In IEEE Transactions on Patterns Analysis and Machine Intelligence, Vol. 42, pp. 381-396. IEEE, March 2002.

[52] 40.P. Baraldi, F. Di Maio, E. Zio, "Unsupervised Clustering for Fault Diagnosis in Nuclear Power Plant Components", International Journal of Computational Intelligence Systems, Vol. 6, No. 4, pp. 764-777, 2013.