

Research Article

Bootstrap and Order Statistics for Quantifying Thermal-Hydraulic Code Uncertainties in the Estimation of Safety Margins

Enrico Zio and Francesco Di Maio

Department of Nuclear Engineering, Polytechnic of Milan, Via Ponzio 34/3, 20133 Milano, Italy

Correspondence should be addressed to Enrico Zio, enrico.zio@polimi.it

Received 30 April 2007; Revised 8 December 2007; Accepted 21 December 2007

Recommended by Cesare Frepoli

In the present work, the uncertainties affecting the safety margins estimated from thermal-hydraulic code calculations are captured quantitatively by resorting to the order statistics and the bootstrap technique. The proposed framework of analysis is applied to the estimation of the safety margin, with its confidence interval, of the maximum fuel cladding temperature reached during a complete group distribution blockage scenario in a RBMK-1500 nuclear reactor.

Copyright © 2008 E. Zio and F. Di Maio. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The traditional safety assessment of nuclear power plants aims at verifying that a number of acceptance criteria are satisfied for a small set of design basis accidents (DBAs). The acceptance criteria are enforced by regulation to ensure that sufficient safety margins are maintained for the defense-in-depth barriers which have been devised to protect from the consequences of the DBAs. The traditional assessment of the safety margins is achieved by performing calculations that are conservative (i.e., pessimistic) both in the hypotheses and models. Such verifications of the existence of “safety margins” are the sufficient condition required on the integrity of the protective barriers.

Currently, this traditional and conservative approach is being challenged by a more realistic analysis, so called best estimate (BE), which aims at adopting realistic models and hypotheses in the calculations of the safety margins associated to the accident scenarios bounding the system response. On one side, the reduction in the conservatism of the system analyses leads to a more efficient plant design and operation. On the other side, the relaxation of the conservatism in the analysis entails that sensitivity and uncertainty analyses are carried out to properly quantify the safety margins while capturing the associated uncertainty. This requires a

revision in probabilistic terms of the concept of safety margins [1, 2] and repeated model runs for the associated sensitivity and uncertainty analyses. Independently of the type of nuclear power plant, these latter pose a significant computational burden, due to the very large computing times of the mechanistic codes used.

Furthermore, the risk-informed decision-making trend to regulation [3] has opened a strong debate on the different role of deterministic and probabilistic approaches in regulatory matters of concern to NPP safety [4]. Historically, the former one has been adopted to verify the robustness of the designed barriers of protection against DBAs, while the latter one has been adopted to verify the limit risk posed by the so-called beyond design basis accidents (BDBAs), which are extreme conditions not enveloped by the DBAs. Nowadays, it is more and more believed that deterministic and probabilistic methods could benefit from each other and that safety analyses could benefit by their integration into one framework.

In particular, the recognized existence of uncertainties in the deterministic codes used for analyzing DBA scenarios leads to the need of estimating safety margins within a probabilistic framework. This is even more when analyzing requests for changes to the licensing bases of the risk-informed decision-making philosophy, because the combination of such changes with the uncertainties in the analysis

could reduce significantly and in an unexpected way the safety margins. This situation may increase the risk of accidents leading to major plant damage so that the uncertain circumstances under which this may occur need to be assessed and kept under control.

In this paper, the probabilistic framework for the calculation of safety margins is adopted by resorting to the order statistics [5, 6] to provide a given confidence on the reliability of the calculated point estimate obtained with a limited number of code runs.

The main originality of the work consists in the use of the bootstrap method [7] for computing the confidence intervals of the estimated safety margins, based on the same limited sample of code runs. This additional information provides a realistic refinement of the estimates that is beneficial to power plant owners; on the other hand, from the viewpoint of the regulatory body, it increases the robustness of the safety case by allowing verification of the fact that uncertainty in the safety margin estimates does not lead to exceedance of the regulatory safety thresholds.

The approach is exemplified on a case study regarding a DBA scenario in a RBMK-1500 nuclear reactor [8].

The paper organization is as follows. In Section 2, the basic principles underpinning the BE nuclear safety analysis in the presence of uncertainties are provided together with an illustration of the bootstrap method for uncertainty estimation. In Section 3, the main characteristics of the RBMK-1500 reactor are briefly introduced, the group distribution header (GDH) complete blockage accident scenario is described, and the RELAP5/MOD3.2 simulations performed to analyze the system response to the accident scenario are presented. In Section 4, the results of the application of the proposed framework for the estimation of the safety margin of the maximum fuel cladding temperature reached during the accident described in Section 3 are provided. Finally, some conclusions are drawn in Section 5.

2. PROBABILISTIC APPROACH TO SAFETY MARGINS CALCULATION

2.1. The mathematical model

A quantitative model for safety analysis of a nuclear power plant may be viewed as composed of three main elements: an input vector $\bar{x} = \{x_1, x_2, \dots, x_m\}$, a BE simulator code, and an output vector $\bar{y} = \{y_1, y_2, \dots, y_p\}$. The elements of the input vector \bar{x} are all the model parameters and input variables needed to calculate one realization of the output variables \bar{y} describing the system response. The simulation code can be regarded as a *black box* which implements the complex, multidimensional, nonlinear function that maps the input vector \bar{x} into the output vector \bar{y} [9]:

$$\bar{y} = T(\bar{x}). \quad (1)$$

For fixed values of \bar{x} , the output values \bar{y} are deterministically computed.

In practice, the input vector \bar{x} is uncertain so that the output vector \bar{y} is also uncertain (see Figure 1).

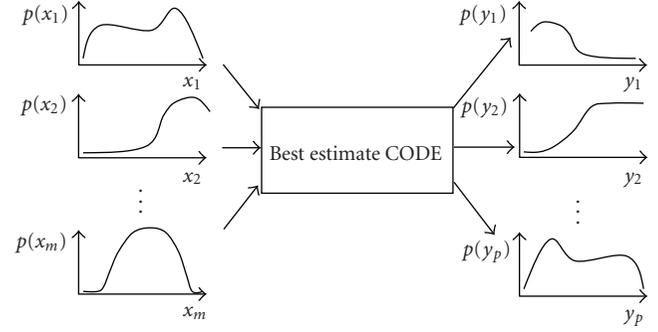


FIGURE 1: Representation of the best estimate code as a black box.

Two different sources of uncertainty are typically considered: randomness due to inherent variability in the system behavior (aleatory uncertainty) and imprecision due to lack of knowledge and information on the system (epistemic uncertainty) [10, 11]. The present work addresses the latter.

The best estimate of input values, vector \bar{x}_0 , defines the nominal state and is typically assumed to represent the expected values $E[\bar{x}]$ of the uncertain input parameters. The output vector corresponding to the nominal state, $\bar{y}_0 = T(\bar{x}_0)$, is the best estimate output and of course is expected to represent a largely safe state well within the safety margins. In mathematical terms, this means that the calculated values of the output variables $\{y_{10}, y_{20}, \dots, y_{p0}\}$ belong to a given set Ψ of predefined safety intervals:

$$\Psi = \{[L_j, U_j], j = 1, \dots, p\} \quad (2)$$

with L_j and U_j being the lower and the upper bounds allowed for the j th safety parameter, respectively. In general, the generic system state represented by the input vector \bar{x} is considered safe if the corresponding output response \bar{y} falls at all times within the safety envelope Ψ defined by (2). By logical extension, in the face of uncertainty affecting the input values, the plant can be regarded safe if its response \bar{y} falls at all times in Ψ for every $\bar{x} \in X$, where X is the set of all possible input vectors \bar{x} distributed according to a given probability distribution [9].

Verification of the above safety conditions entails generating a representative sample of independent input vectors $\bar{x}^{(i)} \in X$, $i = 1, 2, \dots, N$, and running the code for each input vector to generate the corresponding output $\bar{y}^{(i)}$, $i = 1, 2, \dots, N$.

The sample of N independent output vectors $\{\bar{y}^{(1)}, \bar{y}^{(2)}, \dots, \bar{y}^{(N)}\}$ thereby obtained carries information on the probability distribution of the safety output response \bar{y} . From this information, one wants to draw conclusions on the safety conditions of the plant, with the required confidence.

In this work, this objective will be pursued with reference to the accidental scenario of the RBMK-1500 nuclear power plant presented in Section 3. As we will see, in this case, the output variables uncertainty results from the lack of knowledge of some involved input parameter values, physical phenomena, and their modeling [8].

2.2. Formulation of safety margins in presence of uncertainty

The defense-in-depth principle of nuclear safety is founded on the setup of protective barriers between the radioactive products and the environment. Each barrier is a physical device whose integrity is measured with reference to given characteristic safety parameters. When a barrier is subjected to abnormal conditions, some of the related safety parameters may exceed their safety envelope, which results in the failure of the barrier.

With reference to a generic accident scenario k and a characteristic safety variable $y_j(k)$ to be limited from above by an upper threshold limit U_j , the traditional safety margin $M(y_j, k)$ can be defined as follows [2]:

$$M(y_j, k) = \begin{cases} \frac{U_j - y_j}{U_j - y_{j \text{ ref}}} & \text{if } y_{j \text{ ref}} \leq y_j \leq U_j, \\ 0 & \text{if } U_j < y_j(k), \\ 1 & \text{if } y_j(k) < y_{j \text{ ref}}, \end{cases} \quad (3)$$

where $y_{j \text{ ref}}$ is a reference value for $y_j(k)$, typically the best estimate y_{j0} .

The dual definition of the safety margin for a safety parameter y_j to be limited from below by a threshold value L_j is straightforward.

Since $y_j(k)$ is a random variable, because of the function of the input random variables \bar{x} , also $M(y_j, k)$ is a random variable whose values range in the interval $[0, 1]$ by construction.

2.3. Estimation of safety margins

The safety margins are calculated by running the thermal-hydraulic codes used for safety analysis. In presence of uncertainty, a large number of runs of the code may be required to adequately represent the full distribution of the safety parameters in the output. The problem is that the computer runs of the complex models of plant dynamics used for safety analysis are computationally very expensive. To overcome this hurdle, one may resort either to simplified analytical/numerical models, such as those based on lumped effective parameters [12] or to empirical models, for example, artificial neural networks and fuzzy logic systems [13], suitably set up so as to best fit the data available from the plant.

Another possibility to reduce the computational burden, which may be used in combination with the former ones, is to only compute some percentiles of the output distribution, estimated with a limited number of runs. In this case, the confidence in the estimates becomes crucial for decision making and must, thus, be quantified [5, 6, 9, 14].

This latter approach is undertaken in this work and the problem of confidence building and quantification is addressed. A sample of a small number N of input parameter values is drawn by the Monte Carlo method from their probability distributions. The sample of N input vectors thereby obtained is the input to the code which is correspondingly run N times, thus producing a random sample of N output vectors. These can be used to estimate a given percentile of

the safety margin probability distribution. To obtain the desired confidence in the safety margin percentile, the number N of code runs is defined on the basis of the order statistics (OS) methodology, in its nonparametric formulation [5, 6] which applies independently from the type of probability distribution of the output data under study (in this case unknown). As we will see, this amounts to ordering the elements of the random sample by increasing value, the element in the r th place being the statistic of order r , and using the order statistics for estimating the percentiles of the distribution (Section 2.4) with the desired confidence. Following this methodology, the number of runs required can be kept low because only statistical intervals are estimated and not the full probability distributions of the output.

2.4. Estimation of percentiles using order statistics

For ease of illustration, let us refer the discussion to a one dimensional output y , for example, the Pellet Cladding Temperature (PCT).

The N runs of the code, each one with a different input vector $\bar{x}^{(i)}$, produce N output vectors $\bar{y}^{(i)}$, $i = 1, 2, \dots, N$. Let $Y = \{y^{(1)}, y^{(2)}, \dots, y^{(N)}\}$ be the ordered set of values resulting from running the code N times for N different input vectors $X = \{x^{(1)}, x^{(2)}, \dots, x^{(N)}\}$. If the codes were run a very large number of times ($N \rightarrow \infty$), it could be possible to give a sufficiently accurate estimate of the full distribution of the output y and draw probabilistic conclusions on where it lies with respect to the threshold value U (or L) defining the safety limit for the integrity of the protective barriers. This would provide a more realistic assessment than the verification that a single run conservative estimate of y is within the safety envelope (e.g., PCT is less than 2200° F).

Given the computational costs associated with the estimation of the full distribution, one is forced to focus on verifying that with some level of confidence β , a certain percentage γ of the calculated values of y that would be obtained from running the code falls within the safety envelope. The thresholds defining the safety envelope and the values of the confidence and percentage are set by the regulations, considering the risk associated with exceeding the specified range.

With reference to the safety parameter y to be limited from above by U , the approach aims at showing that the m th member y_m of the N sorted outputs has a certain probability β of exceeding the unknown true γ th percentile y_γ [14, 15]. Then, one has a level of confidence β that the actual value of y_γ is less than the value obtained for y_m ; if y_m meets the criterion of being less than the safety threshold U , then the unknown y_γ will do so, too.

2.5. Bootstrapped confidence intervals

Bootstrap methods [7] can be used for constructing confidence intervals for safety margins estimates. In general, given a random sample of size N , it is often of interest to draw inferences about the properties of the distribution of a generic sample estimate $\hat{\theta}$, for example, its confidence intervals. When the sampling distribution of the estimate $\hat{\theta}$ is unknown, one may resort to the bootstrap, a distribution-free

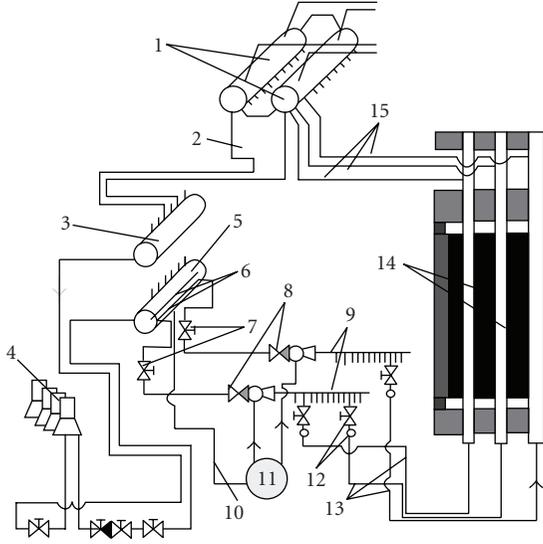


FIGURE 2: Ignalina NPP model nodalization scheme [21].

inference method which requires no prior knowledge about the distribution function of the underlying population [16].

The bootstrap is a computational intensive, nonparametric technique for assessing the accuracy of a parameter estimator which requires very little assumptions or analysis [17]. The basic idea is to generate a reference distribution from the observed data by performing sampling with replacement [18].

Within the order statistics framework, the bootstrap technique builds on the following elements [18, 19].

- (1) An original data sample $Y = \{y^{(1)}, y^{(2)}, \dots, y^{(N)}\}$, with N given from the OS framework with selected m , β , and γ values, is collected from an unknown cumulative distribution function $P(y)$.
- (2) The empirical cumulative distribution function, estimate of $P(y)$, can be constructed from the sample Y , by assigning probability $1/N$ to each value $y^{(i)}$, $i = 1, \dots, N$ in the sample.
- (3) The plug-in principle. Any characteristics of $P(y)$ (e.g., its mean $\theta = E[y]$) can be estimated using the corresponding characteristics of the empirical distribution function. In the case of interest here, θ is the true γ th percentile and $\hat{\theta}$ is its estimate, given with a confidence β .
- (4) Bootstrap sample. A bootstrap sample $Y^* = \{y^{(1)*}, y^{(2)*}, \dots, y^{(N)*}\}$ is defined to be a random independent and identically distributed (i.i.d.) sample of size N drawn from the empirical distribution function with replacement. The sample Y^* can be considered a randomly resampled version of Y . The elements of the bootstrap sample are the same as those of the original data set: some may appear only once in the bootstrap sample, some two or more times, while some others may not appear at all. A large number of independent bootstrap samples are easily generated using a random number generator to select integers between 1 and N with probability $1/N$.

The computer time required for bootstrap replications increases linearly with the number of samples B [16]. In general, 100–200 bootstrap replications are often enough to obtain a good estimate $\hat{\theta}$ [7]. However, much larger values ($B \geq 1000$) are required if one wants to estimate two-sided confidence intervals for θ [16]. Clearly, increasing the number N of the original data in the sample Y allows a more robust estimation.

For the application of interest in the present work, the confidence intervals for the γ th percentile are computed using the γ th percentile estimates obtained from the bootstrap replications through the OS method. To verify that the estimated confidence interval brackets the unknown true percentile of the output distribution, two-sided confidence intervals are computed.

From the B bootstrap replications $\hat{\theta}_b^*$, $b = 1, 2, \dots, B$, the following order statistics are determined: $\hat{\theta}^{*(1)} \leq \hat{\theta}^{*(2)} \leq \dots \leq \hat{\theta}^{*(B)}$. The desired $100(1 - \gamma)\%$ percentile is then $[\hat{\theta}^{*(r)}, \hat{\theta}^{*(s)}]$, where $r = [B\gamma/2]$ and $s = [B(1 - \gamma/2)]$, and $[a]$ denotes the integer part of a . For example, if $B = 1000$, the 95% confidence interval is defined by $[\hat{\theta}^{*(25)}, \hat{\theta}^{*(975)}]$, that is, the 2.5th and 97.5th percentiles of the empirical cumulative distribution function [19, 20].

Extension of the bootstrap procedure to p multiple outputs y_1, y_2, \dots, y_p is straightforward; each element of the sample vector $\bar{Y} = \{Y_1, Y_2, \dots, Y_p\}$, constituted by N samples, has to be resampled B times and these replications are used separately to draw conclusions on each output parameter.

3. THE GDH COMPLETE BLOCKAGE ACCIDENT SCENARIO IN THE RBMK-1500

3.1. The RBMK-1500

The description of the RBMK-1500 makes references to the Ignalina nuclear power plant (NPP) in Lithuania, which operates two units commissioned in 1984 and 1987 [21]. The RBMK-1500 is a graphite-moderated, boiling water, and multichannel reactor. The units at Ignalina are designed to provide a saturated steam at 7.0 MPa. The reactor design of maximum reactor rating is 4800 MWth. Several design features of the RBMK-1500 are rather unique with respect to reactors of western design. The main distinguishing characteristic of the RBMK-type reactor is that each core fuel assembly is housed in an individual pressure tube.

The RBMK-1500 core contains 1661 pressurized fuel channels (FCs), inserted in a graphite block lattice. The main circulation circuit (MCC) is divided into two loops. Figure 2 presents one such loop. The loop has two drum separators (1), which separate steam from the steam-water mixture coming from the core. Eight main circulation pumps (MCPs) are used for the circulation of the cooling water. The MCPs (4), joined in groups of four pumps on each loop (three for normal operation and one in standby), feed the pressure headers (5) on each side of the reactor. Each pressure header provides coolant upwards through the reactor core block to

20 group distribution headers (GDHs) (9). Each of them in turn feeds from 38 to 43 pressurized fuel channels (14) and is connected to the bypass pipeline (10) of the emergency core cooling system (ECCS). The flow in the channel is measured by a ball flow meter (12) and regulated by means of isolation (12) and control valves (7). By flowing through the core, the coolant absorbs about 95% of the total energy released by the fuel elements. The steam-water mixture generated in the FCs flows through the steam-water pipes (15) to the drum separators (DSs) [22].

3.2. The GDH complete blockage scenario

The accident scenario considered in the present study consists of a complete blockage of a GDH, which is one of the scenarios considered in the RBMK reactor safety analysis [8]. This leads to a temporary decrease in the coolant flow rate in the FCs and to a corresponding increase in the cladding temperature which may reach values close to the maximum allowable limit of 973.15 K [8]. For this reason, this scenario is considered safety relevant and a careful analysis of it must be performed [23].

3.3. RELAP simulations of the accident scenario

Given the above mentioned safety concerns related to the GDH blockage scenario, a RELAP5/MOD3.2 model of the RBMK-1500 reactor has been implemented and used to simulate 480 accidental complete blockage transients, generated by sampling the involved input signals, parameters, and initial conditions from proper probability distributions suggested from previous experience and/or skilled operators [8]. The model is assumed to be “sufficiently best estimate” with the nominal (“best estimate”) values of the input parameters. The interested reader may refer to the original works for further details on the RELAP5/MOD3.2 model implementation [24].

In the accident transient simulations, it is assumed that the reactor operates at the stationary power of 2900 MWth. The coolant is supplied through the core by two MCPs in each MCC loop, up to the beginning of the accident (i.e., before the GDH blockage). This reactor state is chosen because in such conditions the reactor cooling of the core is the most complicated. Note that 2900 MWth is the maximum allowable power level when four MCPs are in operation, that is, the worst power and coolant flow rate ratio is conservatively considered in the analysis. For the calculations, it is assumed that the coolant flow rate through the connection pipeline into the GDH is completely stagnated within 0.01 second. This stagnation results in a drop of the pressure in the failed GDH. Under the influence of the pressure difference, the coolant from a MCP pressure header flows through the bypass pipelines into the ECCS header and it is then directed into the failed GDH through ECCS pipelines. In these conditions, fuel channels connected to the GDH are cooled only by the water supplied through the ECCS bypass pipelines. Thus, the fuel cladding and pressure tube wall temperatures sharply increase and then start to decrease after reactor shutdown, which is initiated for protection against the decrease in coolant flow rate through the GDH.

Since the peak of the fuel cladding temperature in the maximum power channel may reach values close to the acceptance criteria temperature (973.15 K for fuel cladding), this parameter is chosen as the model output of interest y , which is thus independent of time, and the relative safety margin is estimated.

The relevant model inputs \bar{x} which may influence the fuel cladding temperature behavior may be divided in [8] to

- (1) initial conditions (coolant pressure, temperature, and flow rate or power),
- (2) RELAP5/MOD3.2 code model parameters, assumptions, and correlations (e.g., different correlations for the calculation of friction loss and heat transfer).

A first preliminary investigation based on engineering judgment led to the identification of the following initial conditions as most important for the postulated accident analysis [8, 21]:

- (i) pressure in the drum separator (continuous variable x_1),
- (ii) coolant flow rate through the MCPs (continuous variable x_2),
- (iii) feed water temperature (continuous variable x_3),
- (vi) amount of steam for in-house needs (continuous variable x_4),
- (v) reactor thermal power (continuous variable x_5).

The uncertainty in modeling has also been challenged by considering alternatives to some of the following hypotheses:

- (1) water packing: it specifies whether the scheme of volume filling with water is to be used (binary variable x_6);
- (2) vertical stratification: it specifies whether the model of two-phase media vertical stratification is enabled or disabled (binary variable x_7);
- (3) modified PV term in the equations: it specifies whether the modified potential pressure energy model is applied or not (binary variable x_8);
- (4) CCFL (counter current flow limit): on/off (binary variable x_9);
- (5) thermal front tracking: on/off (binary variable x_{10});
- (6) mixture level tracking: on/off (binary variable x_{11}).

A last parameter which might have an impact on the temperatures of the fuel cladding and fuel channel wall is the signal to start the reactor protection against the accident initiation, which determines the time when the reactor is shutdown in the fast mode (continuous variable x_{12}). It is assumed that insertion of CPS rods is delayed by 1 second, with respect to the signal.

In conclusion, a total of twelve parameters are considered relevant. Of these, six are continuous (x_1 – x_5 and x_{12}) and six are binary (x_6 – x_{11}). The continuous parameter distributions have been obtained from error specifications in measure devices and from skilled-operator expertise, whereas the binary parameters are set as RELAP5/MOD3.2 inputs, with an arbitrary probability of 0.5 (see Table 1).

TABLE 1: Parameters which are regarded relevant with respect to the behaviour of the fuel cladding temperature [21].

Parameter	Range of parameter distribution		Mean	Standard deviations	Probability distribution
	Min	Max			
Initial conditions					
x_1 (pressure in DS) [Pa]	$6.79E+06$	$6.93E+06$	$6.86E+06$	$3.43E+04$	Normal
x_2 (coolant flow rate) [m ³ /h]	6860	7140	7000	70	Normal
x_3 (feedwater temperature) [K]	458.52	467.78	463.15	2.32	Normal
x_4 (steam for in-house needs) [m ³ /h]	227.7	232.3	230	1.15	Normal
x_5 (reactor thermal power) [W]	$2.81E+09$	$2.99E+09$	$2.90E+09$	$4.5E+07$	Normal
Modelling parameters					
x_6 (water packing)	0 (on)	1 (off)	1 (off)	—	Discrete
x_7 (stratification)	0 (on)	1 (off)	0 (on)	—	Discrete
x_8 (PV term)	0 (on)	1 (off)	0 (off)	—	Discrete
x_9 (CCFL)	0 (on)	1 (off)	0 (off)	—	Discrete
x_{10} (thermal front tracking)	0 (on)	1 (off)	0 (off)	—	Discrete
x_{11} (mixture level tracking)	0 (on)	1 (off)	1 (on)	—	Discrete
x_{12} (scram initiation time)	5.3	6.3	6.0	0.25	Normal

4. RESULTS

The order statistics approach and the bootstrap method illustrated in Section 2 have been applied for the estimation of the safety margin of the maximum fuel cladding temperature reached during the accident scenario of complete blockage of a group distribution header (GDH) of the RBMK-1500 nuclear reactor described in Section 3.

Order statistics approach has been applied to a sample of maximum fuel cladding temperatures obtained by simulation, for estimating the γ th percentile with $\gamma = 0.95$. Then, B replications of the original sample have been obtained by the bootstrap method and correspondingly, B estimates of the γ th percentile have been collected by the OS applied to each bootstrapped sample. Finally, confidence intervals for the real γ th percentile have been evaluated.

The total number of patterns available from the RELAP5/MOD3.2 simulations described in Section 3.3 is 480. All 480 samples available have been used to build the reference empirical distributions. In Figures 3 and 4, the empirical cumulative distribution function and the histogram constructed from the simulation data are shown. These distributions are assumed as reference distributions for comparison with the results obtained with the proposed method of estimation of the percentiles defining the safety margin.

In practice, smaller samples (e.g., of 59, 93, and 311 data) are used with OS for percentile estimation and these are indeed the values which will be considered here.

In the OS analysis, the values of the confidence β and coverage γ have been chosen to be both equal to 0.95. From the empirical distribution functions, the 95th percentile y_{95} turns out to be 958.00 K. On the other hand, by regulation the pellet-cladding safety threshold limit U is set equal to 973.15 K [8]. Furthermore, the reference pellet-cladding temperature y_{ref} has been taken equal to the sample mean of the 480 RELAP5/MOD3.2 values at the initial time of the accident transients; this turns out to be 570.15 K.

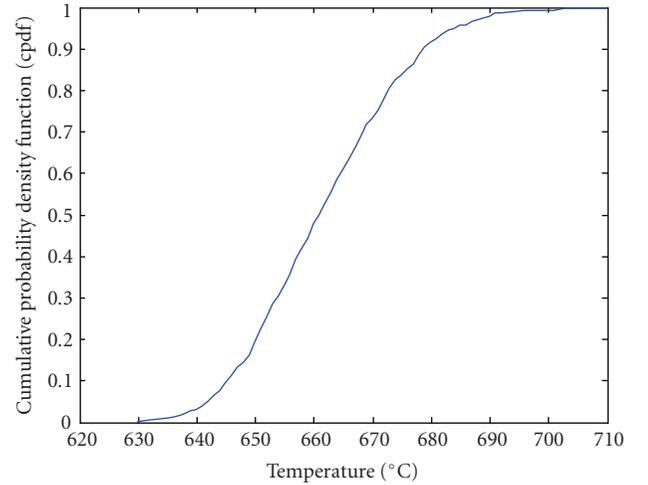


FIGURE 3: Empirical distribution function constructed from 480 RELAP5/MOD3.2 simulations.

From the entire sample of 480 accidental transients, N have been uniformly sampled, with $N = 59, 93, 311$. According to OS theory, the number m of values in the sample that are at least required to lie beyond the “extent” γ of the cumulative probability, define the sample size N needed to assure the desired values of confidence β and coverage γ (see Table 2) [14]. Here the values for $m = 1, 2, 10$ and $\beta = \gamma = 0.95$ have been considered for test.

The results for the point estimation of the 95th percentile obtained by the OS method applied to the sample $Y = \{y^{(1)}, y^{(2)}, \dots, y^{(N)}\}$ of the maximum pellet-cladding temperatures reached during the accident scenario are given in Table 3.

The null safety margin obtained for $m = 1$ is due to one run among the $N = 59$ that is above the threshold limit $U = 973.15$ K. This leads to a very conservative estimate: increasing m and correspondingly, increasing the number of

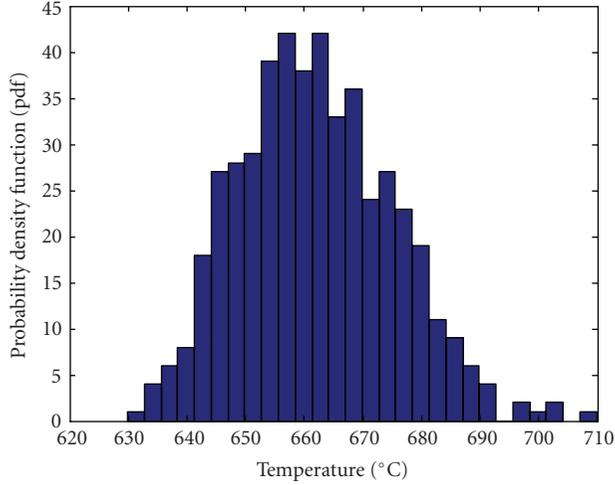


FIGURE 4: Histogram constructed from 480 RELAP5/MOD3.2 simulations.

TABLE 2: Sample size N that guarantees a confidence value $\beta = 0.95$ and a coverage value $\gamma = 0.95$, for different numbers m of values lying above y_{95} .

	Sample size: N
$m = 1$	59
$m = 2$	93
$m = 10$	311

TABLE 3: Results for the point estimation of the 95th percentile and the corresponding safety margin obtained through the order statistics method.

	95th percentile estimation	Empirical coverage	Safety margin
$m = 1$	976.00 K	0.998	0.00
$m = 2$	962.00 K	0.975	$2.77E - 02$
$m = 10$	962.00 K	0.975	$2.77E - 02$

values that are requested at least to lie beyond the “extent” y of the cumulative probability; the estimation \hat{y}_{95} of the percentile tends to narrow to the “true” y_{95} (958.00 [K]), thus providing a less conservative estimate of the safety margin.

Incidentally, the estimates for $m = 2$ and $m = 10$ agree, due to the large number of the 480 RELAP5/MOD3.2 simulations that reach a maximum pellet-cladding temperature of 962.00 K, which, thus, has a large probability of appearing in the random sample.

Because of the limitation on the sample size used in the estimation, the safety acceptance criteria cannot be based solely on the best estimate results. Hence, the uncertainty of the estimated safety margin must be properly informed, for example, by computing its confidence interval.

This is achieved by applying the bootstrap algorithm for the estimation of the confidence interval of the 95th percentile, with a number of bootstrap samples $B = 1000$. As an example, Figure 5 shows, for the case of $N = 59$, the his-

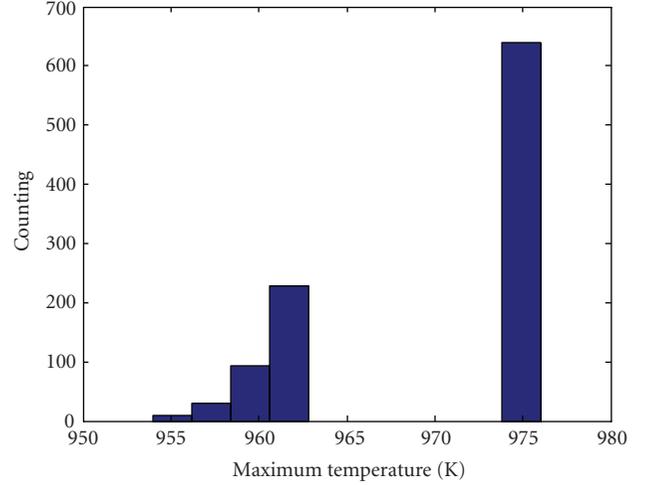


FIGURE 5: Histogram of the $B = 1000$ bootstrap estimates of the 95th percentile for $N = 59$.

ogram of the B estimated values of the 95th percentile which are strongly not normally distributed.

Table 4 reports the confidence interval obtained with the bootstrap procedure applied to the estimates of the 95th percentiles and the safety margin of Table 3. Increasing m increases the confidence on the estimated percentile value, as demonstrated by the shrinking of the confidence interval.

Notice that even if the safety margin estimates for $m = 2$ and $m = 10$ coincide, their confidence interval values disagree due to the narrowing of the bootstrap percentile estimates, that determine the confidence intervals, to the value $y_{95} = 958.00$ K.

Furthermore, increasing m , the estimate of the confidence interval becomes statistically more reliable itself because in each bootstrapped sample the number of values used for the ensemble averages becomes larger.

Considering the most reliable estimation with $N = 311$ and $m = 10$, \hat{y}_{95} turns out to be equal to 962.00 K, higher than $y_{95} = 958.00$ K. Because the estimate \hat{y}_{95} bracketed by its confidence interval $[956.00, 976.00]$ meets the safety criterion of being less than $U = 973.15$ K, one can be 95% confident that the pellet-cladding temperature will be below such value at least 95% of the cases.

5. CONCLUSIONS

In this work, an approach for estimating safety margins has been purported, accounting for the uncertainties associated in the computation of the relevant safety parameters. Order statistics is exploited to support a limited number of calculations. An original utilization of the bootstrap method allows computing confidence intervals on the percentile estimates. This additional information refines the estimates while at the same time makes it more robust for regulatory decisions.

The procedure has been applied to the maximum fuel cladding temperature reached during a complete group distribution blockage scenario in a RBMK-1500 nuclear reactor. The method has been demonstrated effectively in that it

TABLE 4: Estimated 95th percentile, corresponding confidence interval, and safety margin confidence interval obtained by the bootstrap method with $B = 1000$.

	95th percentile estimation	95th percentile confidence interval	Safety margin confidence interval
$m = 1$	976.00 K	[958.00, 976.00]	[0.00, $3.76E - 02$]
$m = 2$	962.00 K	[956.00, 976.00]	[0.00, $4.26E - 02$]
$m = 10$	962.00 K	[956.00, 964.00]	[$2.27E - 02$, $4.26E - 02$]

is capable of indicating the system safety conditions properly, accounting for the uncertainties in the model parameters and in the estimate itself. The application confirms that the proper number of scenarios N upon which the relevant estimates are based is fundamental for building the required confidence in the safety of the system.

NOTATION AND LIST OF ACRONYMS

OS:	Order statistics
BE:	Best estimate
NPP:	Nuclear power plant
RBMK:	Reaktor Bolshoi Moshchnosty Kanalny
DBA:	Design basis accident
BDBA:	Beyond design basis accident
GDH:	Group distribution header
MCC:	Main circulation circuit
MCP:	Main circulation pump
DS:	Drum separator
PCT:	Pellet-cladding temperature
ECCS:	Emergency core cooling system
\bar{x} :	Input values vector
x_m :	m th element of the input vector
\bar{y} :	Output values vector
y_p :	p th element of output vector
$T(\bar{x})$:	Function that maps the input vector \bar{x} into the output vector \bar{y}
\bar{x}_0 :	Nominal state of the input values vector
$E[\bar{x}]$:	Expected values of the uncertain input parameters
\bar{y}_0 :	Nominal state of the output vector
y_{p0} :	p th element of the nominal output vector
Ψ :	Set of predefined safety intervals
L_j :	Lower threshold for the j th output parameter
U_j :	Upper threshold for the j th output parameter
$\bar{x}^{(i)}$:	i th element of the representative sample of independent input vectors
$\bar{y}^{(i)}$:	i th element of the representative sample of independent output vectors
K :	Index of the accident scenario
$M(y_j, k)$:	Safety margin relative to the j th safety parameter for the k th accident scenario
$y_j(k)$:	j th safety parameter relative to the k th accident scenario
$y_{j \text{ ref}}$:	Reference value for $y_j(k)$, typically the best estimate y_{j0}
N :	Number of simulations
r :	Position of the ordered sample of simulations
Y :	Ordered set of values resulting from running the code N times

β :	Confidence value
γ :	Coverage value
m :	Number of values that lie beyond the γ coverage extent
y_γ :	Real γ th percentile
y_m :	Estimated γ th percentile
$\hat{\theta}$:	Generic sample estimate
Y^* :	Bootstrap output values sample
$\hat{\theta}_b^*$:	Bootstrap estimate replication
y_{95} :	Real 95th percentile
\hat{y}_{95} :	95th percentile estimate

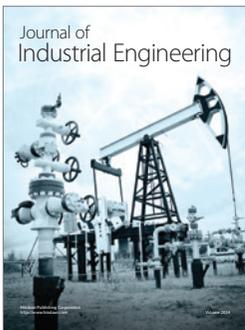
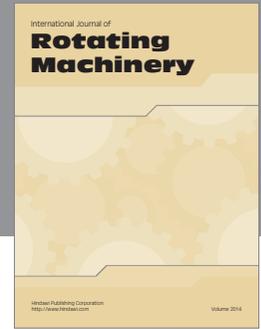
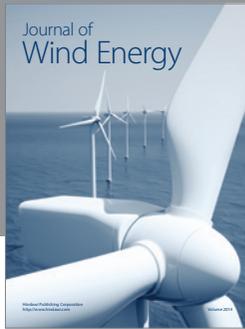
ACKNOWLEDGMENTS

The authors wish to thank Drs. Vytis Kopustinskas and Rolandas Urbonas of the Lithuanian Energy Institute for providing the transient simulation data and Professor S. Martorell of the Polytechnic University of Valencia for the insightful discussion. Also, the authors are thankful to the anonymous referees whose remarks have allowed significantly improving the paper.

REFERENCES

- [1] M. Gavrilas, J. Meyer, B. Youngblood, D. Prelewicz, and R. Beaton, "A generalized framework for assessment of safety margins in nuclear power plants," in *Proceedings of the International Meeting on Updates in Best Estimate Methods in Nuclear Installation Safety Analysis (BE '04)*, pp. 28–36, Washington, DC, USA, November 2004.
- [2] S. Martorell, Private Communication, 2007.
- [3] UNRSC, An approach for using probabilistic risk assessment in risk-informed decisions on plant-specific changes to the licensing basis, RG 1.174, 1998.
- [4] S. Martorell, Y. Nebot, J. F. Vilanueva, et al., "Safety margins estimation method considering uncertainties within the risk-informed decision-making framework," in *Proceedings of the American Nuclear Society's Topical Meeting on Reactor Physics (PHYSOR '06)*, vol. 2006, Vancouver, BC, Canada, September 2006.
- [5] S. S. Wilks, "Determination of sample sizes for setting tolerance limits," *Annals of Mathematical Statistics*, vol. 12, no. 1, pp. 91–96, 1941.
- [6] S. S. Wilks, "Statistical prediction with special reference to the problem of tolerance limits," *Annals of Mathematical Statistics*, vol. 13, no. 4, pp. 400–409, 1942.
- [7] B. Efron and R. J. Tibshiriani, *An Introduction to the Bootstrap, Monographs on Statistics and Applied Probability 57*, Chapman & Hall, New York, NY, USA, 1993.
- [8] V. Kopustinskas, R. Urbonas, A. Kaliakta, and R. Bolado-Lavin, "Input parameters distribution sensitivity study of RBMK-1500 reactor one group distribution header complete

- blockage scenario,” in *Proceedings of the European Safety & Reliability Conference (ESREL '05)*, pp. 1125–1132, Tri Cityá, Poland, June 2005.
- [9] A. Guba, M. Makai, and L. Pál, “Statistical aspects of best estimate method—I,” *Reliability Engineering and System Safety*, vol. 80, no. 3, pp. 217–232, 2003.
- [10] G. Apostolakis, “The concept of probability in safety assessments of technological systems,” *Science*, vol. 250, no. 4986, pp. 1359–1364, 1990.
- [11] J. C. Helton and W. L. Oberkampf, “Alternative representations of epistemic uncertainty,” *Reliability Engineering and System Safety*, vol. 85, no. 1–3, pp. 1–10, 2004.
- [12] M. Marseguerra, E. Zio, and R. Canetta, “Using genetic algorithms for calibrating simplified models of nuclear reactor dynamics,” *Annals of Nuclear Energy*, vol. 31, no. 11, pp. 1219–1250, 2004.
- [13] M. Marseguerra, R. Masini, E. Zio, and G. Cojazzi, “Variance decomposition-based sensitivity analysis via neural networks,” *Reliability Engineering and System Safety*, vol. 79, no. 2, pp. 229–238, 2003.
- [14] W. T. Nutt and G. B. Wallis, “Evaluation of nuclear safety from the outputs of computer codes in the presence of uncertainties,” *Reliability Engineering and System Safety*, vol. 83, no. 1, pp. 57–77, 2004.
- [15] A. Wald, “An extension of Wilks method for setting tolerance limits,” *Annals of Mathematical Statistics*, vol. 14, no. 1, pp. 45–55, 1943.
- [16] G. Heinrich, “Changing times, testing times: A bootstrap analysis of poverty and inequality using the PACO database,” CERT Discussion Papers 9802, Centre for Economic Reform and Transformation, Heriot Watt University, Edinburgh, UK, 1998.
- [17] A. M. Zoubir and B. Boashash, “The bootstrap and its application in signal processing,” *IEEE Signal Processing Magazine*, vol. 15, no. 1, pp. 56–76, 1998.
- [18] M. J. Fortin, G. M. Jacquez, and B. Shipley, “Computer-intensive methods,” in *Encyclopedia of Environmetrics*, A. El-Shaarawi and W. W. Piegorsch, Eds., vol. 1, pp. 399–402, John Wiley & Sons, Chichester, UK, 2002.
- [19] C. Z. Mooney and R. D. Duval, *Bootstrapping: A Nonparametric Approach to Statistical Inference*, Quantitative Applications in the Social Sciences, Sage, Newbury Park, CA, USA, 1993.
- [20] C. Swanepoel, “The teaching and practical implementation of the non-parametric bootstrap,” in *Proceedings of the 6th International Conference on Teaching Statistics (ICOTS 6 '02)*, Cape Town, South Africa, July 2002.
- [21] E. Ušpuras and A. Kaliatka, *Accident and Transient Processes at NPPs with Channel-Type Reactors: Monography*, vol. 28 of *Thermophysics*, Lithuanian Energy Institute, Kaunas, Lithuania, 2006.
- [22] A. Kaliatka, M. Vaisnoras, and V. Vileiniskis, “Best estimate analysis of group distribution header blockage events in RBMK-1500 reactors,” *Kerntechnik*, vol. 70, no. 5-6, pp. 270–276, 2005.
- [23] R. Urbonas, A. Kaliatka, and E. Ušpuras, “Comparison of modern analysis approaches based on the postulated GDH blockage event at Ignalina NPP,” *Energetika*, vol. 2, pp. 14–21, 2003.
- [24] R. Urbonas, E. Ušpuras, and A. Kaliatka, “State-of-the-art computer code RELAP5 validation with RBMK-related separate phenomena data,” *Nuclear Engineering and Design*, vol. 225, no. 1, pp. 65–81, 2003.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

