HAL
open science

How to effectively compute the reliability of a thermal-hydraulic nuclear passive system

Enrico Zio, Nicola Pedroni

To cite this version:

Enrico Zio, Nicola Pedroni. How to effectively compute the reliability of a thermal-hydraulic nuclear passive system. Nuclear Engineering and Design, 2011, 241 (1), pp.310-327. 10.1016/j.nucengdes.2010.10.029 . hal-00609623

HAL Id: hal-00609623

https://centralesupelec.hal.science/hal-00609623v1

Submitted on 27 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How to Effectively Compute the Reliability of a Thermal-Hydraulic Nuclear Passive System

E. Zio[1, 2] and N. Pedroni[2]
[1]*Ecole Centrale Paris-Supelec, Paris, France*
E-mail: enrico.zio@ecp.fr, enrico.zio@supelec.fr

[2]*Energy Department, Politecnico di Milano*
Via Ponzio, 34/3 – 20133 Milan, Italy
Fax: +39-02-2399-6309; Phone: +39-02-2399-6340
E-mail address: enrico.zio@polimi.it

Abstract

The computation of the reliability of a thermal-hydraulic (T-H) passive system of a nuclear power plant can be obtained by i) Monte Carlo (MC) sampling the uncertainties of the system model and parameters, ii) computing, for each sample, the system response by a mechanistic T-H code and iii) comparing the system response with pre-established safety thresholds, which define the success or failure of the safety function. The computational effort involved can be prohibitive because of the large number of (typically long) T-H code simulations that must be performed (one for each sample) for the statistical estimation of the probability of success or failure.

The objective of this work is to provide operative guidelines to effectively handle the computation of the reliability of a nuclear passive system. Two directions of computation efficiency are considered: from one side, efficient Monte Carlo Simulation (MCS) techniques are indicated as a means to performing robust estimations with a limited number of samples: in particular, the Subset Simulation (SS) and Line Sampling (LS) methods are identified as most valuable; from the other side, fast-running, surrogate regression models (also called response surfaces or meta-models) are indicated as a valid replacement of the long-running T-H model codes: in particular, the use of bootstrapped Artificial Neural Networks (ANNs) is shown to have interesting potentials, including for uncertainty propagation.

The recommendations drawn are supported by the results obtained in an illustrative application of literature.

Keywords: Nuclear passive system, functional failure analysis, computational cost, efficient Monte Carlo Simulation, fast-running regression model.

1 Introduction

Nuclear safety has expanded its considerations to severe accidents and increased its requirements for guaranteeing effective safety functions. This explains the interest in passive systems [Ahn et al., 2010; Kim et al., 2010], which all innovative reactor concepts make use of, to a large extent in combination with active systems [Mackay et al., 2008; Mathews et al., 2008 and 2009].

According to the International Atomic Energy Agency (IAEA) definitions, a passive component does not need external input (especially energy) to operate [IAEA, 1991]. Then, the term "passive" identifies a system which is composed entirely of passive components and structures, or a system, which uses active components in a very limited way to initiate subsequent passive operation. The currently accepted categorization of passive systems, developed by the IAEA, is summarized in Table 1 [IAEA, 1991].

Passive systems are expected to contribute significantly to nuclear safety by combining peculiar characteristics of simplicity, reduction of human interaction and reduction or avoidance of external electrical power and signals input [Nayak et al., 2008a and b; Nayak et al., 2009]. On the other hand, the assessment of the effectiveness of passive systems must include considerations on their reliability; these have to be drawn in the face of lack of data on some underlying phenomena, scarce or null operating experience of these systems over the wide range of conditions encountered during operation and less guaranteed performance as compared to active safety systems [Pagani et al., 2005; Burgazzi, 2007a].

Indeed, although passive systems are credited a higher reliability with respect to active ones, because of the reduced unavailability due to hardware failure and human error, the uncertainties involved in the actual *operation* of passive systems in the field and their *modeling* are usually larger than in active systems. Two different sources of uncertainties are usually considered in passive system analysis: randomness due to intrinsic variability in the behavior of the system (aleatory uncertainty) and imprecision due to lack of data on some underlying phenomena (e.g., natural circulation) and to scarce or null operating experience over the wide range of conditions encountered during operation [Apostolakis, 1990; Helton and Oberkampf, 2004].

As a consequence of these uncertainties, in practice there is a nonzero probability that the physical phenomena involved in the passive system operation lead to failure of performing the intended safety function even if i) safety margins are present and ii) no hardware failures occur. In fact, deviations in the natural forces and in the conditions of the underlying physical principles from the expected ones can impair the function of the system itself: this event is referred to in the literature

as *functional failure* [Burgazzi, 2003]. The quantification of the probability of this occurrence is an issue of concern both for the "nominal" passive systems (e.g., the ESBWR operating in nominal conditions) [Juhn et al., 2000; Rohde et al. 2008] and the "emergency" passive systems (e.g., accumulators, isolation condensers, etc.) [Chung et al., 2008]. In the following, the discussion will focus on the latter type of systems.

The occurrence of functional failures is especially critical in type B passive systems, i.e., those involving moving working fluids and referred to as Thermal-Hydraulic (T-H) passive systems (Table 1). The reason lies behind the small driving forces engaging passive operation and the complex and delicate T-H phenomena determining the system performance. For performing their accident prevention and/or mitigation functions, these passive systems rely exclusively on natural forces, e.g. gravity or natural convection, not generated by external power sources. Because the magnitude of the natural forces which drive operation is relatively small, counter-forces (e.g., friction) cannot be ignored because of comparable magnitude. This leads to uncertainty in the actual T-H system performance which must be evaluated by a specific, systematic and rigorous methodology[1].

In recent years, several methodologies have been proposed in the literature to quantify the probability that nuclear passive systems fail to perform their functions [Burgazzi, 2007b; Zio and Pedroni, 2009a]. A number of methods adopt the system reliability analysis framework. In [Aybar et al., 1999], a dynamic methodology based on the cell-to-cell mapping technique has been used for the reliability analysis of an inherently safe Boiling Water Reactor (BWR). In [Burgazzi, 2007a], the failure probability is evaluated as the probability of occurrence of different independent failure modes, a priori identified as leading to the violation of the boundary conditions and/or physical mechanisms needed for successful passive system operation. In [Burgazzi, 2002], modeling of the passive system is simplified in terms of the modeling of the unreliabilities of the hardware components of the system: this is done by identifying the hardware components failures that degrade the natural mechanisms which the passive system relies upon and associating the corresponding components unreliabilities. This concept is also at the basis of the Assessment of Passive System ReliAbility (APSRA) approach which has been applied to the reliability analysis of the natural circulation-based Main Heat Transport (MHT) system of an Indian Heavy Water Reactor (HWR) [Nayak et al., 2008a and b; Nayak et al., 2009].

[1] Notice that in the following, the discussion will focus on Type B passive systems, i.e., those involving moving working fluids and referred to as T-H passive systems; thus, the locution "passive system" will implicitly mean "T-H passive system" in the remainder of the paper.

An alternative approach is founded on the introduction of the concept of *functional failures*, within the reliability physics framework of load-capacity exceedance [Burgazzi, 2003, 2007a and c, 2008 and 2009]: a passive system fails to perform its function due to deviations from its expected behavior which lead the load imposed on the system to overcome its capacity. In [Woo and Lee, 2009a, b and 2010; Han and Yang, 2010], this concept is at the basis of the estimation of the functional failure probability of passive decay heat removal systems of Very High Temperature Reactors (VHTRs). It also provides the basis for the methodologies known as Reliability Evaluation of PAssive Safety (REPAS) systems [D' Auria et al., 2002; Jafari et al., 2003; Zio et al., 2003] and Reliability Methods for Passive Safety (RMPS) functions [Marquès et al., 2005], developed and employed for the analysis of passive Residual Heat Removal Systems (RHRSs) of Light Water Reactors (LWRs). It has also been used to evaluate the failure probabilities of decay heat removal systems in Gas-cooled Fast Reactors (GFRs) [Pagani et al., 2005; Bassi and Marquès, 2008; Mackay et al., 2008; Patalano et al., 2008; Zio and Pedroni, 2009b, c and 2010; Pedroni et al., 2010; Zio et al., 2010], sodium-cooled Fast Breeder Reactors (FBRs) [Mathews et al., 2008 and 2009; Arul et al., 2009 and 2010] and the lead-cooled, fast spectrum Flexible Conversion Ratio Reactor (FCRR) [Fong et al., 2009]. In all these analyses, the passive system is modeled by a detailed, mechanistic T-H system code and the probability of not performing the required function is estimated based on a Monte Carlo (MC) sample of code runs which propagate the *epistemic* (state-of-knowledge) uncertainties in the model describing the system and the numerical values of its parameters. Because of the existence of these uncertainties, it is possible that even if no hardware failure occurs, the system may not be able to accomplish its mission[2].

The functional failure-based approach provides in principle the most realistic assessment of the T-H passive system, thanks to the flexibility of Monte Carlo simulation which does not suffer from any T-H model complexity and, therefore, does not force to resort to simplifying approximations: for this reason, the functional failure-based approach will be taken here as reference. On the other hand, such approach requires considerable and often prohibitive computational efforts. The reason is twofold. First, a large number of Monte Carlo-sampled T-H model evaluations must generally be carried out for an accurate uncertainty propagation and functional failure probability estimation.

[2] It is worth mentioning also the work performed by Lee and co-workers who took up the problem of passive system functional reliability assessment focusing on the idea of *identifying* the limit state function of the system (essentially referring to the generic structural reliability paradigm of load-capacity exceedance described above) as a *prelude* to the quantification of the functional reliability itself [Aumeier, 1994; Aumeier and Lee, 1993 and 1994; Aumeier et al., 1995 and 2006; Lee et al., 1993-1995]. However, since the focus of the present paper is on the *efficient computation* of the passive system functional reliability (*given* the limit state function of the system and proper input probability distributions representing the uncertainties in the system model and parameters), no further details are given here for brevity; the interested reader is thus referred to the cited references.

Since the number of simulations required to obtain a given accuracy depends on the magnitude of the failure probability to be estimated, with the computational burden increasing with decreasing functional failure probability [Schueller, 2007 and 2009], this poses a significant challenge for the typically quite small (e.g., less than 10^{-4}) probabilities of functional failure of T-H passive safety systems. Second, long calculations (several hours) are typically necessary for each run of the detailed, mechanistic T-H code (one code run is required for each sample of values drawn from the uncertainty distributions) [Fong et al., 2009; Pourgol-Mohamad et al., 2010][3].

Finally, notice that for the same reasons a high computational burden is associated also to the sensitivity analysis process, i.e., the identification of the model parameters that contribute the most to the uncertainty in the performance of the passive system and consequently to its functional failure [Saltelli et al., 2008; Marrel et al., 2009].

Thus, efficient simulation techniques must be sought to perform robust functional failure probability estimation, uncertainty propagation and sensitivity analysis while reducing as much as possible the number of T-H code simulations and the associated computational time.

The objective of the present paper is to show how the computational issues associated to the functional reliability assessment of nuclear passive systems can be effectively handled. Two conceptual directions of computation efficiency are considered: efficient Monte Carlo Simulation techniques for performing robust estimations based on a limited number of samples drawn (i.e., T-H code simulations); fast-running, surrogate regression models (also called response surfaces or meta-models) in replacement of the long-running T-H model codes.

Within this conceptual framework, *different* computational methods are recommended for efficiently tackling the *different* phases of the functional reliability assessment of nuclear passive systems: in particular, an optimized Line Sampling (LS) method [Zio and Pedroni, 2010] is recommended for functional failure probability estimation, whereas the use of Subset Simulation (SS) [Au and Beck, 2001; Au and Beck, 2003b] and bootstrapped Artificial Neural Networks (ANNs) [Efron and Thibshirani, 1993; Zio, 2006] is suggested for uncertainty propagation and sensitivity analysis.

These recommendations are arrived at on the basis of i) a critical review of the methods available in the literature on the subject and ii) the experience of the authors in nuclear passive systems functional reliability assessments [Zio and Pedroni, 2009a-c and 2010; Pedroni et al., 2010; Zio et al., 2010].

[3] For example, the computer code RELAP5-3D, which is used to describe the thermal-hydraulic behavior of nuclear systems, may take up to twenty hours per run in some applications.

The remainder of the paper is organized as follows. In Section 2, the main sources and types of uncertainties involved in the operation and modeling of nuclear passive systems are recalled. In Section 3, the reliability analysis of nuclear passive systems is framed in terms of the concept of functional failure. In Section 4, the two conceptual directions considered for reducing the computational burden associated to the reliability assessment of nuclear passive systems (i.e., advanced MCS and empirical regression modeling) are presented and critically analyzed on the basis of a literature review. In Section 5, techniques are recommended to effectively tackle the computational burden associated to the different phases of the reliability assessment; results of the application of the proposed techniques to a case study of literature are also shown. Finally, guidelines and recommendations are summarized in the concluding Section.

2 Sources and types of uncertainties in the operation and modeling of nuclear passive systems

Uncertainties in the operation and modelling of nuclear passive systems must be accounted for in their reliability evaluations within a Probabilistic Risk Assessment (PRA) framework [Burgazzi, 2004; Pagani et al., 2005; Burgazzi, 2007a-c].

To effectively represent and model these uncertainties, it is useful to distinguish two kinds: “aleatory” and “epistemic” [Apostolakis, 1990; Helton and Oberkampf, 2004; USNRC, 2009]. The former refers to phenomena occurring in a random way: probabilistic modeling offers a sound and efficient way to describe such occurrences. The latter captures the analyst’s confidence in the PRA model by quantifying the degree of belief of the analysts on how well it represents the actual system; it is also referred to as *state-of-knowledge* or *subjective* uncertainty and can be reduced by gathering information and data to improve the knowledge on the system behavior.

As might be expected, the uncertainties affecting the operation of nuclear passive systems (Table 2) are both of aleatory kind, because of the randomness in the occurrence of some phenomena, and of epistemic nature, because of the limited knowledge on some phenomena and processes and the paucity of the relative operational and experimental data available [Burgazzi, 2007a].

Aleatory uncertainties concern, for instance, the occurrence of an accident scenario, the time to failure of a component or the variation of the actual geometrical dimensions (due to differences between the as-built system and its design upon which the analysis is based) and material properties (affecting the failure modes, e.g. concerning undetected leakages and heat losses) [NUREG-1150, 1990; Helton, 1998; USNRC, 2002; Burgazzi, 2007a-c]. Two examples of classical probabilistic

models used to describe this kind of uncertainties in PRAs are the Poisson model for events randomly occurring in time (e.g., random variations of the operating state of a valve) and the binomial model for events occurring "as the immediate consequence of a challenge" (e.g., failures on demand) [NUREG-CR-6850, 2005]. The effects of these uncertainties are then propagated onto the risk measure, e.g. by Monte Carlo simulation based on Importance Sampling or Stratified Sampling [Hofer et al., 2002; Cacuci and Ionescu-Bujor, 2004; Krzykacz-Hausmann, 2006]. The contribution of aleatory uncertainty to nuclear passive systems failure is quite clear: for example, natural circulation could be altered by a random disturbance in the system geometry or by a random variation of the operating state of a component [Pagani et al., 2005].

In the present paper, the representation and propagation of aleatory uncertainties are not considered, the focus being on epistemic uncertainty [Pagani et al., 2005; Bassi and Marques, 2008; Mackay et al., 2008; Mathews et al., 2008; Patalano et al., 2008; Arul et al., 2009 and 2010].

Epistemic uncertainty is associated to the lack of knowledge about the properties and conditions of the phenomena (i.e., natural circulation) underlying the behavior of the passive systems. This uncertainty manifests itself in the model representation of the system behavior, in terms of both (*model*) uncertainty in the hypotheses assumed and (*parameter*) uncertainty in the values of the parameters of the model [Cacuci and Ionescu-Bujor, 2004; Helton et al., 2006; Patalano et al., 2008].

Model uncertainty arises because mathematical models are simplified representations of real systems and, therefore, their results may be affected by error or bias. Model uncertainty also includes the fact that the model could be too simplified and therefore would neglect some important phenomena affecting the final result. This latter type of uncertainty is sometimes identified independently from model uncertainty and is known as *completeness* uncertainty [USNRC, 2009].

Model uncertainty may for example involve the correlations adopted to describe the T-H phenomena, which are subject to errors of approximation. Such uncertainties may for example be captured by a multiplicative model [Zio and Apostolakis, 1996; Patalano et al., 2008]:

$$z = c(\boldsymbol{x}) \cdot \varepsilon, \tag{1}$$

where z is the real value of the quantity to be predicted (e.g. heat transfer coefficients, friction factors, Nusselt numbers or thermal conductivity coefficients), $c(\cdot)$ is the mathematical model of the correlation (i.e., the result of the correlation as computed by the T-H code), $\boldsymbol{x}$ is the vector of correlating variables and ε is the associated multiplicative error factor: as a result, the uncertainty in the quantity z to be predicted is translated into an uncertainty in the multiplicative error factor ε. This error is commonly classified as representing *model* uncertainty.

Furthermore, uncertainty affects the values of the *parameters* used to describe the system (e.g., power level, pressure, cooler wall temperature, material conductivity, …), e.g. owing to errors in their measurement or insufficient data and information. For example, according to industry practice and experience, an error of 2% is usually considered in the determination of the power level in a reactor, due to uncertainties in the measurements. As a consequence, the power level is usually known only to a certain level of precision, i.e., epistemic uncertainty is associated with it.
Both model and parameter uncertainties associated to the current state of knowledge of the system can be represented by subjective probability distributions within a Bayesian approach to PRA [Apostolakis, 1990, 1995 and 1999]. In current PRAs, the effect of these uncertainties is often propagated on the risk measure by Latin Hypercube Sampling (LHS) [Helton and Davis, 2003].

Epistemic uncertainties affect also the identification of the *failure criterion* to be adopted for the system under analysis: for instance, reactor parameters (e.g., the maximal cladding temperature) as well as passive system variables (e.g., the thermal power exchanged in a cooler) could be equally adopted as indicators of the safety performance of the passive system; furthermore, the failure thresholds may be established as point-targets (e.g., a specific quantity of liquid must be delivered within a fixed time) or time-varying targets or even integral targets over a defined mission time (e.g., the system must reject at least a given value of thermal power during the entire system intervention) [Jafari et al., 2003; Marques et al., 2005].
Finally, state-of-knowledge uncertainty affects the identification of the possible *failure modes* and related *causes* and *consequences*, such as leaks (e.g., from pipes and pools), deposit thickness on components surfaces (e.g., pipes or heat exchangers), presence of non-condensable gases, stresses, blockages and material defects [Burgazzi, 2007a]. The identification of all the relevant modes/causes of failure in terms of *critical parameters* for the passive system performance/stability and the assessment of the relative uncertainty may be attempted by commonly used hazard identification procedures, like HAZard and OPerability (HAZOP) analysis and Failure Mode and Effect Analysis (FMEA) [Burgazzi, 2004 and 2006].
The contribution of epistemic uncertainties to the definition of the reliability/failure probability of nuclear passive systems can be qualitatively explained as follows. If the analyst is not fully confident on the validity of the correlations adopted to estimate, e.g., the design value of the heat transfer coefficient in the core during natural convection (e.g., due to the paucity of experimental data available in support of the use of a particular correlation), he/she admits that in a *real* accident scenario the *actual* value of the heat transfer coefficient in the core might deviate from the nominal/design one (i.e., different from the value computed by a deterministic correlation). If this

variation (accepted as plausible by the analyst) were to take place during an accident scenario, it may cause the passive system to fail performing its safety function; based on the current state of knowledge of the heat transfer phenomenon in the core under the expected conditions, the likelihood of the heat transfer coefficient variation is to be quantified for *estimating* the reliability/failure probability. A future improvement in the state of knowledge, e.g. due to the collection of data and information useful to improve the characterization of the heat transfer phenomenon, would lead to a change in the epistemic uncertainty distribution describing the likelihood of the various values of heat transfer coefficient and eventually to a more accurate *estimate* of the system reliability/failure probability [Pagani et al., 2005; Bassi and Marques, 2008; Mackay et al., 2008; Mathews et al., 2008 and 2009; Patalano et al., 2008; Arul et al., 2009 and 2010; Fong et al., 2009].

In the present paper, *only* epistemic uncertainties are considered in the estimation of the reliability/failure probability of nuclear passive systems [Pagani et al., 2005; Bassi and Marques, 2008; Mackay et al., 2008; Mathews et al., 2008; Patalano et al., 2008; Arul et al., 2009 and 2010].

3 Functional failure analysis of nuclear passive systems

The essential steps for the conceptual development of the functional failure analysis of nuclear passive systems are briefly reported below [Marquès et al., 2005]:

1. Detailed modeling of the system response by means of a deterministic, best-estimate (typically long-running) T-H code.
2. Identification of the vector $\boldsymbol{x} = \{x_1, x_2, \ldots, x_j, \ldots, x_{n_i}\}$ of parameters/variables, models and correlations (i.e., the inputs to the T-H code) which contribute to the uncertainty in the vector $\boldsymbol{y} = \{y_1, y_2, \ldots, y_l, \ldots, y_{n_o}\}$ of the outputs of the best-estimate T-H calculations (Section 2).
3. Propagation of the uncertainties associated to the identified relevant parameters, models and correlations $\boldsymbol{x}$ (step 2. above) through the deterministic, long-running T-H code in order to provide a complete representation (in terms of Probability Density Functions-PDFs, Cumulative Distribution Functions-CDFs and so on) of the *uncertainty* associated to the vector $\boldsymbol{y}$ of the outputs (step 2. above) of the deterministic, best-estimate T-H code.
4. Estimation of the *functional failure probability* of the passive system conditional on the current state of knowledge about the phenomena involved (step 2. above) [Pagani et al., 2005; Bassi and Marques, 2008; Mackay et al., 2008; Mathews et al., 2008 and 2009; Patalano et al., 2008; Arul et al., 2009 and 2010; Fong et al., 2009; Zio and Pedroni, 2009a-c and 2010; Pedroni et al., 2010; Zio et al., 2010]. Formally, let $Y(\boldsymbol{x})$ be a single-valued scalar

variable indicator of the performance of the passive system (e.g., the fuel peak cladding temperature) and α_Y a threshold value defining the corresponding failure criterion (e.g., a limit value imposed by regulating authorities)[4]. For illustrating purposes, let us assume that the passive system operates as long as $Y(\boldsymbol{x}) < \alpha_Y$; equivalently, introducing a variable called Performance Function (PF) as $g_x(\boldsymbol{x}) = Y(\boldsymbol{x}) - \alpha_Y$, failure occurs if $g_x(\boldsymbol{x}) > 0$. The probability $P(F)$ of system functional failure can then be expressed by the multidimensional integral:

$$P(F) = \iint \ldots \int I_F(\boldsymbol{x}) q(\boldsymbol{x}) d\boldsymbol{x} \tag{2}$$

where $q(\cdot)$ is the joint Probability Density Function (PDF) representing the uncertainty in the parameters $\boldsymbol{x}$, F is the failure region (where $g_x(\cdot) > 0$) and $I_F(\cdot)$ is an indicator function such that $I_F(\boldsymbol{x}) = 1$, if $\boldsymbol{x} \in F$ and $I_F(\boldsymbol{x}) = 0$, otherwise. The MCS procedure for estimating the functional failure probability entails that a *large* number N_T of samples of the values of the system parameters $\boldsymbol{x}$ be drawn from the corresponding probability distributions and used to evaluate $Y(\boldsymbol{x})$ by running the T-H code. An estimate $\hat{P}(F)^{N_T}$ of the probability of failure $P(F)$ can then be computed by dividing the number of times that $Y(\boldsymbol{x}) > \alpha_Y$ by the total number of samples N_T.

5. Perform a sensitivity study to determine the contribution of the individual uncertain parameters (i.e., the inputs to the T-H code) $\{x_j: j = 1, 2, \ldots, n_i\}$ to the uncertainty in the outputs of the T-H code $\{y_l: l = 1, 2, \ldots, n_o\}$ (and in the performance function $Y(\boldsymbol{x})$ of the passive system) and consequently to the functional failure probability of the T-H passive system. As is true for uncertainty propagation (step 4. above), sensitivity analysis relies on *multiple* (e.g., many thousands) evaluations of the code for different combinations of system inputs.

In this work, we propose to tackle the computational burden posed by the uncertainty propagation, failure probability estimation and sensitivity analysis of steps 3. – 5. above in two effective ways (Section 4): from one side, efficient Monte Carlo Simulation techniques can be employed to perform robust estimations with a limited number of input samples (Section 4.1); from the other side, fast-running, surrogate regression models (also called response surfaces or meta-models) can be used to replace the long-running T-H model code (Section 4.2).

---

[4] Note that the choice of a *single-valued* performance function does not reduce the generality of the approach, because any multidimensional vector of physical quantities (i.e., the vector $\boldsymbol{y}$ of the outputs of the T-H code in this case) can be conveniently re-expressed as a scalar parameter by resorting to suitable min-max transformations: see [Au and Beck, 2001 and 2003b; Zio and Pedroni, 2009b, c and 2010] for details.

4 Handling the computational issues associated to the functional reliability assessment of nuclear passive systems

In this Section, the two approaches considered for dealing with the computational issue associated to the functional reliability assessment of nuclear passive systems are summarized: in Section 4.1, various Monte Carlo Simulation techniques are synthetically described; in Section 4.2, empirical regression modeling is presented as a means to build fast-running, surrogate models for replacing the long-running T-H model codes. Both approaches are critically reviewed on the basis of the available literature.

4.1 Advanced Monte Carlo Simulation methods

As previously stated, the computational issues described in the previous Section 3 can be tackled from one side by resorting to efficient simulation techniques that perform robust estimations with a limited number of input samples, thus with an associated low computational time.

One such technique is the Importance Sampling (IS) method [Kalos and Whitlock, 1986; Au and Beck, 2003a; Au, 2004; Schueller et al., 2004]. This technique amounts to replacing the original Probability Density Function (PDF) of the uncertain variables with an Importance Sampling Distribution (ISD) chosen so as to generate samples that lead to failure more frequently [Au and Beck, 2003a]. IS has the capability of considerably reducing the variance of the estimates compared with standard Monte Carlo Simulation (MCS), provided that the ISD is chosen similar to the theoretical optimal one. In practice, substantial insights on the system behaviour and extensive modelling work may be required to identify a “good” ISD, e.g. by setting up complex kernel density estimators [Au and Beck, 2003a], by identifying the design point of the problem [Au, 2004] or simply by tuning the parameters of the ISD based on expert judgment and trial-and-error [Pagani et al., 2005]. Overall, this increases the effort associated to the simulation; furthermore, there is always the risk that an inappropriate choice of the ISD may lead to worse estimates compared to Standard MCS [Schueller et al., 2004].

Another technique is Stratified Sampling. This technique requires dividing the sample space into several non-overlapping subregions (referred to as “strata”) and calculating the probability of each subregion; the (stratified) sample is then obtained by randomly sampling a predefined number of outcomes from *each* stratum [Helton and Davis, 2003; Cacuci and Ionescu-Bujor, 2004]. By so doing, the full coverage of the sample space is ensured while maintaining the probabilistic character of random sampling. A major issue related to the implementation of Stratified Sampling lies in defining the strata and calculating the associated probabilities, which may require considerable a priori knowledge. As a remark, notice that the widely used event tree techniques in nuclear reactor

Probabilistic Risk Assessment (PRA) can be seen as defining and implementing Stratified Sampling of accident events and scenarios [Cacuci and Ionescu-Bujor, 2004].

A popular compromise between plain random sampling (i.e., standard MCS) and Importance/Stratified Sampling is offered by Latin Hypercube Sampling (LHS), which is commonly used in PRA [Morris, 2000] for efficiently generating random samples [MacKay et al., 1979; Helton and Davis, 2003; Helton et al., 2005; Sallaberry et al., 2008]. The effectiveness of LHS, and hence its popularity, derives from the fact that it provides a dense *stratification* over the range of each uncertain variable, with a relatively small sample size, while preserving the desirable probabilistic features of simple random sampling; moreover, there is no necessity to determine strata and strata probabilities like in Stratified Sampling [Helton and Davis, 2003]. For these reasons LHS is frequently adopted for efficiently propagating *epistemic* uncertainties in PRA problems [NUREG-1150, 1990; Helton, 1998; Hofer et al., 2002; Krzykacz-Hausmann, 2006; Helton and Sallaberry, 2009].
On the other hand, LHS is very efficient for estimating mean values and standard deviations in complex reliability problems [Olsson et al., 2003], but only slightly more efficient than standard MCS for estimating small failure probabilities [Pebesma and Heuvelink, 1999], like those expected for passive safety systems.

Recently, Subset Simulation (SS) [Au and Beck, 2001; Au and Beck, 2003b] and Line Sampling (LS) [Koutsourelakis et al., 2004; Pradlwarter et al., 2005] have been proposed as advanced Monte Carlo Simulation methods for efficiently tackling the multidimensional problems of structural reliability. These methods have proved efficient also in the estimation of the functional failure probability of T-H passive systems [Zio and Pedroni, 2009b, c and 2010]. Indeed, structural reliability problems are also formulated within a functional failure framework of analysis, in which the systems fail whenever the load applied (i.e., the stress) exceeds their capacity (i.e., the resistance) [Schueller and Pradlwarter, 2007]. This makes the two methods suitable for application to the functional reliability analysis of nuclear passive systems, where the failure is specified in terms of one or more safety variables (e.g., temperatures, pressures, flow rates, ...) crossing the safety thresholds specified by the regulating authorities [Bassi and Marques, 2008; Mackay et al., 2008; Mathews et al., 2008; Patalano et al., 2008].
More specifically, in the SS approach, the functional failure probability is expressed as a product of conditional probabilities of some chosen intermediate and thus more frequent events. The problem of evaluating the small probabilities of functional failures is thus tackled by performing a sequence

of simulations of more frequent events in their conditional probability spaces; the necessary conditional samples are generated through successive Markov Chain Monte Carlo (MCMC) simulations [Metropolis et al., 1953], in a way to gradually populate the intermediate conditional regions until the final functional failure region is reached.

In the LS method, *lines*, instead of random *points*, are used to probe the failure domain of the high-dimensional problem under analysis [Pradlwarter et al., 2005]. An "important direction" is optimally determined to point towards the failure domain of interest and a number of conditional, one-dimensional problems are solved along such direction, in place of the high-dimensional problem [Pradlwarter et al., 2005]. The approach has been shown to perform better than standard MCS in a wide range of reliability applications [Koutsourelakis et al., 2004; Schueller et al., 2004; Pradlwarter et al., 2005 and 2007; Schueller and Pradlwarter, 2007; Lu et al., 2008; Valdebenito et al., 2010; Zio and Pedroni, 2009c and 2010]. Furthermore, if the boundaries of the failure domain of interest are not too rough (i.e., almost linear) and the "important direction" is almost perpendicular to them, the variance of the failure probability estimator could be ideally reduced to zero [Koutsourelakis et al., 2004][5].

In the present paper, particular focus is devoted to SS and LS: for this reason, synthetic descriptions of these techniques and an illustrative application to the functional failure analysis of a T-H passive system are reported in Section 5.

4.2 Empirical regression modeling

Another way to tackle the computational issues associated to the reliability analysis of nuclear passive systems is that of resorting to fast-running, surrogate regression models, also called response surfaces or meta-models, to approximate the input/output function implemented in the long-running system model code, and then substitute it in the passive system reliability analysis [Storlie et al., 2008].

The construction of such regression models entails running the system model code a predetermined, *reduced* number of times (e.g., 50-100) for specified values of the uncertain input variables and collecting the corresponding values of the output of interest; then, statistical techniques are employed for calibrating/adapting the internal *parameters/coefficients* of the response surface of the regression model in order to fit the input/output data generated in the previous step.

---

[5] Apart from efficient MC techniques, there exist methods based on nonparametric order statistics [Wilks, 1942] that propagate uncertainties through mechanistic T-H codes with reduced computational burden, especially if only one- or two-sided confidence intervals are needed for particular statistics (e.g., the 95th percentile) of the outputs of the code. For example, the so-called *coverage* [Guba et al., 2003; Makai and Pal, 2006] and *bracketing* [Nutt and Wallis, 2004] approaches can be used to identify the number of sample code runs required to obtain a given *confidence* level on the estimates of prescribed statistics of the code outputs.

Putting it in a formal framework, let us consider a generic meta-model to be built for performing the task of nonlinear regression, i.e., estimating the nonlinear relationship between a vector of input variables $\boldsymbol{x} = \{x_1, x_2, ..., x_j, ..., x_{n_i}\}$ and a vector of output targets $\boldsymbol{y} = \{y_1, y_2, ..., y_l, ..., y_{n_o}\}$, on the basis of a *finite* (and possibly *small*) set of input/output data examples (i.e., patterns), $D_{train} = \{(\boldsymbol{x}_p, \boldsymbol{y}_p), p = 1, 2, ..., N_{train}\}$ [Zio, 2006]. It can be assumed that the target vector $\boldsymbol{y}$ is related to the input vector $\boldsymbol{x}$ by an unknown nonlinear deterministic function $\boldsymbol{\mu}_y(\boldsymbol{x})$ corrupted by a noise vector $\boldsymbol{\varepsilon}(\boldsymbol{x})$, i.e.,

$$\boldsymbol{y}(\boldsymbol{x}) = \boldsymbol{\mu}_y(\boldsymbol{x}) + \boldsymbol{\varepsilon}(\boldsymbol{x}). \tag{3}$$

As introduced in Section 3, in the present case of T-H passive system functional failure probability assessment the vector $\boldsymbol{x}$ contains the relevant uncertain system parameters/variables, the nonlinear deterministic function $\boldsymbol{\mu}_y(\boldsymbol{x})$ represents the complex, long-running T-H mechanistic model code (e.g., RELAP5-3D), the vector $\boldsymbol{y}(\boldsymbol{x})$ contains the output variables of interest for the analysis and the noise $\boldsymbol{\varepsilon}(\boldsymbol{x})$ represents the errors introduced by the numerical methods employed to calculate $\boldsymbol{\mu}_y(\boldsymbol{x})$ [Storlie et al., 2009]; for simplicity, in the following we assume $\boldsymbol{\varepsilon}(\boldsymbol{x}) = \boldsymbol{0}$ [Secchi et al., 2008]. Thus, the objective of the regression task is to estimate $\boldsymbol{\mu}_y(\boldsymbol{x})$ in (3) by means of a regression function $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ depending on a set of parameters $\boldsymbol{w}^*$ to be properly determined on the basis of the available data set D_{train}. The algorithm used to calibrate the set of parameters $\boldsymbol{w}^*$ is obviously dependent on the nature of the regression model adopted, but in general it aims at minimizing the mean (absolute or quadratic) error between the output targets of the original T-H code, $\boldsymbol{y}_p = \boldsymbol{\mu}_y(\boldsymbol{x}_p)$, $p = 1, 2, ..., N_{train}$, and the output vectors of the regression model, $\hat{\boldsymbol{y}}_p = \boldsymbol{f}(\boldsymbol{x}_p, \boldsymbol{w}^*)$, $p = 1, 2, ..., N_{train}$; for example, the Root Mean Squared Error (RMSE) is commonly adopted to this purpose [Zio, 2006].

Several examples can be found in the open literature concerning the application of surrogate meta-models in reliability problems. In [Bucher and Most, 2008; Gavin and Yau, 2008; Liel et al., 2009], polynomial Response Surfaces (RSs) are employed to evaluate the failure probability of structural systems; in [Arul et al., 2009 and 2010; Fong et al., 2009; Mathews et al., 2009], *linear* and *quadratic* polynomial RSs are employed for performing the reliability analysis of T-H passive systems in advanced nuclear reactors; in [Deng, 2006; Hurtado, 2007; Cardoso et al., 2008; Cheng et al., 2008], learning statistical models such as Artificial Neural Networks (ANNs), Radial Basis Functions (RBFs) and Support Vector Machines (SVMs) are trained to provide *local* approximations of the failure domain in structural reliability problems; in [Volkova et al., 2008;

Marrel et al., 2009], Gaussian meta-models are built to calculate global sensitivity indices for a complex hydrogeological model simulating radionuclide transport in groundwater.

However, when using the approximation of the system output provided by an empirical regression model, an additional source of *model* uncertainty is introduced which needs to be evaluated, particularly in safety critical applications like those related to nuclear power plant technology. In this paper we propose to resort to *bootstrapped* regression models [Efron and Thibshirani, 1993], i.e., ensembles of regression models, constructed on different data sets bootstrapped from the original one [Zio, 2006; Storlie et al., 2009]. In fact, the ensemble framework of regression modeling allows quantifying the model uncertainty associated to the estimates provided by the regression models in terms of *confidence intervals*.

The bootstrap method is a distribution-free inference method which requires no prior knowledge about the distribution function of the underlying population [Efron and Thibshirani, 1993]. The basic idea is to generate samples from the observed data by sampling with replacement from the original data set [Efron and Thibshirani, 1993]: each of these bootstrapped data sets is used to build a bootstrapped regression model which is used to calculate the reliability quantity of interest (e.g., the passive system failure probability in this case). From the theory and practice of ensembles of empirical models, it can be shown that the estimates given by bootstrapped regression models is in general more accurate than the estimate of the best regression model in the bootstrap ensemble of regression models [Zio, 2006; Cadini et al., 2008].

Some examples of the application of the bootstrap method for the evaluation of the uncertainties associated to the output of regression models in safety-related problems can be found in the literature: in [Zio, 2006], bootstrapped ANNs are trained to predict nuclear transients processes; in [Cadini et al., 2008; Secchi et al., 2008], the model uncertainty, quantified in terms of a *standard deviation*, is used to “correct” the ANN output in order to provide conservative estimates for important safety parameters in nuclear reactors (i.e., percentiles of the pellet cladding temperature); finally, in [Storlie et al., 2009], the bootstrap procedure is combined with different regression techniques, e.g. Multivariate Adaptive Regression Spline (MARS), Random Forest (RF) and Gradient Boosting Regression (GBR), to calculate confidence intervals for global sensitivity indices of the computationally demanding model of a nuclear waste repository.

In the present paper, particular emphasis is given to bootstrapped ANN regression models: for this reason, a synthetic description of this technique and an illustrative application to the functional failure analysis of a T-H passive system is reported in Section 5.

5 Recommendations for reducing the computational burden associated to the functional reliability analysis of nuclear passive systems

In this Section i) the different phases of the functional reliability analysis of nuclear passive systems are considered: in particular, the estimation of the functional failure probability (Section 5.1), the uncertainty propagation (Section 5.2) and sensitivity analysis (Section 5.3) phases; ii) on the basis of the literature review and the considerations made in the previous Section 4, techniques are recommended to efficiently tackle the computational burden associated to *each* of these analyses; iii) guidelines on the recommended techniques are provided, with illustrative applications to the functional reliability analysis of a nuclear passive system of literature [Pagani et al., 2005].

5.1 Functional failure probability estimation

If the analyst is only interested in an accurate and precise estimation of the (typically small) functional failure probability of the T-H passive system (modelled by a *long-running*, *nonlinear* and *non-monotonous* T-H code), then the use of the Line Sampling technique is strongly suggested.

In extreme synthesis, the computational steps of the algorithm are [Pradlwarter et al., 2005 and 2007]:

1. From the original multidimensional joint probability density function $q(\cdot):\Re^n \rightarrow [0,\infty)$, sample N_T vectors $\{\boldsymbol{x}^k : k=1,2,...,N_T\}$, with $\boldsymbol{x}^k = \{x_1^k, x_2^k, ..., x_j^k, ..., x_n^k\}$.
2. Transform the N_T sample vectors $\{\boldsymbol{x}^k : k=1,2,...,N_T\}$ defined in the original (i.e., physical) space into N_T samples $\{\boldsymbol{\theta}^k : k=1,2,...,N_T\}$ defined in the so-called "standard normal space", where each random variable is represented by an independent central unit Gaussian distribution; also the PFs $g_x(\cdot)$ defined in the physical space have to be transformed into $g_\theta(\cdot)$ in the standard normal space [Huang and Du, 2006].
3. In the standard normal space, determine a *unit* vector $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, ..., \alpha_j, ..., \alpha_n\}^T$ (hereafter also called "important unit vector" or "important direction") pointing towards the failure domain F of interest.
4. Reduce the problem of computing the high-dimensional failure probability integral (2) to a number of conditional one-dimensional problems, solved along the "important direction" $\boldsymbol{\alpha}$ in the standard normal space: in particular, estimate N_T conditional "one-dimensional" failure probabilities $\{\hat{P}(F)^{1\boldsymbol{D},k} : k=1,2,...,N_T\}$, corresponding to each one of the standard

normal samples $\left\{\boldsymbol{\theta}^k : k = 1, 2, ..., N_T\right\}$ obtained in step 2. Notice that $2 \cdot N_T$ or $3 \cdot N_T$ system performance analyses (i.e., runs of the T-H model code) have to be carried out to calculate *each* of the N_T conditional one-dimensional failure probability estimates $\left\{\hat{P}(F)^{1\boldsymbol{D},k} : k = 1, 2, ..., N_T\right\}$ (see Pradlwarter et al., 2005 and 2007 for details).

5. Compute the unbiased estimator $\hat{P}(F)^{N_T}$ for the failure probability $P(F)$ and its variance $\sigma^2\left[\hat{P}(F)^{N_T}\right]$ as:

$$\hat{P}(F)^{N_T} = 1/N_T \cdot \sum_{k=1}^{N_T} \hat{P}(F)^{1\boldsymbol{D},k} , \tag{4}$$

$$\sigma^2\left[\hat{P}(F)^{N_T}\right] = 1/N_T(N_T - 1) \cdot \sum_{k=1}^{N_T} \left(\hat{P}(F)^{1D,k} - \hat{P}(F)^{N_T}\right)^2 . \tag{5}$$

The LS method here outlined can significantly reduce the variance (5) of the estimator (4) of the failure probability integral (2) [Koutsourelakis et al., 2004]; however, its efficiency depends on the determination of the important direction $\boldsymbol{\alpha}$ (step 3. above).

With respect to this issue, four methods have been proposed in the open literature to estimate the important direction $\boldsymbol{\alpha}$ for LS. In [Koutsourelakis et al., 2004], the important unit vector $\boldsymbol{\alpha}$ is computed as the normalized "center of mass" of the failure domain F of interest; in [Koutsourelakis et al., 2004; Valdebenito et al., 2010], the important unit vector $\boldsymbol{\alpha}$ is taken as pointing in the direction of the "design point" in the standard normal space; in [Pradlwarter et al. 2005], the direction of $\boldsymbol{\alpha}$ is identified as the normalized gradient of the performance function $g_\theta(\cdot)$ in the standard normal space; finally, in a previous paper by the authors [Zio and Pedroni, 2010], the important direction $\boldsymbol{\alpha}$ is taken as the one minimizing the variance (5) of the failure probability estimator (4). This latter method produces more accurate and precise failure probability estimates than those provided by the other three techniques of literature and, for this reason, its adoption is recommended for the estimation of the small failure probabilities of T-H passive systems.

In more details, in [Zio and Pedroni, 2010] the *optimal* important direction $\boldsymbol{\alpha}^{opt}$ for Line Sampling is defined as the one minimizing the variance $\sigma^2\left[\hat{P}(F)^{N_T}\right]$ (5) of the LS failure probability estimator $\hat{P}(F)^{N_T}$ (4). Notice that $\boldsymbol{\alpha}^{opt}$ can be expressed as the normalized version of a proper vector $\boldsymbol{\theta}^{opt}$ in the standard normal space, i.e., $\boldsymbol{\alpha}^{opt} = \boldsymbol{\theta}^{opt} / \left\|\boldsymbol{\theta}^{opt}\right\|_2$. Thus, in order to search for a physically meaningful important unit vector $\boldsymbol{\alpha}^{opt}$ (i.e., a vector that optimally points towards the failure

domain F of interest), $\boldsymbol{\theta}^{opt}$ should belong to the failure domain F of interest, i.e. $\boldsymbol{\theta}^{opt} \in F$ or, equivalently, $g_\theta\left(\boldsymbol{\theta}^{opt}\right) > 0$.

In mathematical terms, the optimal LS important direction $\boldsymbol{\alpha}^{opt}$ is obtained by solving the following nonlinear constrained minimization problem:

$$\begin{aligned} &\text{Find } \boldsymbol{\alpha}^{opt} = \boldsymbol{\theta}^{opt} / \left\|\boldsymbol{\theta}^{opt}\right\|_2 : \ \sigma^2\left[\hat{P}(F)^{N_T}\right] = \min_{\boldsymbol{\alpha}=\boldsymbol{\theta}/\|\boldsymbol{\theta}\|_2} \left\{\sigma^2\left[\hat{P}(F)^{N_T}\right]\right\} \\ &\text{subject to } \boldsymbol{\theta} \in F \left(\text{i.e., } g_\theta(\boldsymbol{\theta}) > 0\right). \end{aligned} \tag{6}$$

The conceptual steps of the procedure for solving (6) are [Zio and Pedroni, 2010]:

1. An optimization algorithm proposes a candidate solution $\boldsymbol{\alpha} = \boldsymbol{\theta}/\|\boldsymbol{\theta}\|_2$ to (6): for example, probabilistic search algorithms like Genetic Algorithms (GAs) [Konak et al., 2006; Marseguerra et al., 2006] are particularly suitable for multivariate nonlinear problems like those involving nuclear passive safety systems [Zio and Pedroni, 2010].
2. The LS failure probability estimator $\hat{P}(F)^{N_T}$ (4) and the associated variance $\sigma^2\left[\hat{P}(F)^{N_T}\right]$ (5) are calculated using the unit vector $\boldsymbol{\alpha} = \boldsymbol{\theta}/\|\boldsymbol{\theta}\|_2$ proposed as important direction in step 1. above.
3. The variance $\sigma^2\left[\hat{P}(F)^{N_T}\right]$ obtained in step 2. above is the objective function to be *minimized*.
4. The feasibility of the proposed solution $\boldsymbol{\alpha} = \boldsymbol{\theta}/\|\boldsymbol{\theta}\|_2$ is checked by evaluating the system PF $g_\theta(\cdot)$ (i.e., by running the system model code) in correspondence of $\boldsymbol{\theta}$: if the proposed solution $\boldsymbol{\alpha} = \boldsymbol{\theta}/\|\boldsymbol{\theta}\|_2$ is not feasible (i.e., if $\boldsymbol{\theta} \notin F$ or, equivalently, $g_\theta(\boldsymbol{\theta}) \le 0$), it is *penalized* by *increasing* the value of the corresponding objective function $\sigma^2\left[\hat{P}(F)^{N_T}\right]$.
5. Steps 1. − 4. are repeated until a predefined stopping criterion is met and the optimization algorithm identifies the *optimal* unit vector $\boldsymbol{\alpha}^{opt} = \boldsymbol{\theta}^{opt} / \left\|\boldsymbol{\theta}^{opt}\right\|_2$.

Notice that i) the optimization search requires the iterative evaluation of hundreds or thousands of possible solutions $\boldsymbol{\alpha} = \boldsymbol{\theta}/\|\boldsymbol{\theta}\|_2$ to (6) and ii) $2 \cdot N_T$ or $3 \cdot N_T$ system performance analyses (i.e., runs of the system model code) have to be carried out to calculate the objective function $\sigma^2\left[\hat{P}(F)^{N_T}\right]$ for *each* proposed solution (step 2. and 3. above); as a consequence, the computational effort associated to this technique would be absolutely prohibitive with a system model code requiring hours or even minutes to run a single simulation. Hence, for practical applicability, one has to resort to a

regression model as a fast-running approximator of the original system model code for performing the calculations in steps 2. and 4. above, to make the computational cost acceptable.

The regression model suggested is the classical three-layered feed-forward ANN [Bishop, 1995]. In order to improve the *accuracy* in the approximation of the system PF $g_\theta(\cdot)$ (needed for an accurate estimation of the LS important direction $\boldsymbol{\alpha}$), the employed ANN models can be trained by a properly devised *sequential*, *two-step* algorithm based on error back-propagation, as proposed in [Zio and Pedroni, 2010]. In extreme synthesis, a *first-step* ANN regression model is built using a set of input/output data examples. The resulting ANN model is used (instead of the original, long-running system model code) to provide an *approximation* to the *design point* of the problem: this is meant to provide an approximate, rough indication of the real location of the failure domain F of interest. Subsequently, a new data set is randomly generated *centred* on the approximate design point previously identified: a *second-step* ANN model is then constructed on these newly generated data set. This should result in an ANN regression model which is more accurate in proximity of the failure domain F of interest, thus providing reliable estimates of the system PF $g_\theta(\cdot)$ for the identification of the LS important direction $\boldsymbol{\alpha}$ [Zio and Pedroni, 2010].

For completeness, we report some of the results obtained in a previous work by the authors [Zio and Pedroni, 2010], in which the optimized LS method described above is applied for the estimation of the *small* functional failure probability $P(F)$ of the passive decay heat removal system of a Gas-cooled Fast Reactor (GFR) of literature [Pagani et al., 2005] (notice that in this example $P(F) = 3.541 \cdot 10^{-4}$). A detailed description of the system is not reported here for brevity: the interested reader is referred to [Pagani et al., 2005] for details.

Further, the benefits coming from the use of the proposed method is shown by means of a comparison between the estimation *accuracies* and *precisions* of the following simulation methods: i) standard Monte Carlo Simulation (MCS); ii) Latin Hypercube Sampling (LHS) [Helton and Davis, 2003]; iii) standard Importance Sampling (IS) [Au and Beck, 2003a; Au, 2004]; iv) a combination of standard Importance Sampling (IS) and Latin Hypercube Sampling (LHS) (hereafter referred to as IS + LHS) [Olsson et al., 2003]; v) Subset Simulation (SS) [Au and Beck, 2001 and 2003b]; vi) optimized Line Sampling (LS) [Zio and Pedroni, 2010]; vii) a combination of optimized Line Sampling (LS) and Latin Hypercube Sampling (LHS) (hereafter referred to as LS + LHS) [Zio and Pedroni, 2010]. Part of the results used in the comparison are derived from the manipulation of results previously obtained by the authors [Zio and Pedroni, 2009b, c and 2010].

In order to properly represent the randomness of the probabilistic simulation methods i)-vii) adopted and provide a statistically meaningful comparison between their performances in the

estimation of the system failure probability $P(F)$, S = 2000 independent runs of each method have been carried out. In each simulation s = 1, 2, …, S, the percentage relative absolute error ε_s between the true (reference) value of the system failure probability $P(F)$ and the corresponding estimate $\hat{P}(F)_s^{N_T}$ obtained with N_T samples is computed as follows:

$$\varepsilon_s = \frac{\left| P(F) - \hat{P}(F)_s^{N_T} \right|}{P(F)} \cdot 100 \text{ , } s = 1, 2, \ldots, S \tag{7}$$

The accuracies of the simulation methods of interest in the estimation of $P(F)$ are then compared in terms of the mean percentage relative absolute error $\bar{\varepsilon}$ over S = 2000 runs:

$$\bar{\varepsilon} = \frac{1}{S} \cdot \sum_{s=1}^{S} \varepsilon_s \tag{8}$$

The quantity (8) provides a measure of the percentage relative absolute error in the estimation of the failure probability $P(F)$ made *on average in a single run* by the simulation method with N_T samples; obviously, the lower $\bar{\varepsilon}$, the higher the *accuracy* of the method.

The failure probability estimates $\hat{P}(F)_s^{N_T}$, s = 1, 2, …, S, are then used to build a *bootstrapped* 95% Confidence Interval (CI) for the failure probability estimator $\hat{P}(F)^{N_T}$, i.e.,

$$\left[L_{CI,\hat{P}(F)^{N_T}}, U_{CI,\hat{P}(F)^{N_T}} \right] \tag{9}$$

where $U_{CI,\hat{P}(F)^{N_T}}$ and $L_{CI,\hat{P}(F)^{N_T}}$ are the 2.5th and 97.5th percentiles, respectively, of the *bootstrapped empirical distribution* of the failure probability estimator $\hat{P}(F)^{N_T}$. The percentage relative *width* $\overline{w}_{CI}$ of the bootstrapped 95% Confidence Interval (CI) of the LS failure probability estimator $\hat{P}(F)^{N_T}$ is then computed as

$$\overline{w}_{CI} = \frac{U_{CI,\hat{P}(F)^{N_T}} - L_{CI,\hat{P}(F)^{N_T}}}{P(F)} \cdot 100 \tag{10}$$

Obviously, the lower $\overline{w}_{CI}$, the higher the *precision* of the method.

Finally, in addition to the accuracy and precision of the failure probability estimator, also the computational time associated to the simulation method has to be taken into account. To this aim, the FOM can be used:

$$\text{FOM} = \frac{1}{\sigma^2\left(\hat{P}(F)^{N_T}\right) \cdot t_{comp}} \approx \frac{1}{\hat{\sigma}^2\left(\hat{P}(F)^{N_T}\right) \cdot t_{comp}} \tag{11}$$

where t_{comp} is the computational time required by the simulation method and $\sigma^2\left[\hat{P}(F)^{N_T}\right]$ is defined in (5). Since $\sigma^2\left(\hat{P}(F)^{N_T}\right) \propto N_T$ and approximately $t_{comp} \propto N_T$, the FOM is independent of N_T. Obviously, the higher the FOM, the higher the computational efficiency of the method.

Table 3 reports the values of the performance indicators $\bar{\varepsilon}$ (8), $\overline{w}_{CI}$ (10) and FOM (11) obtained with N_T = 1850 samples by the simulation methods i)-vii) (notice that since N_T is the same for all the simulation methods, performance indicators $\bar{\varepsilon}$ (8) and $\overline{w}_{CI}$ (10) can be compared fairly). The number of T-H code runs required by each method is also reported: actually, when a single run of the system model code lasts several hours (which is often the case for passive safety systems) the total number of simulations is the *critical* parameter which determines the overall computational cost (i.e., t_{comp}) associated to the method. In particular, $N_{c,P(F)}$ is the number of code runs used by the algorithm *only* to *estimate* the failure probability $P(F)$; instead, $N_{c,add}$ is the number of *additional* code runs required to *set up* the method: for example, for IS and IS + LHS, $N_{c,add}$ code runs are used to build the Importance Sampling Density (ISD) by identification of the "design point" of the problem [Au, 2004]; instead, for LS and LS + LHS, $N_{c,add}$ code runs are used to identify the important direction $\boldsymbol{\alpha}$ by minimization of the variance of the LS failure probability estimator [Zio and Pedroni, 2010].

It can be seen that the optimized Line Sampling methods (i.e., both LS and LS + LHS) provide more accurate and precise failure probability estimates than the other methods: actually, the mean percentage errors $\bar{\varepsilon}$ are about 13 to 380 times lower than those of the other methods, whereas the percentage 95% CI widths $\overline{w}_{CI}$ are about 16 to 278 times lower than those of the other methods. Finally, although the computational cost associated to the optimized Line Sampling methods is higher than that of the other methods (because the total number of T-H code runs is more than three times larger), the overall computational efficiency of the method is significantly higher: actually, the FOM is about 2 to 4 *orders of magnitude* larger than that of the other methods.

The previous example has served to demonstrate that the optimized LS methods indeed provide more accurate and precise failure probability estimates than the other simulation methods considered. However, this must be achieved with a *small* number of samples (and, thus, of T-H model evaluations: say, few tens or hundreds depending on the application), because in practice the T-H computer codes require several hours to run a single simulation [Fong et al., 2009]. Thus, we consider here a practical situation where the number $N_{c,P(F)}$ of T-H code runs allowed for estimating

the *small* failure probability $P(F) = 3.541 \cdot 10^{-4}$ is set to few tens (e.g., 30 in this case). The results are summarized in Table 4.

It can be seen that even in this case the optimized Line Sampling methods (i.e., both LS and LS + LHS) provide more accurate and precise failure probability estimates than the other methods: actually, the mean percentage errors $\bar{\varepsilon}$ are about 6 to 44 times lower than those of the other methods, whereas the percentage 95% CI widths $\overline{w}_{CI}$ are about 6 to 163 times lower than those of the other methods. Finally, the global efficiency of the method is significantly higher: actually, the FOM is about 1 to 3 *orders of magnitude* larger than that of the other methods.
These results confirm the recommendation of adopting this method.

5.2 Uncertainty analysis

The objective of the uncertainty analysis is to propagate the uncertainty associated to the input parameters $\boldsymbol{x} = \{x_1, x_2, \ldots, x_j, \ldots, x_{ni}\}$ through the deterministic, long-running T-H code in order to quantify the uncertainty associated to the output variables $\boldsymbol{y} = \{y_1, y_2, \ldots, y_l, \ldots, x_{no}\}$ of interest and to the performance function $Y(\boldsymbol{x})$ of the passive system (e.g., computing Probability Density Functions-PDFs, Cumulative Distribution Functions-CDFs and percentiles).
In all fairness, notice that the strongly recommended LS technique allows *only* the (efficient) calculation of the failure probability of the passive system, but it does not allow a *complete* uncertainty propagation: actually, no Probability Density Functions (PDFs), Cumulative Distribution Functions (CDFs) or percentiles of the T-H code outputs of interest can be identified in a single simulation run. Thus, if the analyst is interested in propagating the uncertainty onto the output, two options are recommended:

1. in the (unlikely) case that the T-H model is sufficiently simple and requires *seconds or minutes* to run, the use of the Subset Simulation (SS) algorithm may represent the optimal choice (Section 5.2.1);
2. in those (more realistic) cases where the T-H model requires many hours, or days, to perform a single evaluation, the use of fast-running surrogate regression models (e.g., bootstrapped Artificial Neural Networks-ANNs, in this work) instead of the long-running original T-H code seems mandatory (Section 5.2.2).

These recommendations are further explained and motivated below.

5.2.1 Uncertainty propagation using Subset Simulation

The idea underlying the Subset Simulation (SS) method is to convert the simulation of an event (e.g., the rare failure event) into a sequence of simulations of intermediate conditional events

corresponding to subsets (or subregions) of the uncertain input parameter space (for example, if a passive decay heat removal system in a nuclear reactor is assumed to fail when the fuel peak cladding temperature exceeds 725 °C, then plausible intermediate conditional events could be represented by the peak cladding temperature exceeding 350, 500 and 650 °C, respectively). During simulation, the conditional samples (lying in the intermediate subsets or subregions) are generated by means of properly designed Markov chains; by so doing, the conditional samples gradually populate the successive intermediate subsets (or subregions) up to the target (failure) region [Au and Beck, 2001; Au and Beck, 2003b].

In synthesis, the SS algorithm proceeds as follows. First, N vectors $\{\boldsymbol{x}_0^k: k = 1, 2, \ldots, N\}$ are sampled by standard MCS, i.e., from the original probability density function $q(\cdot)$. The corresponding values of the response variable $\{Y(\boldsymbol{x}_0^k): k = 1, 2, \ldots, N\}$ are then computed and the first threshold value y_1 (identifying the first intermediate conditional event) is chosen as the $(1 - p_0)N^{\text{th}}$ value in the increasing list of values $\{Y(\boldsymbol{x}_0^k): k = 1, 2, \ldots, N\}$. With this choice of y_1, there are now p_0N samples among $\{\boldsymbol{x}_0^k: k = 1, 2, \ldots, N\}$ whose response $Y(\boldsymbol{x})$ lies in the intermediate subregion $F_1 = \{\boldsymbol{x}: Y(\boldsymbol{x}) > y_1\}$. Starting from each one of these samples, Markov Chain Monte Carlo (MCMC) simulation is used to generate $(1 - p_0)N$ additional conditional samples in the intermediate subregion $F_1 = \{\boldsymbol{x}: Y(\boldsymbol{x}) > y_1\}$, so that there are a total of N conditional samples $\{\boldsymbol{x}_1^k: k = 1, 2, \ldots, N\} \in F_1$. Then, the intermediate threshold value y_2 is chosen as the $(1 - p_0)N^{\text{th}}$ value in the ascending list of $\{Y(\boldsymbol{x}_1^k): k = 1, 2, \ldots, N\}$ to define $F_2 = \{\boldsymbol{x}: Y(\boldsymbol{x}) > y_2\}$. The p_0N samples lying in F_2 function as 'seeds' for sampling $(1 - p_0)N$ additional conditional samples lying in F_2, making up a total of N conditional samples $\{\boldsymbol{x}_2^k: k = 1, 2, \ldots, N\} \in F_2$. This procedure is repeated until the samples lying in the intermediate subregion $F_{m\text{-}1} = \{\boldsymbol{x}: Y(\boldsymbol{x}) > y_{m\text{-}1}\}$ are generated to yield $y_m > y$ as the $(1 - p_0)N^{\text{th}}$ value in the ascending list of $\{Y(\boldsymbol{x}_{m\text{-}1}^k): k = 1, 2, \ldots, N\}$ [Au and Beck, 2001; Au and Beck 2003b; Au, 2005; Au et al., 2007].

The superior efficiency of SS with respect to standard MCS in the uncertainty propagation task has been widely demonstrated in the open literature: the interested reader may refer to [Au and Beck, 2001; Au and Beck, 2003b] for mathematical details, to [Ching et al., 2005; Katafygiotis and Cheung, 2005 and 2007; Au, 2007; Au et al., 2007; Pradlwarter et al., 2007] for illustrative applications to high-dimensional (i.e., $n \geq 100$) structural reliability problems and to [Zio and Pedroni, 2009b] for an application to the functional failure analysis of a T-H passive system.

For completeness, we report some of the results previously obtained by the authors [Zio and Pedroni, 2009b] in the use of the SS method to propagate the uncertainties through the T-H model of the passive decay heat removal system of a Gas-cooled Fast Reactor (GFR) analyzed in the

previous Section 5.1 [Pagani et al., 2005]. Nine uncertain input parameters $\{x_j: j = 1, 2, \ldots, 9\}$ are taken into account and two safety variables $\{y_l: l = 1, 2\}$ (i.e., the hot- and average-channel temperatures of the naturally circulating coolant leaving the core) are considered as outputs of interest of the T-H system model code. The output variables $\{y_l: l = 1, 2\}$ are then used to generate a single-valued system performance indicator (or critical response variable) $Y(\boldsymbol{x})$ for the evaluation of passive system failure; further details can be found in [Pagani et al., 2005; Zio and Pedroni, 2009b].
The performance of SS is compared to that of LHS: notice that LHS has been chosen as benchmark method due to its popularity and wide use in Probabilistic Risk Assessment (PRA) [Helton and Davis, 2003; Sallaberry et al., 2008; Helton and Sallaberry, 2009]. Following the approach presented in [Au et al., 2007] and subsequently used in [Zio and Pedroni, 2009b], Figure 1, left shows the empirical Cumulative Distribution Function (CDF) of the performance function $Y(\boldsymbol{x})$ of the passive decay heat removal system considered; in addition, Figure 1, right focuses on the portion of CDF where the cumulative probability ranges between 0.999 and 1. The results produced by SS with a total of N_T = 1850 samples (i.e., T-H code runs) are shown in solid lines, whereas those produced by LHS with the same number of samples/T-H code runs (i.e., N_T = 1850) are shown in dashed lines. The dot-dashed lines correspond to the results obtained by LHS with N_T = 500000 samples/T-H code runs: this number of samples is largely sufficient for efficiently estimating the CDF even where the cumulative probability ranges between 0.999 and 1: thus, the corresponding results are taken as benchmarks.

Notice that the results from SS are satisfactorily close to the reference solution in all the probability ranges considered. On the contrary, LHS with 1850 samples is not able to produce accurate results for values of the cumulative probability very close to 1 (Figure 1, right). This is due to the fact that with 1850 samples there are on average only $1850 \cdot (1 - 0.999) = 1850 \cdot 0.001 \sim 2$ samples in Figure 1, right. In contrast, SS (due to successive conditional MCMC simulations) generates 1850 and 500 conditional samples in Figure 1, left and right, respectively, giving enough information for an efficient estimation of the CDF.
Then, the 99.9th percentile of the performance function $Y(\boldsymbol{x})$ of the passive system is estimated by SS with 1850 samples (obtaining 1120.1 °C) and LHS with 1850 (obtaining 1095.3 °C) and 500000 samples (obtaining 1118.9 °C). It can be seen that the estimate of the 99.9th percentile produced by SS with 1850 samples is very *accurate* and close to the reference one, i.e., the one computed by LHS with 500000 samples: however, this result is obtained with a computational effort which is $500000/1850 \approx 270$ times lower; on the contrary, the percentile identified by LHS with 1850 samples is much lower than the reference one.

Finally, to assess quantitatively the statistical properties and the *precision* of the 99.9th percentile estimates produced by SS with 1850 samples and LHS with 1850 samples, S = 100 independent runs have been carried out for each simulation method and the empirical 95% Confidence Intervals (CIs) of the 99.9th percentile estimates thereby obtained have been computed: the obtained CIs are [1068.9, 1183.0] and [1012.4, 1242.1] for SS and LHS, respectively. It can be seen that the width of the 95% CI produced by SS is about 2 times lower than that of LHS: thus, conversely, the precision of the estimate is 2 times higher.

As a final remark, it is worth noting that for SS (differently from LS) there does not seem to exist any indication that it is possible to reduce the number of samples (i.e., the number of T-H model code evaluations) to below a few hundreds. Actually, referring to the computational flow of SS described above, at least N = 100 samples have to be generated in *each* subset F_i, i = 1, 2, …, m, to produce reliable estimates in the uncertainty propagation phase: thus, if high quantiles (e.g., the 99.9th or 99.99th percentiles) have to be estimated (which is often the case for passive safety systems), then an amount of about $N \cdot m = 100 \cdot 3 = 300$ or $N \cdot m = 100 \cdot 4 = 400$ samples have to be generated, respectively. As a consequence, if the T-H model requires many hours, or days, to perform a single evaluation, SS is not suitable.

5.2.2 Uncertainty propagation using bootstrapped Artificial Neural Networks

In those cases where the T-H model requires many hours, or days, to perform a single evaluation, the use of fast-running surrogate regression models instead of the long-running original T-H code becomes somewhat mandatory: because calculations with the surrogate model can be performed quickly, the problem of long simulation times is circumvented.

Here, the use of Artificial Neural Networks (ANNs) is recommended for this task. In extreme synthesis, ANNs are computing devices inspired by the function of the nerve cells in the brain [Bishop, 1995]. They are composed of many parallel computing units (called *neurons* or *nodes*) arranged in different *layers* and interconnected by weighed connections (called *synapses*). Each of these computing units performs a few simple operations and communicates the results to its neighbouring units. From a mathematical viewpoint, ANNs consist of a set of nonlinear (e.g., sigmoidal) basis functions with adaptable parameters $\boldsymbol{w}^*$ that are adjusted by a process of *training* (on many different input/output data examples), i.e., an iterative process of regression error minimization [Rumelhart et al., 1986]. ANNs have been demonstrated to be universal approximants of *continuous* nonlinear functions (under mild mathematical conditions) [Cybenko, 1989], i.e., in principle, an ANN model with a properly selected architecture can be a consistent estimator of any continuous nonlinear function, e.g. any nonlinear T-H code simulating the system of interest.

Further details about ANN regression models are not reported here for brevity; the interested reader may refer to the cited references and the copious literature in the field. The particular type of ANN considered in this paper is the classical three-layered feed-forward ANN trained by the error back-propagation algorithm.

Notice that the recommendation of using ANN regression models is mainly based on i) theoretical considerations about the (mathematically) demonstrated capability of ANN regression models of being *universal* approximants of continuous nonlinear functions (e.g., any nonlinear T-H code simulating the system of interest) [Cybenko, 1989] and ii) the experience of the authors' in the use of ANN regression models for propagating the uncertainties through T-H model codes simulating passive safety systems [Pedroni et al., 2010; Zio et al., 2010]: for example, in [Pedroni et al., 2010], both the accuracy and precision of ANN regression models in estimating the percentiles of the temperature of the naturally circulating coolant in a passive decay heat removal system have been compared and shown to be superior to those of simple quadratic Response Surface (RS) regression models. Since no further comparisons with other types of regression models have been performed by the authors yet, no additional proofs of the superiority of ANNs with respect to other regression models can be provided at present, in general terms.

To evaluate the additional source of *model* uncertainty introduced by the ANN empirical regression model the use of an ensemble of ANN regression models, constructed on different data sets bootstrapped from the original one is recommended [Zio, 2006; Storlie et al., 2009]. The bootstrap method is a distribution-free inference method which requires no prior knowledge about the distribution function of the underlying population [Efron and Thibshirani, 1993]. The basic idea is to generate a sample from the observed data by sampling with replacement from the original data set [Efron and Thibshirani, 1993]: each of these bootstrapped data sets is used to build a bootstrapped regression model which is used to calculate the quantity of interest (e.g., in this case of uncertainty propagation, the quantity of interest may be represented by the vector $\boldsymbol{y}$ of the outputs of the T-H model code, by the performance function of the passive system $Y(\boldsymbol{x})$ and by their percentiles). In this context, the bootstrap algorithm is used to quantify, in terms of *confidence intervals*, the model uncertainty associated to the estimates provided by the ANN regression models. Recall also that from the theory and practice of ensemble empirical models, it can be shown that the estimates given by bootstrapped ANN regression models are in general more accurate than the estimate of the best ANN regression model in the bootstrap ensemble of ANN regression models [Zio, 2006; Cadini et al., 2008].

In synthesis, the following steps must be undertaken to perform uncertainty propagation by means of bootstrapped ANNs [Zio, 2006; Storlie et al., 2009]:

1. Generate a data set D_{train} of training input/output data examples by sampling a (possibly *reduced*) number N_{train} of independent input parameters values $\boldsymbol{x}_p$, p = 1, 2, ..., N_{train}, and calculating the corresponding set of N_{train} output vectors $\boldsymbol{y}_p = \boldsymbol{\mu}_y(\boldsymbol{x}_p)$ through the mechanistic T-H system code.
2. Generate a set D_{val} of validation input/output data examples (different from D_{train}) by sampling a (possibly *reduced*) number N_{val} of independent input parameters values $\boldsymbol{x}_p$, p = 1, 2, ..., N_{val}, and calculating the corresponding set of N_{val} output vectors $\boldsymbol{y}_p = \boldsymbol{\mu}_y(\boldsymbol{x}_p)$ through the mechanistic T-H system code.
3. Build an ANN regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ using the training and validation data sets D_{train} and D_{val}; in particular, the training data set D_{train} is used to calibrate the internal parameters $\boldsymbol{w}^*$ of the regression model, whereas the validation data set D_{val} is used to monitor the accuracy of the ANN model during the training procedure in order to avoid *overfitting* of the training data according to the so-called *early stopping* method. In practice, the RMSE is computed on D_{val} at different iterative stages of the training procedure: at the beginning of training, this value decreases as does the RMSE computed on the training set D_{train}; later in the training, if the ANN regression model starts overfitting the data, the RMSE calculated on the validation set D_{val} starts increasing and training must be stopped [Bishop, 1995].
4. Measure the accuracy of the constructed regression model constructed in step 3. by computing proper numerical figures (e.g., the commonly adopted coefficient of determination R^2 and RMSE) for each output y_l, l = 1, 2, ..., n_o, on a *new* data set $D_{test} = \{(\boldsymbol{x}_p, \boldsymbol{y}_p), p = 1, 2, ..., N_{test}\}$ of size N_{test}, purposely generated for *testing* the regression model built [Marrel et al., 2009], and thus different from those used for training and validation.
5. Use the regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$, in place of the original T-H model code, to provide a point estimate $\hat{Q}$ of the quantity Q of interest (e.g., in this case of uncertainty propagation, the quantity Q may be represented by the vector $\boldsymbol{y}$ of the outputs of the T-H model code, by the performance function of the passive system $Y(\boldsymbol{x})$ and by their percentiles).
6. Build an ensemble of B (e.g., B = 500-1000) regression models $\{\boldsymbol{f}_b(\boldsymbol{x}, \boldsymbol{w}_b^*), b = 1, 2, ..., B\}$ on the basis of bootstrap data sets $D_{train,b} = \{(\boldsymbol{x}_{p,b}, \boldsymbol{y}_{p,b}), p = 1, 2, ..., N_{train}\}$, b = 1, 2, ..., B,

generated by performing random sampling *with replacement* from the original training data set $D_{train} = \{(\boldsymbol{x}_p, \boldsymbol{y}_p), p = 1, 2, ..., N_{train}\}$.

7. Use each of the bootstrapped regression models $\boldsymbol{f}_b(\boldsymbol{x}, \boldsymbol{w}_b^*)$, b = 1, 2, ..., B, to calculate an estimate $\hat{Q}_b$, b = 1, 2, ..., B, for the quantity Q of interest: by so doing, a bootstrap-based *empirical* probability distribution for the quantity Q is produced which is the basis for the construction of the corresponding confidence intervals.
8. Calculate the so-called Bootstrap Bias Corrected (BBC) point estimate $\hat{Q}_{BBC}$ for Q (see [Baxt and White, 1995] for details) and the corresponding two-sided Bootstrap Bias Corrected (BBC)-100·(1 - α)% Confidence Interval (CI) (using the bootstrap-based *empirical* probability distribution for the quantity Q obtained in step 7. above).

The complete and detailed bootstrap algorithm is not reported here for brevity; some technical details can be found in [Efron and Thibshirani, 1993; Zio, 2006; Cadini et al., 2008; Secchi et al., 2008; Storlie et al., 2009; Pedroni et al., 2010; Zio et al., 2010].

For completeness, we report some of the results obtained in a previous work by the authors [Pedroni et al., 2010], in which bootstrapped ANNs are used to propagate the uncertainties through the T-H model of Section 5.1 [Pagani et al., 2005]; again, the performance of bootstrapped ANNs is compared to that of LHS.

Figure 2, left shows the empirical Cumulative Distribution Function (CDF) of the performance function $Y(\boldsymbol{x})$ of the passive decay heat removal system considered; in addition, Figure 2, right focuses on the portion of CDF where the cumulative probability ranges between 0.95 and 1. The results obtained with N_T = 500000 estimations from B = 1000 by bootstrapped ANNs (built on $N_{code} = N_{train} + N_{val} + N_{test}$ = 80 + 20 + 10 = 110 input/output examples, i.e., T-H code runs) are shown in solid lines, whereas those produced by LHS with the *same* number of T-H code runs (i.e., $N_T = N_{code}$ = 110) are shown in dashed lines. Notice that the comparison between these two approaches is fair because the number N_{code} of runs of the original T-H system model code (and thus the associated *overall computational effort*) is the same (i.e., N_{code} = 110); however, for LHS the *few* system model code runs are *directly* used to produce the CDF of interest, whereas for bootstrapped ANNs they are used to build the regression models, which are in turn employed to produce the CDF estimate. The dot-dashed lines correspond to the results obtained by LHS with $N_T = N_{code}$ = 500000 samples (i.e., T-H code runs): this number of samples is largely sufficient for efficiently estimating the CDF even where the cumulative probability ranges between 0.95 and 1: thus, the corresponding results are taken as *benchmarks*.

The bootstrapped ANNs are shown to be quite *reliable* and *accurate*, as the CDF produced is satisfactorily close to the reference one (i.e., the one produced by LHS with $N_T = N_{code} = 500000$ samples) in all the probability ranges considered. Also, the bootstrapped ANN results are obtained at a much lower computational effort: actually, the number N_{code} of T-H code runs (i.e., 110) is about 4500 times lower than that of the reference case (i.e., 500000). The overall CPU time required by the use of bootstrapped ANNs (i.e., on average 2.22 h) is about 180 times lower than that required by the use of the original T-H model code (i.e., on average 409 h).

Further, it can be seen that the bootstrapped ANNs built on $N_{code} = N_{train} + N_{val} + N_{test} = 80 + 20 + 10 = 110$ input/output examples (i.e., T-H code runs) outperform LHS with the same number $N_{code} = 110$ of T-H code simulations: actually, LHS is not able to produce accurate results, in particular for values of the cumulative probability very close to 1 (Figure 2, right).

The two approaches are further compared in the estimation of the 95th percentile of the performance function $Y(\boldsymbol{x})$ of the passive decay heat removal system. The BBC point estimate of the 95th percentile of the performance function $Y(\boldsymbol{x})$ obtained with $N_T = 500000$ estimations from $B = 1000$ bootstrapped ANNs and with $N_T = 110$ and 500000 estimations from LHS are 802.5 °C, 824.1 °C and 794.2 °C, respectively. It can be seen that the estimate produced by the bootstrapped ANNs is quite close to the *reference* one, i.e., the one obtained by LHS with 500000 samples: the corresponding percentage Relative Absolute Error (RAE) is 1.04%; on the contrary, the percentile identified by LHS with $N_T = 110$ samples is considerably larger: the corresponding percentage RAE is 3.74%. It can be seen that the percentage RAE produced by the bootstrapped ANNs is 3.6 times lower than that of LHS with $N_T = 110$ samples: thus, conversely, the *accuracy* of the estimate is 3.6 times higher.

Finally, to assess quantitatively the statistical properties and the *precision* of the 95th percentile estimates produced by the methods considered, the 95% Confidence Interval (CI) associated to the estimates are evaluated. In particular, $S = 1000$ independent runs of LHS with $N_T = N_{code} = 110$ samples are carried out and the *empirical* 95% Confidence Interval (CI) of the 95th percentile estimate thereby obtained has been computed: it turns out to be [785.6, 868.2]; on the contrary, the BBC 95% CI is produced by $B = 1000$ bootstrapped ANN regression models constructed on $N_{code} = 110$ data examples according to steps 1. – 8. above: it turns out to be [777.06, 818.92]. It can be seen that the width of the CI produced by bootstrapped ANNs is about 2 times lower than that of LHS: thus, conversely, the precision of the estimate is 2 times higher.

5.3 Sensitivity analysis

For safety-critical systems, like nuclear passive systems, the task of sensitivity analysis is fundamental for reliability/failure probability assessment and safety decision-making and assurance [Helton and Sallaberry, 2009]. In particular, in the functional failure analysis of a T-H passive system, sensitivity analysis can be a useful tool for identifying the uncertain parameters (i.e., the uncertain inputs to the T-H code) that contribute most to the variability of the model outputs (i.e., the coolant outlet temperatures): this information is important for the identification of those parameter and hypothesis uncertainties that are most relevant in determining system failure [Saltelli et al., 2008; Volkova et al., 2008; Marrel et al., 2009].

In general, the sensitivity analysis outcomes provide two important insights. On the one side, the analyst is able to identify those parameters/variables whose epistemic uncertainty plays a major role in determining the functional failure of the T-H passive system: consequently, his/her efforts can be focused on increasing the state-of-knowledge on these important parameters/variables and the related physical phenomena (for example, by the collection of experimental data one may achieve an improvement in the state-of-knowledge on the correlations used to model the heat transfer process in natural convection and a corresponding reduction in the uncertainty); on the opposite side, the analyst can identify those parameters/variables that are not important and may be excluded from the modeling and analysis.

The options recommended for performing sensitivity analysis are the same as those proposed for uncertainty analysis (Section 5.2), as explained below.

5.3.1 Sensitivity analysis using Subset Simulation

The Markov chain samples generated by SS can be used not only for estimating the conditional probabilities but also to infer the probable scenarios that will occur in the case of failure [Au, 2005]. Intuitively, from the comparison of the probability density function $q(x_j \mid F)$ of the uncertain parameter x_j, j = 1, 2, …, n_i, conditional to the occurrence of failure F, with the unconditional probability density function $q(x_j)$, an indication can be obtained on how important is the parameter x_j in affecting the system failure. Formally, for any given value of x_j the Bayes' theorem reads,

$$P(F \mid x_j) = \frac{q(x_j \mid F)}{q(x_j)} P(F), \; j = 1, 2, \ldots, n_i \tag{12}$$

so that $P(F \mid x_j)$ is insensitive to x_j when $q(x_j \mid F) \sim q(x_j)$, i.e. when the conditional probability density function $q(x_j \mid F)$ is similar in shape to the PDF $q(x_j)$ [Au and Beck, 2003; Au, 2005; Au et al., 2007]. The effectiveness of this approach for sensitivity analysis has been demonstrated by a

number of studies conducted in the field of structural reliability: for example, in [Au and Beck, 2003] and [Au, 2005], the approach has been effectively used to address a 1500-dimensional problem concerning a steel frame subject to stochastic ground motion; in [Au et al., 2007] the method has been applied to perform a compartment fire risk analysis where seven uncertain parameters were considered, whereas in [Zio and Pedroni, 2009b] it has been applied to perform the sensitivity analysis of the model of Section 5.1.

In this latter work, the sensitivity of the passive system performance to the n_i = 9 uncertain input parameters has been studied by examining the change of the sample distributions $q(x_j|F_i)$, j = 1, 2, …, n_i, i = 1, 2, …, m, at different conditional levels F_i, i = 1, 2, …, m. The histograms of the conditional samples of two of the nine uncertain parameters (i.e., x_2, the pressure level established in the guard containment after the LOCA, and x_8, the friction factor in mixed convection) at different conditional levels for a single SS run are shown in Figure 3, left. It can be seen that the performance of the passive system is strongly sensitive to the pressure level established in the guard containment after the LOCA, as indicated by the significant leftward shift of its empirical conditional distribution (histograms) from the unconditional one (solid lines). A slight sensitivity of the passive system performance is also observed with respect to the correlation errors in the friction factor (rightward shift) in mixed convection.
The information contained in the empirical conditional distributions $q(x_j|F_i)$, j = 1, 2, …, n_i, i = 1, 2, …, m, can then be used to *refine* the sensitivity information by obtaining the distribution of the system failure probability conditional on the values of the individual uncertain input parameters, i.e. $P(F|x_j)$, according to (12) (Figure 3, right): this information is relevant because it quantifies how the failure probability $P(F)$ of the passive system would change if the value of the uncertain parameter x_j were set to a given value (e.g., if its epistemic uncertainty were reduced).

Note that SS presents the advantage over other standard techniques of sensitivity analysis, of being directly "embedded" in the computation of the failure probability: the SS algorithm produces the empirical conditional distributions of Figure 3 *during the simulation* that is performed to compute the functional failure probability of the passive system. In other words, *while* estimating the functional failure probability of the system, sensitivity analysis results are produced that can be readily visualized for identification and ranking of the most important variables.

5.3.2 Sensitivity analysis using bootstrapped Artificial Neural Networks

Bootstrapped ANNs are used to replace the original T-H code in the *multiple* (e.g., many thousands) system performance evaluations (for different combinations of system inputs) required by

sensitivity analysis; thus, in principle, bootstrapped ANNs could be used in the development of any of the sensitivity analysis methods available in the open literature.

Here we recommend the use of bootstrapped ANNs for computing first- and total-order Sobol sensitivity indices [Sobol, 1993] for the vector $\boldsymbol{y}$ of the outputs of the T-H code and for the passive system performance function $Y(\boldsymbol{x})$: see [Zio et al., 2010] for a preliminary analysis of this kind.

By definition, the first-order Sobol sensitivity index S_j^l, $j = 1, 2, \ldots, n_i$, $l = 1, 2, \ldots, n_o$, quantifies the proportion of the variance of the output y_l, $l = 1, 2, \ldots, n_o$, that can be attributed to the variance of the uncertain input variable x_j *alone*, i.e., *without* taking into account *interactions* with other input variables; on the contrary, the total-order Sobol sensitivity index S_{Tj}^l, $j = 1, 2, \ldots, n_i$, $l = 1, 2, \ldots, n_o$, quantifies the proportion of the variance of the output y_l, $l = 1, 2, \ldots, n_o$, that can be attributed to the variance of the uncertain input variable x_j taking into account the *interactions* (of *all* the *orders*) with all the other input variables. A thorough description of these sensitivity measures goes beyond the scope of this work: mathematical details can be found in [Saltelli, 2002a, b; Saltelli et al., 2008].

As pointed out in [Saltelli, 2002a], the sensitivity indices S_j^l and S_{Tj}^l have the advantage of being *global* because the effect of the *entire* distribution of the parameter whose uncertainty importance is evaluated, is considered; moreover, this sensitivity index is also "model free" because its computation is independent from assumptions about the model form, such as linearity, additivity and so on. The drawback of this approach relies in the computational burden associated to its calculation: actually, thousands or millions of system model evaluations are frequently required for the evaluation of Sobol indices through Monte Carlo-based techniques [Saltelli, 2002a; Saltelli et al., 2008].

For completeness, we complete the results obtained in a previous work by the authors [Zio et al., 2010] (in which bootstrapped ANNs were applied for computing first-order Sobol indices for one of the outputs of the model of Section 5.1) by computing first- and total-order Sobol indices S_j^Y and S_{Tj}^Y for the performance function $Y(\boldsymbol{x})$ of the model of the T-H passive system of Section 5.1. The algorithm proposed by [Saltelli, 2002a] has been implemented to obtain the "true" (i.e., reference) values of the first- and total-order Sobol sensitivity indices S_j^Y and S_{Tj}^Y for the input variables x_j, $j = 1, 2, \ldots, 9$: these values obtained with $N_T = 110000$ runs of the original T-H model code are reported for reference in Table 5 (in parentheses).

Table 5 reports also the BBC point estimates $\hat{S}_{j,BBC}^Y$ and $\hat{S}_{Tj,BBC}^Y$ for S_j^Y and S_{Tj}^Y, $j = 1, 2, \ldots, 9$, obtained with $N_T = 110000$ estimations from $B = 1000$ bootstrapped ANN models built on $N_{code} = N_{train} + N_{val} + N_{test} = 80 + 20 + 10 = 110$ input/output examples, i.e., T-H code runs; the Table also shows the corresponding Bootstrap Bias Corrected (BBC)-95% Confidence Intervals (CIs): the

information conveyed by these intervals is important when few data are used to train the bootstrapped ANNs and the consequent confidence of the analyst on the Sobol index point estimates $\hat{S}^{Y}_{j,BBC}$ and $\hat{S}^{Y}_{Tj,BBC}$ is poor, like in the present case.

It can be seen that bootstrapped ANNs are quite *accurate* because the BBC point estimates produced are satisfactorily close to the reference values; moreover, ANNs are sufficiently *precise* since the BBC 95% CIs are quite *narrow* around the reference values.
Finally, notice that the computational cost associated to the use of bootstrapped ANNs is much lower than that required by the use of the original T-H code: actually, the computational times associated to both analyses have been of 2.12 h and 92 h, respectively, on a Pentium 4 CPU 3.00GHz.

6 Conclusions

The assessment of the reliability of T-H passive systems is a crucial issue to be resolved for their extensive use in future nuclear power plants. The reliance of T-H passive systems on inherent physical principles makes their reliability evaluation quite difficult to accomplish, if compared to classical system reliability analysis, due to the lack of data which makes current knowledge of passive system operation somewhat poor, thus introducing large *uncertainties* in the analysis. These uncertainties are both of aleatory and epistemic nature and are mainly due to poor understanding and imprecise modelling of the phenomena affecting the T-H performance of the system and of the relative physical correlations, environmental and boundary conditions used.
These issues may in principle be detrimental for the public acceptance of future reactor designs, which conversely are expected to offer an overall, guaranteed level of safety higher than the one of the currently operating nuclear fleet, especially thanks to the adoption of passive systems.
Thus, there is a strong need for the development and demonstration of consistent methodologies and approaches for T-H passive systems reliability assessment.

As a further step forward in this direction, in this paper the computational issues associated with assessing the reliability of T-H passive systems have been considered. The copious use of expert judgement and subjective assumptions during the assessment process leads to the need of propagating the associated uncertainties by simulating several times the system response under different working conditions: this can be done by Monte Carlo sampling the uncertainties in the system model and parameters, and simulating the corresponding passive system response with a mechanistic T-H computer code. However, this approach requires considerable computational

efforts. The reason is twofold. First, a large number of Monte Carlo-sampled T-H model evaluations must generally be carried out for an accurate estimation of the functional failure probability. Since the number of simulations required to obtain a given accuracy depends on the magnitude of the failure probability to be estimated, with the computational burden increasing with decreasing functional failure probability, this poses a significant challenge for the typically quite small (e.g., less than 10^{-4}) probabilities of functional failure of T-H passive safety systems. Second, long calculations (several hours) are typically necessary for each run of the detailed, mechanistic T-H code (one code run is required for each sample of values drawn from the uncertainty distributions).

These computational issues can be tackled in two different ways. From one side, efficient Monte Carlo Simulation techniques can be employed to perform robust estimations with a limited number of input samples; from the other side, fast-running, surrogate regression models (also called response surfaces or meta-models) can be used to replace the long-running T-H model code.

Different approaches have been considered and compared with reference to a case study of literature involving the natural convection cooling in a Gas-cooled Fast Reactor (GFR) after a Loss of Coolant Accident (LOCA) [Pagani et al., 2005].

On the basis of the results obtained in the present and previous works by the authors [Zio and Pedroni, 2009a-c and 2010; Pedroni et al., 2010; Zio et al., 2010], the following guidelines and recommendations can be drawn:

- If the interest is only in an accurate and precise estimation of the (typically small) functional failure probability of the T-H passive system (modelled by a *long-running*, *nonlinear* and *non-monotonous* T-H code), then the following approach is recommended (Section 5.1):
    a. build an Artificial Neural Network (ANN) regression model using a *sequential*, *two-step* training algorithm on a *reduced* number of examples (e.g., around one hundred) of the input/output nonlinear relationships underlying the original system model code;
    b. use the ANN model as a fast-running surrogate of the original system model code in the determination of the LS important direction; the technique recommended for this is that based on the *minimization* of the variance of the LS failure probability estimator by means of Genetic Algorithms: the motivation is that since it relies directly on the definition of the *optimal* LS important direction, it produces more accurate and precise failure probability estimates than those provided by the other techniques proposed in the literature;

c. estimate the functional failure probability of the T-H passive system by means of Line Sampling with a *small* number of samples (e.g., few tens); the accuracy and precision of the estimates can be enhanced by combining Line Sampling with Latin Hypercube Sampling.

It is worth remarking once more that the LS technique allows *only* the calculation of the failure probability of the passive system, whereas it does not allow a *complete* uncertainty propagation.

- If the analyst is interested also in the uncertainty propagation (i.e., determination of the PDFs, CDFs, percentiles of the T-H code outputs of interest and so on) and sensitivity analysis, two options are recommended:

  1. the SS method offers a feasible means because it generates a *large* amount of conditional (failure) samples by *sequential* Markov Chain Monte Carlo (MCMC) simulations developed in different *subsets* of the uncertain input space. This allows producing the PDFs and CDFs of *all* the T-H code outputs of interest (e.g., peak cladding temperatures, pressures, mass flow rates and so on) in a *single* simulation run. Moreover, the conditional samples distributions in different *subsets* of the uncertain input space can be used to study the sensitivity of the passive system performance to the uncertain system input parameters: the informative measure of the importance of a given parameter in determining the failure of the system is the deviation of its conditional distribution from the unconditional one.

     On the other hand, differently from the LS method, there does not seem to exist any indication that it is possible to reduce the number of samples (i.e., the number of T-H model code evaluations) to below a few hundreds. Actually, at least one hundred samples have to be generated in *each* subset to produce reliable failure probability estimates: thus, if the failure probabilities to be estimated are 10^{-4} or 10^{-5} (which is often the case for passive safety systems), then an amount of 400 or 500 samples have to be generated, respectively. As a consequence, if the T-H model requires many hours, or days, to perform a single evaluation, SS is not suitable; on the other hand, if the T-H model is sufficiently simple and requires *seconds or minutes* to run, SS may represent the optimal choice.

  2. in those (realistic) cases where the T-H model requires many hours, or days, to perform a single evaluation, the use of fast-running surrogate regression models (e.g., ANNs, quadratic RSs, …) instead of the long-running original T-H code seems mandatory. The following procedure is recommended:

a. run the T-H system model code a predetermined, *reduced* number of times (e.g., 50-100) for specified values of the uncertain input variables;
b. collect the corresponding values of the output of interest;
c. employ statistical techniques for calibrating/adapting the internal *parameters/coefficients* of the response surface of the regression model in order to fit the input/output data generated in the previous steps;
d. use the empirical regression model built at step c. to estimate the quantities of interest: in this paper, the estimation of i) the CDF of the passive system performance function, ii) its 95th and 99.9^{th} percentiles and iii) first- and total-order Sobol sensitivity indices has been illustrated;
e. use the bootstrap procedure to quantify, in terms of confidence intervals, the uncertainties associated to the estimates provided by the empirical regression models.

It is worth pointing out that the selection of a surrogate regression model suitable to replace the complex, nonlinear T-H code in the uncertainty propagation process is quite a difficult task: actually, such selection is heavily dependent on the particular application at hand, so that no *general* rules are available to this aim.

In the present paper, ANN regression models have been recommended on the basis of i) theoretical considerations about the (mathematically) demonstrated capability of ANN regression models of being *universal* approximants of continuous nonlinear functions (e.g., any nonlinear T-H code simulating the system of interest) [Cybenko, 1989] and ii) the experience of the authors in the use of ANN regression models for propagating the uncertainties through T-H model codes simulating passive safety systems [Pedroni et al., 2010; Zio et al., 2010]. However, since no detailed and systematic comparisons with other types of regression models (except for quadratic Response Surfaces [Pedroni et al., 2010]) have been performed by the authors yet, no additional proofs of the superiority of ANNs with respect to other regression models can be provided at present. Future research will be devoted to address this issue, although it is arguably optimistic to think that a general statement in this direction can be reached.

Finally, a general remark is in order to drive the reader towards a correct interpretation of the numerical results obtained and of the recommendations drawn in the present paper. Actually, one may interpret that the failure probabilities and sensitivity indices computed by means of the

methodologies described and recommended throughout the paper are *the* failure probabilities and sensitivity indices associated to the "*real*" T-H passive system under analysis (i.e., those quantities that would characterize the behavior of the T-H passive system in its operation during a *real* accidental transient). However, in order for this to be true, the T-H code employed in the analyses would need to be flawless and comprehensive of *all* the relevant failure modes of the real T-H passive system, *all* aleatory uncertainties would need to be modeled perfectly, and *all* epistemic uncertainties would need to be well characterized. This is obviously not so and it seems in order to acknowledge that the computational methods described and recommended throughout the paper can "only" do as much, driving the T-H code with its limitations (even if very detailed and extremely demanding to run). In other words, the paper has addressed the quantification of passive system functional reliability "only" from the computational viewpoint, i.e., to the extent that the relevant failure modes are captured in the T-H model code being driven, and to the extent that the input uncertainty distributions are appropriate. Even after consistency checks are run and statistical confidence bounds are established on the results, issues may remain concerning the possibility of "extending" the results obtained in the analyses to the "*actual*" behavior of the "*real*" T-H passive system during an accidental transient, because of the model incomplete representation of reality.

FIGURE CAPTION PAGE

*Figure 1. Empirical CDF of the performance function Y(**x**) of the passive system in [Pagani et al., 2005]. Solid lines: SS with N_T = 1850 samples; dashed lines: LHS with N_T = 1850 samples; dot-dashed lines: reference LHS with N_T = 500000 samples*

*Figure 2. Empirical CDF of the performance function Y(**x**) of the passive system in [Pagani et al., 2005]. Solid lines: N_T = 500000 estimations from B= 1000 bootstrapped ANNs built on N_{code} = N_{train} + N_{val} + N_{test} = 80 + 20 + 10 = 110 input/output examples (i.e., T-H code runs); dashed lines: LHS with N_T = N_{code} = 110 samples (i.e., T-H code runs); dot-dashed lines: reference LHS with N_T = N_{code} = 500000 samples (i.e., T-H code runs)*

Figure 3. Sensitivity analysis by SS. Left: empirical conditional distributions of uncertain input parameters x_2 and x_8 at different conditional levels (histograms) compared to their unconditional distributions (solid lines); right: distribution of the system failure probability conditional on the values of the individual uncertain input parameters x_2 and x_8, i.e., $P(F|x_2)$ and $P(F|x_8)$

TABLES

Category	Description
A	Physical barriers and static structures (e.g., concrete buildings)
B	Moving working fluid (e.g., cooling by free convection)
C	Moving mechanical parts (e.g., check valves)
D	External signals and stored energy (e.g., scram systems)

Table 1. Categorization of passive systems [IAEA, 1991]

	Categories of uncertainties	
ALEATORY	Occurrence of accident scenarios	
	Failure time of mechanical components	
	Variation of geometrical dimensions	
	Variation of material properties	
EPISTEMIC	T-H analysis	Model (correlations)
		Parameters
	System failure analysis	Failure criteria
		Failure modes (critical parameters)

Table 2. Categories of uncertainties associated to nuclear passive systems reliability assessment

Functional failure probability ("True" value, $P(F) = 3.541 \cdot 10^{-4}$)					
			Performance indicators ($N_T = 1850$; $S = 2000$)		
Method	$N_{c,P(F)}$	$N_{c,add}$	$\bar{\varepsilon}$ [%]	$\overline{w}_{CI}$ [%]	***FOM***
Standard MCS (i)	$N_T = 1850$	0	101.681	305.874	$1.081 \cdot 10^3$
LHS (ii)	$N_T = 1850$	0	96.652	305.870	$1.222 \cdot 10^3$
IS (iii)	$N_T = 1850$	110	3.803	18.601	$6.159 \cdot 10^5$
IS + LHS (iv)	$N_T = 1850$	110	3.564	17.970	$7.121 \cdot 10^5$
SS (v)	$N_T = 1850$	0	35.760	183.180	$6.414 \cdot 10^3$
LS (vi)	$3 \cdot N_T = 5550$	110	0.517	2.322	$1.329 \cdot 10^7$
LS + LHS (vii)	$3 \cdot N_T = 5550$	110	0.268	1.102	$8.295 \cdot 10^7$

Table 3. Values of the performance indicators $\bar{\varepsilon}$ (8), $\overline{w}_{CI}$ (10) and FOM (11) obtained with N_T = 1850 samples by methods i)-vii) in the estimation of the functional failure probability P(F) of the passive system in [Pagani et al., 2005]

Functional failure probability ("True" value, $P(F) = 3.541 \cdot 10^{-4}$)					
			Performance indicators ($N_{c,P(F)} = 30$; $S = 2000$)		
Method	N_T	$N_{c,add}$	$\bar{\varepsilon}$ [%]	$\overline{w}_{CI}$ [%]	*FOM*
Standard MCS (i)	$N_{c,P(F)} = 30$	0	206.150	$3.943 \cdot 10^3$	911.541
LHS (ii)	$N_{c,P(F)} = 30$	0	183.080	$3.492 \cdot 10^3$	$1.162 \cdot 10^3$
IS (iii)	$N_{c,P(F)} = 30$	110	29.049	139.280	$1.474 \cdot 10^5$
IS + LHS (iv)	$N_{c,P(F)} = 30$	110	27.182	134.170	$1.679 \cdot 10^5$
SS (v)	/	0	/	/	/
LS (vi)	$N_{c,P(F)}/3 = 10$	110	7.016	36.278	$2.338 \cdot 10^6$
LS + LHS (vii)	$N_{c,P(F)}/3 = 10$	110	4.684	24.154	$5.029 \cdot 10^6$

Table 4. Values of the performance indicators $\bar{\varepsilon}$ (8), $\overline{w}_{CI}$ (10) and FOM (11) obtained with $N_{c,P(F)}$ = 30 T-H code runs by methods i)-vii) in the estimation of the functional failure probability P(F) of the passive system in [Pagani et al., 2005]

Sensitivity analysis using bootstrapped ANNs				
	S_i^Y		S_{Ti}^Y	
Parameters	$\hat{S}_{j,BBC}^Y$ **("reference")**	**BBC-95% CI**	$\hat{S}_{Tj,BBC}^Y$ **("reference")**	**BBC-95% CI**
x_1	$7.6774\cdot10^{-3}$ ($8.6372\cdot10^{-3}$)	[$4.751\cdot10^{-4}$, $8.971\cdot10^{-3}$]	0.0113 (0.0121)	[$9.001\cdot10^{-3}$, 0.0195]
x_2	0.7879 (0.7928)	[0.7792, 0.8158]	0.8259 (0.8391)	[0.8188, 0.8553]
x_3	0.0496 (0.0516)	[0.0331, 0.0510]	0.0546 (0.0434)	[0.0391, 0.0570]
x_4	$3.3248\cdot10^{-6}$ ($8.4218\cdot10^{-6}$)	[0, $8.317\cdot10^{-5}$]	$2.226\cdot10^{-3}$ ($3.0575\cdot10^{-3}$)	[$3.231\cdot10^{-4}$, $4.385\cdot10^{-3}$]
x_5	0.0651 (0.0522)	[0.0583, 0.0767]	0.0711 (0.0833)	[0.0655, 0.0809]
x_6	$1.2317\cdot10^{-4}$ ($6.5814\cdot10^{-5}$)	[0, $3.718\cdot10^{-4}$]	$2.2169\cdot10^{-3}$ ($3.1948\cdot10^{-3}$)	[$3.179\cdot10^{-4}$, $4.862\cdot10^{-3}$]
x_7	$2.4542\cdot10^{-5}$ ($6.0669\cdot10^{-5}$)	[0, $4.239\cdot10^{-4}$]	$2.2013\cdot10^{-3}$ ($3.0618\cdot10^{-3}$)	[$3.447\cdot10^{-4}$, $4.621\cdot10^{-3}$]
x_8	0.0527 (0.0522)	[0.0500, 0.0677]	0.0827 (0.0832)	[0.0718, 0.0955]
x_9	$1.5848\cdot10^{-6}$ ($5.9493\cdot10^{-6}$)	[0, $8.168\cdot10^{-5}$]	$2.1968\cdot10^{-3}$ ($3.0531\cdot10^{-3}$)	[$3.455\cdot10^{-4}$, $4.333\cdot10^{-3}$]

*Table 5. Bootstrap Bias Corrected (BBC) point estimates $\hat{S}_{j,BBC}^Y$ and $\hat{S}_{Tj,BBC}^Y$, j = 1, 2, …, 9, and BBC-95% Confidence Intervals (CIs) of the first- and total-order Sobol sensitivity indices S_j^Y and S_{Tj}^Y, j = 1, 2, …, 9, calculated he performance function Y(**x**) of the model of the T-H passive system in [Pagani et al., 2005]*

FIGURES

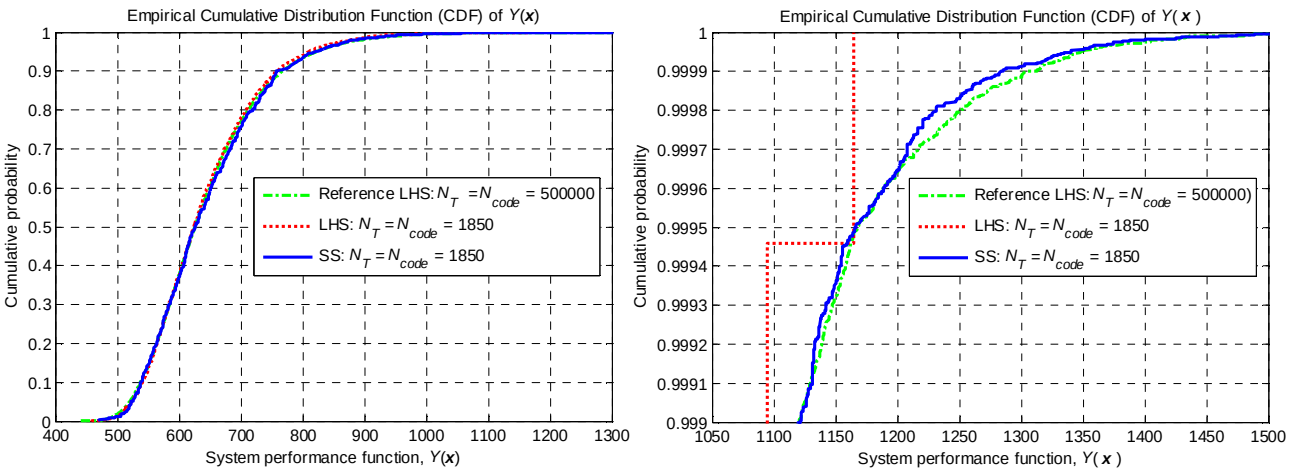

Figure 1

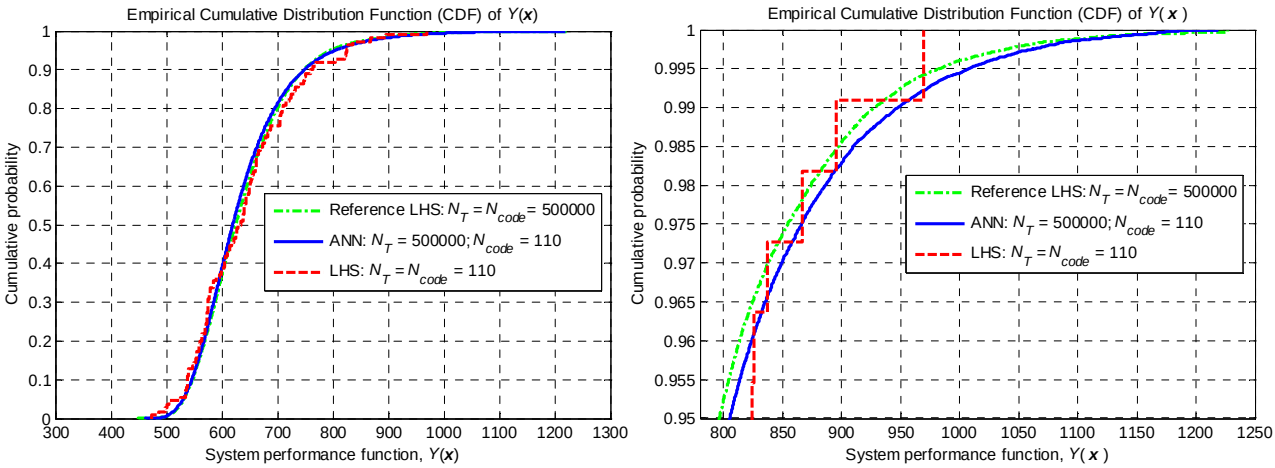
Empirical Cumulative Distribution Function (CDF) of Y(x)
Cumulative probability
System performance function, Y(x)
Reference LHS: N_T = N_code = 500000
ANN: N_T = 500000; N_code = 110
LHS: N_T = N_code = 110
Empirical Cumulative Distribution Function (CDF) of Y(x)
Cumulative probability
System performance function, Y(x)
Reference LHS: N_T = N_code = 500000
ANN: N_T = 500000; N_code = 110
LHS: N_T = N_code = 110

Figure 2

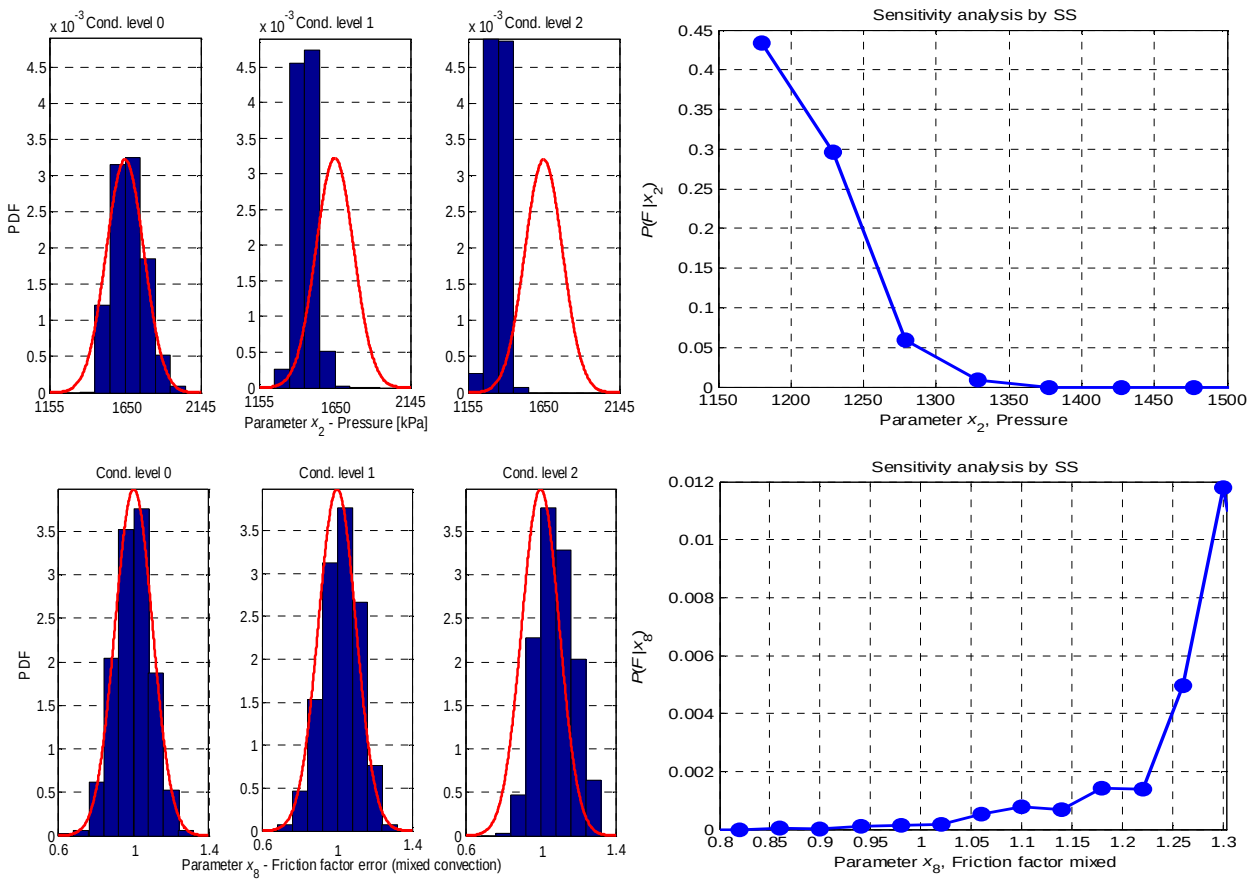
x 10^-3 Cond. level 0
x 10^-3 Cond. level 1
x 10^-3 Cond. level 2
PDF
Parameter x2 - Pressure [kPa]
Sensitivity analysis by SS
P(F|x2)
Parameter x2, Pressure
Cond. level 0
Cond. level 1
Cond. level 2
PDF
Parameter x8 - Friction factor error (mixed convection)
Sensitivity analysis by SS
P(F|x8)
Parameter x8, Friction factor mixed

Figure 3

References

[Ahn et al., 2010] Ahn, SK, Kim, IS, Oh, KM, 2010. Deterministic and risk-informed approaches for safety analysis of advanced reactors: Part I, deterministic approaches. Reliability Engineering and System Safety, 95: 451-458.

[Apostolakis, 1990] Apostolakis, G. E., 1990. The concept of probability in safety assessment of technological systems. Science, 250, 1359.

[Apostolakis, 1995] Apostolakis, G. E., 1995. A commentary on model uncertainty. Proc. Workshop on Model Uncertainty: Its Characterization and Quantification, pp. 13-22. In: Mosleh, A., Siu, N., Smidts, C., and Lui, C., Eds, Center for Reliability Engineering, University of Maryland, College Park, Maryland; also published as NUREG/CP-0138, U.S. Nuclear Regulatory Commission, Washington, D.C. (1994).

[Apostolakis, 1999] Apostolakis, G. E., 1999. The distinction between aleatory and epistemic uncertainties is important: an example from the inclusion of ageing effects into PSA. Proc. Int. Topl. Mtg. Probabilistic Safety Assessment (PSA '99), Washington, D.C., August 22-26, 1999, pp. 135-142, American Nuclear Society, La Grange Park, Illinois.

[Arul et al., 2009] Arul, A.J., Iyer, N.K., Velusamy, K., 2009. Adjoint operator approach to functional reliability analysis of passive fluid dynamical systems. Reliability Engineering and System Safety, 94, pp. 1917-1926.

[Arul et al., 2010] Arul, A.J., Kannan Iyer, N., Velusamy, K., 2010. Efficient reliability estimate of passive thermal hydraulic safety system with automatic differentiation. Nuclear Engineering and Design, doi:10.1016/j.nucengdes.2010.05.012.

[Au, 2004] Au, SK, 2004. Probabilistic Failure Analysis by Importance Sampling Markov Chain Simulation. Journal of Engineering Mechanics, 130(3): 303-311.

[Au, 2005] Au, S. K., 2005. Reliability-based design sensitivity by efficient simulation. Computers and Structures, 83, pp. 1048-1061.

[Au, 2007] Au, S. K., 2007. Augmented approximate solutions for consistent reliability analysis. Probabilistic Engineering Mechanics, 22, pp. 77-87.

[Au and Beck, 2001] Au, S. K. and Beck, J. L., 2001. Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Engineering Mechanics, 16(4), pp. 263-277.

[Au and Beck, 2003a] Au, S. K. and Beck, J. L., 2003. Importance sampling in high dimensions. Structural Safety, 25(2), pp. 139-163.

[Au and Beck, 2003b] Au, S. K. and Beck, J. L., 2003. Subset Simulation and its application to seismic risk based on dynamic analysis. Journal of Engineering Mechanics, 129(8), pp. 1-17.

[Au et al., 2007] Au, S. K., Wang, Z. H., Lo, S. M., 2007. Compartment fire risk analysis by advanced Monte Carlo simulation. Engineering Structures, 29(9), pp. 2381-2390.

[Aumeier, 1994] Aumeier, S.E., 1994. Probabilistic techniques for multi-component system diagnostics and surveillance. Ph.D. dissertation, University of Michigan, Ann Arbor, MI (USA).

[Aumeier and Lee, 1993] Aumeier, S.E., Lee, J.C., 1993. A Monte Carlo technique for system diagnostics, monitoring, and surveillance. Trans. Am. Nucl. Soc., 69, pp. 235-236.

[Aumeier and Lee, 1994] Aumeier, S.E., Lee, J.C., 1994. Probabilistic system diagnostics and surveillance. Proc. ARS '94 – Intl. Topical Meeting on Advanced Reactor Safety, Pittsbourgh (PA), April 17-21, 1994.

[Aumeier et al., 1995] Aumeier, S.E., Lee J.C. and Akcasu, A.Z., 1995. Probabilistic techniques using Monte Carlo Sampling for multi-component system diagnostics. Proc. International conference on mathematics and computations, reactor physics, and environmental analyses, Portland, OR (USA), 30 April - 4 May 1995.

[Aumeier et al., 2006] Aumeier, S.E., Alpay B. and Lee J.C., 2006. Probabilistic Techniques for Diagnosis of Multiple Component Degradations. Nuclear Science and Engineering, 153, 101-123.

[Aybar et al., 1999] Aybar, H. S., Aldemir, T., 1999. Dynamic probabilistic reliability and safety assessment: an application to IS-BWR. In: Proceedings of the Seventh International conference on Nuclear Engineering, Tokio, Japan.

[Bassi and Marquès, 2008] Bassi, C., Marquès, M., 2008. Reliability assessment of 2400 MWth gas-cooled fast reactor natural circulation decay heat removal in pressurized situations. Science and Technology of Nuclear Installations, Special Issue "Natural Circulation in Nuclear Reactor Systems", Hindawi Publishing Corporation, Paper 87376.

[Baxt and White, 1995] Baxt, W. G. and White, H., 1995. Bootstrapping confidence intervals for clinic input variable effects in a network trained to identify the presence of acute myocardial infarction. Neural Computation, 7, pp. 624-638.

[Bishop, 1995] Bishop, C. M., 1995. Neural Networks for pattern recognition. Oxford University Press.

[Bucher and Most, 2008] Bucher, C., Most, T., 2008. A comparison of approximate response function in structural reliability analysis. Probabilistic Engineering Mechanics, vol. 23, pp. 154-163.

[Burgazzi, 2002] Burgazzi, L., 2002. Passive system reliability analysis: a study on the isolation condenser. Nuclear Technology, 139 (July), 3-9.

[Burgazzi, 2003] Burgazzi, L., 2003. Reliability evaluation of passive systems through functional reliability assessment. Nuclear Technology, 144, 145.

[Burgazzi, 2004] Burgazzi, L., 2004. Evaluation of uncertainties related to passive systems performance. Nuclear Engineering and Design, 230, 93-106.

[Burgazzi, 2006] Burgazzi, L., 2006. Failure mode and effect analysis application for the safety and reliability analysis of a thermal-hydraulic passive system. Nuclear Technology, 146, pp. 150-158.

[Burgazzi, 2007a] Burgazzi, L., 2007. Addressing the uncertainties related to passive system reliability. Progress in Nuclear Energy, 49, 93-102.

[Burgazzi, 2007b] Burgazzi, L., 2007. State of the art in reliability of thermal-hydraulic passive systems. Reliability Engineering and System Safety, 92, pp. 671-675.

[Burgazzi, 2007c] Burgazzi, L., 2007. Thermal-hydraulic passive system reliability-based design approach. Reliability Engineering and System Safety, 92(9), pp. 1250-1257.

[Burgazzi, 2008] Burgazzi, L., 2008. About time-variant reliability analysis with reference to passive systems assessment. Reliability Engineering and System Safety, 93(11), pp. 1682-1688.

[Burgazzi, 2009] Burgazzi, L., 2009. Evaluation of the dependencies related to passive system failure. Nuclear Engineering and Design, 239(12), pp. 3048-3053.

[Cacuci and Ionescu-Bujor, 2004] Cacuci, D. G., Ionescu-Bujor, M., 2004. A comparative review of sensitivity and uncertainty analysis of large scale systems – II: Statistical methods. Nuclear Science and Engineering (147), pp. 204-217.

[Cadini et al., 2008] Cadini, F., Zio, E., Kopustinskas, V., Urbonas, R., 2008. An empirical model based bootstrapped neural networks for computing the maximum fuel cladding temperature in a RBMK-1500 nuclear reactor accident. Nuclear Engineering and Design, 238, pp. 2165-2172.

[Cardoso et al., 2008] Cardoso, J. B., De Almeida, J. R., Dias, J. M., Coelho, P. G., 2008. Structural reliability analysis using Monte Carlo simulation and neural networks. Advances in Engineering Software, 39, pp. 505-513.

[Cheng et al., 2008] Cheng, J., Li, Q. S., Xiao, R. C., 2008. A new artificial neural network-based response surface method for structural reliability analysis. Probabilistic Engineering Mechanics, 23, pp. 51-63.

[Ching et al., 2005] Ching, J., Beck, J. L., Au, S. K., 2005. Hybrid subset simulation method for reliability estimation of dynamical systems subject to stochastic excitation. Probabilistic Engineering Mechanics, 20, pp. 199-214.

[Chung et al., 2008] Chung, Y. J., Lee, S. W., Kim S. H., Kim, K. K., 2008. Passive cooldown performance of a 65MW integral reactor, vol. 238, pp. 1681-1689.

[Cybenko, 1989] Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Mathematics of Control Signals Systems, 2, pp. 303-314.

[D' Auria et al., 2002] D' Auria, F., Bianchi, F., Burgazzi, L., Ricotti, M. E., 2002. The REPAS study: reliability evaluation of passive safety systems. In: Proceedings of the 10th International Conference on Nuclear Engineering ICONE 10-22414, Arlinghton, VA USA, April 14-18.

[Deng, 2006] Deng, J., 2006. Structural reliability analysis for implicit performance function using radial basis functions. International Journal of Solids and Structures, 43, pp. 3255-3291.

[Efron and Thibshirani, 1993] Efron, B. and Thibshirani, R. J., 1993. An introduction to the bootstrap. Monographs on statistics and applied probability 57. Chapman and Hall, New York.

[Fong et al., 2009] Fong, C. J., Apostolakis, G. E., Langewish, D. R., Hejzlar, P., Todreas, N. E., Driscoll, M. J., 2009. Reliability analysis of a passive cooling system using a response surface with an application to the flexible conversion ratio reactor. Nuclear Engineering and Design, 239(12), pp. 2660-2671.

[Gavin and Yau, 2008] Gavin, H. P., Yau, S. C., 2008. High-order limit state functions in the response surface method for structural reliability analysis. Structural Safety, 30, pp. 162-179.

[Guba et al., 2003] Guba, A., Makai, M., Pal, L., 2003. Statistical aspects of best estimate method-I. Reliability Engineering and System Safety, 80, pp. 217-232.

[Han and Yang, 2010] Han, SJ, Yang, JE, 2010. A quantitative evaluation of reliability of passive systems within probabilistic safety assessment framework for VHTR. Annals of Nuclear Energy, 37: 345-358.

[Helton, 1998] Helton, J. C., 1998. Uncertainty and sensitivity analysis results obtained in the 1996 performance assessment for the waste isolation power plant, SAND98-0365, Sandia National Laboratories.

[Helton and Oberkampf, 2004] Helton, J. and Oberkampf, W., 2004. Alternative representations of epistemic uncertainties. Reliability Engineering and System Safety, 85 (Special Issue).

[Helton and Davis, 2003] Helton J. C, Davis, F. J., 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliability Engineering and System Safety, 81, pp. 23-69.

[Helton et al., 2005] Helton, J.C., Davis, F.J., Johnson, J.D., 2005. A Comparison of Uncertainty and Sensitivity Analysis Results Obtained with Random and Latin Hypercube Sampling. Reliability Engineering & System Safety, Vol. 89, 3, pp. 305-330.

[Helton et al., 2006] Helton J. C, Johnson J. D., Sallaberry C. J., Storlie C. B., 2006. Survey on sampling-based methods for uncertainty and sensitivity analysis. Reliability Engineering and System Safety, 91, pp. 1175-1209.

[Helton and Sallaberry, 2009] Helton, J.C., Sallaberry, C., 2009. Computational implementation of sampling-based approaches to the calculation of expected dose in performance assessments for the proposed high-level radioactive waste repository at Yucca Mountain, Nevada. Reliability Engineering and System Safety, 94, pp. 699-721.

[Hofer et al., 2002] Hofer, E., Kloos, M., Krzykacz-Hausmann, B., Peschke, J., Woltereck, M., 2002. An approximate epistemic uncertainty analysis approach in the presence of epistemic and aleatory uncertainties. Reliability Engineering and System safety, 77, pp. 229-238.

[Huang and Du, 2006] Huang, B., Du, X., 2006. A robust design method using variable transformation and Gauss-Hermite integration. International Journal for Numerical Methods in Engineering, 66, pp. 1841-1858.

[Hurtado, 2007] Hurtado, J. E., 2007. Filtered importance sampling with support vector margin: a powerful method for structural reliability analysis. Structural Safety, 29, pp. 2-15.

[IAEA, 1991] IAEA, 1991. Safety related terms for advanced nuclear plant. IAEA TECDOC-626.

[Jafari et al., 2003] Jafari, J., D' Auria, F., Kazeminejad, H., Davilu, H., 2003. Reliability evaluation of a natural circulation system. Nuclear Engineering and Design, 224, 79-104.

[Juhn et al., 2000] Juhn, P. E., Kupitz, J., Cleveland, J., Cho, B., Lyon, R. B., 2000. IAEA activities on passive safety systems and overview of international development. Nuclear Engineering and Design, 201, pp. 41-59.

[Kalos and Whitlock, 1986] Kalos, M.H. and Whitlock, P.A., 1986. Monte Carlo methods. Vol. 1: basics. Wiley-Interscience, New York, NY, USA.

[Katafygiotis and Cheung, 2005] Katafygiotis, L., Cheung, S. H., 2005. A two-stage subset simulation-based approach for calculating the reliability of inelastic structural systems subjected to Gaussian random excitations. Computer Methods in Applied Mechanics and Engineering, 194, pp. 1581-1595.

[Katafygiotis and Cheung, 2007] Katafygiotis, L., Cheung, S. H., 2007. Application of spherical subset simulation method and auxiliary domain method on a benchmark reliability study. Structural Safety, 29, pp. 194-207.

[Kim et al., 2010] Kim, IS, Ahn, SK, Oh, KM, 2010. Deterministic and risk-informed approaches for safety analysis of advanced reactors: Part II, Risk-informed approaches. Reliability Engineering and System Safety, 95: 459-468.

[Konak et al., 2006] Konak, A., Coit, D. W., Smith, A. E., 2006. Multi-objective optimization using genetic algorithms: A tutorial. Reliability Engineering and System Safety, Vol. 91, Issue 9, pp. 992-1007.

[Koutsourelakis et al., 2004] Koutsourelakis, P. S., Pradlwarter, H. J., Schueller, 2004. Reliability of structures in high dimensions, Part I: algorithms and application. Probabilistic Engineering Mechanics (19), pp. 409-417.

[Krzykacz-Hausmann, B., 2006] Krzykacz-Hausmann, B., 2006. An approximate sensitivity analysis of results from complex computer models in the presence of epistemic and aleatory uncertainties. Reliability Engineering and System safety, 91, pp. 1210-1218.

[Lee et al., 1993] Lee, J.C., Aumeier, S.E., Patton, B.W., Rank, P.J., 1993. Simulation-based diagnostics and control for nuclear power plants. DOE/ER-75712-1, Progress Report for the period April 15, 1992 – April 14, 1993.

[Lee et al., 1994] Lee, J.C., Aumeier, S.E., Patton, B.W., Rank, P.J., 1993. Simulation-based diagnostics and control for nuclear power plants. DOE/ER-75712-2, Progress Report for the period April 15, 1993 – April 14, 1994.

[Lee et al., 1995] Lee, J.C., Aumeier, S.E., Patton, B.W., Rank, P.J., 1993. Simulation-based diagnostics and control for nuclear power plants. DOE/ER-75712-3, Final Report for the period April 15, 1992 – April 14, 1995.

[Liel et al., 2009] Liel, A. B., Haselton, C. B., Deierlein, G. G., Baker, J. W., 2009. Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. Structural Safety, 31(2), pp. 197-211.

[Lu et al., 2008] Lu, Z., Song, S., Yue, Z., Wang, J., 2008. Reliability sensitivity method by Line Sampling. Structural Safety, vol. 30, pp. 517-532.

[MacKay et al., 1979] MacKay, M. D., Beckman, R. J., Conover, W. J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21(2), pp. 239-245.

[Mackay et al., 2008] Mackay F. J., Apostolakis, G. E., Hejzlar, P., 2008. Incorporating reliability analysis into the design of passive cooling systems with an application to a gas-cooled reactor. Nuclear Engineering and Design, 238(1), pp. 217-228.

[Makai and Pal, 2006] Makai, M., Pal, L., 2006. Best estimate method and safety analysis II. Reliability Engineering and System Safety, 91, pp. 222-232.

[Marquès et al., 2005] Marquès, M., Pignatel, J. F., Saignes, P., D' Auria, F., Burgazzi, L., Müller, C., Bolado-Lavin, R., Kirchsteiger, C., La Lumia, V., Ivanov, I., 2005. Methodology for the reliability evaluation of a passive system and its integration into a probabilistic safety assessment. Nuclear Engineering and Design, 235, 2612-2631.

[Marrel et al., 2009] Marrel, A., Iooss, B., Laurent, B., Roustant, O., 2009. Calculations of Sobol indices for the Gaussian process metamodel. Reliability Engineering and System Safety, vol., 94, pp. 742-751.

[Marseguerra et al., 2006] Marseguerra, M., Zio, E., Martorell, S., 2006. Basics of genetic algorithms optimization for RAMS applications. Reliability Engineering and System Safety, Vol. 91, No. 9, pp. 977-991.

[Mathews et al., 2008] Mathews, T. S., Ramakrishnan, M., Parthasarathy, U., John Arul, A., Senthil Kumar, C., 2008. Functional reliability analysis of safety grade decay heat removal system of Indian 500 MWe PFBR. Nuclear Engineering and Design, 238(9), pp. 2369-2376.

[Mathews et al., 2009] Mathews, T.S., Arul, A.J., Parthasarathy, U., Kumar, C.S., Ramakrishnan, M., Subbaiah, K.V., 2009. Integration of functional reliability analysis with hardware reliability: An application to safety grade decay heat removal system of Indian 500 MWe PFBR. Annals of Nuclear Energy, 36, pp. 481-492.

[Metropolis et al., 1953] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. and Taller, A. H., 1953. Equations of state calculations by fast computing machines. Journal of Chemical Physics, 21(6), pp. 1087-1092.

[Morris, 2000] Morris, M. D., 2000. Three technometrics experimental design classics. Technometrics, 42(1), pp. 26-27.

[Nayak et al., 2008a] Nayak, A. K., Gartia, M. R., Antony, A., Vinod, G., Sinha, R. K., 2008. Passive system reliability analysis using the APSRA methodology. Nuclear Engineering and Design, 238, pp. 1430-1440.

[Nayak et al., 2008b] Nayak, A. K., Jain, V., Gartia, M. R., Srivastava, A., Prasad, H., Anthony, A., Gaikwad, A. J., Bhatia, S. K., Sinha, R. K., 2008. Reliability assessment of passive isolation condenser system using APSRA methodology. Annals of Nuclear Energy, 35, pp. 2270-2279.

[Nayak et al., 2009] Nayak, A. K., Jain, V., Gartia, M. R., Prasad, H., Anthony, A., Bhatia, S. K., Sinha, R. K., 2009. Reliability assessment of passive isolation condenser system of AHWR using APSRA methodology. Reliability Engineering and System Safety, 94, pp. 1064-1075.

[NUREG-1150, 1990] NUREG-1150, 1990. Severe accident risk: an assessment for five US nuclear power plants, US Nuclear Regulatory Commission.

[NUREG-CR-6850, 2005] NUREG-CR-6850, 2005. EPRI/NRC-RES Fire PRA methodology for nuclear power facilities, Volume 2: detailed methodology. US Nuclear Regulatory Commission.

[Nutt and Wallis, 2004] Nutt, W. T., Wallis, G. B., 2004. Evaluations of nuclear safety from the outputs of computer codes in the presence of uncertainties. Reliability Engineering and System Safety, 83, pp. 57-77.

[Olsson et al., 2003] Olsson, A., Sabdberg, G., Dahlblom, O., 2003. On Latin hypercube sampling for structural reliability analysis. Structural Safety, 25, pp. 47-68.

[Pagani et al., 2005] Pagani, L., Apostolakis, G. E. and Hejzlar, P., 2005. The impact of uncertainties on the performance of passive systems. Nuclear Technology, 149, 129-140.

[Patalano et al., 2008] Patalano, G., Apostolakis, G. E., Hejzlar, P., 2008. Risk-informed design changes in a passive decay heat removal system. Nuclear Technology, vol. 163, pp. 191-208.

[Pebesma and Heuvelink, 1999] Pebesma E. J., Heuvelink, G. B. M., 1999. Latin hypercube sampling of Gaussian random fields. Technometrics, 41(4), pp. 203-212.

[Pedroni et al., 2010] Pedroni, N., Zio, E., Apostolakis, G. E., 2010. Comparison of bootstrapped Artificial Neural Networks and quadratic Response Surfaces for the estimation of the functional failure probability of a thermal-hydraulic passive system. Reliability Engineering and System Safety, 95(4): 386-395.

[Pourgol-Mohamad et al., 2010] Pourgol-Mohamad, M., Mosleh, A., Modarres, M., 2009. Methodology for the use of experimental data to enhance model output uncertainty assessment in thermal hydraulics codes. Reliability Engineering and System Safety, Volume 95(2), pp. 77-86.

[Pradlwarter et al., 2005] Pradlwarter, H. J., Pellissetti, M. F., Schenk, C. A., Schueller, G. I., Kreis, A., Fransen, S., Calvi, A., Klein, M., 2005. Realistic and efficient reliability estimation for aerospace structures. Computer Methods in Applied Mechanics and Engineering, 194, pp. 1597-1617.

[Pradlwarter et al., 2007] Pradlwarter, H. J., Schueller, G. I., Koutsourelakis, P. S., Charmpis, D. C., 2007. Application of line sampling simulation method to reliability benchmark problems. Structural Safety, 29, pp. 208-221.

[Rohde et al. 2008] Rohde, M., Marcel, C. P., Manera, A., Van der Hagen, T. H. J. J., Shiralkar, B., 2008. Investigating the ESBWR stability with experimental and numerical tools: a comparative study. Nuclear Engineering and Design, 240(2), pp. 275-384.

[Rumelhart et al., 1986] Rumelhart, D. E., Hinton, G. E., Williams, R. J., 1986. Learning internal representations by error back-propagation. In Rumelhart, D. E. and McClelland, J. L. (Eds.), Parallel distributed processing: exploration in the microstructure of cognition (vol. 1). Cambridge (MA): MIT Press.

[Sallaberry et al., 2008] Sallaberry, C. J., Helton, J. C., Hora, S. C., 2008. Extension of Latin Hypercube samples with correlated variables. Reliability Engineering and System Safety, 93(7), pp. 1047-1059.

[Saltelli, 2002a] Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. Comput. Phys. Commun., vol. 145, pp. 280-297.

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

[Saltelli et al., 2008] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. *Global sensitivity analysis. The Primer*. John Wiley and Sons Ltd.

[Schueller, 2007] Schueller, G. I., 2007. On the treatment of uncertainties in structural mechanics and analysis. Computers and Structures, 85, pp. 235-243.

[Schueller, 2009] Schueller, G. I., 2009. Efficient Monte Carlo simulation procedures in structural uncertainty and reliability analysis - recent advances. Journal of Structural Engineering and Mechanics, 32(1):1–20.

[Schueller and Pradlwarter, 2007] Schueller, G. I., Pradlwarter, H. J., 2007. Benchmark study on reliability estimation in higher dimensions of structural systems – An overview. Structural Safety, 29(3), pp. 167-182.

[Schueller et al., 2004] Schueller, G. I., Pradlwarter, H. J., Koutsourelakis, P. S., 2004. A critical appraisal of reliability estimation procedures for high dimensions. Probabilistic Engineering Mechanics, 19, pp. 463-474.

[Secchi et al., 2008] Secchi, P., Zio, E., Di Maio, F., 2008. Quantifying uncertainties in the estimation of safety parameters by using bootstrapped artificial neural networks. Annals of Nuclear Energy, 35, pp. 2338-2350.

[Sobol, 1993] Sobol, I. M., 1993. Sensitivity analysis for nonlinear mathematical model. Math. Modelling Comput. Exp., 1, pp. 407-414.

[Storlie et al., 2008] Storlie, C. B., Helton, J. C., 2008. Multiple predictor smooting methods for sensitivity analysis: Description of techniques. Reliability Engineering and System Safety, vol. 93, pp. 28-54.

[Storlie et al., 2009] Storlie C.B., Swiler L.P., Helton J.C., Sallaberry C.J., 2009. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. Reliability Engineering and System Safety 2009; 94: 1735-1763.

[USNRC, 2002] USNRC, 2002. An approach for using probabilistic risk assessment in risk-informed decisions on plant-specific changes to the licensing basis. NUREG-1.174 – Revision 1, US Nuclear Regulatory Commission, Washington, DC.

[USNRC, 2009] USNRC, 2009. Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-Informed Decision Making. NUREG-1855, US Nuclear Regulatory Commission, Washington, DC.

[Valdebenito et al., 2010] Valdebenito, M. A., Pradlwarter, H. J., Schueller, G. I., 2010. The role of the design point for calculating failure probabilities in view of dimensionality and structural nonlinearities. Structural Safety, 32(2), pp. 101-111.

[Volkova et al., 2008] Volkova, E., Iooss, B., Van Dorpe, F., 2008. Global sensitivity analysis for a numerical model of radionuclide migration from the RRC "Kurchatov Institute" redwaste disposal site. Stoch Environ Res Assess, 22: pp. 17-31.

[Wilks, 1942] Wilks, S. S., 1942. Statistical prediction with special reference to the problem of tolerance limits. Annals of Mathematical Statistics,13, pp. 400-409.

[Woo and Lee, 2009a] Woo, TH, Lee, UC, 2009. Dynamical reliability of the passive system in the very high temperature gas cooled reactor. Annals of Nuclear Energy, 36: 1299-1306.

[Woo and Lee, 2009b] Woo, TH, Lee, UC, 2009. Passive system reliability in the nuclear power plants (NPPs) using statistical modeling. Nuclear Engineering and Design, 239(12): 3014-3020.

[Woo and Lee, 2010] Woo, TH, Lee, UC, 2010. The statistical analysis of the passive system reliability in the Nuclear Power Plants (NPPs). Progress in Nuclear Energy, 52: 456-461.

[Zio, 2006] Zio, E., 2006. A study of the bootstrap method for estimating the accuracy of artificial neural networks in predicting nuclear transient processes. IEEE Transactions on Nuclear Science, 53(3), pp.1460-1470.

[Zio and Apostolakis, 1996] Zio, E., and Apostolakis, G. E., 1996. Two methods for the structured assessment of model uncertainty by experts in performance assessment in radioactive waste repositories. Reliability Engineering and System Safety, Vol. 54, No. 2, 225-241.

[Zio et al., 2003] Zio, E., Cantarella, M., Cammi, A., 2003. The analytic hierarchy process as a systematic approach to the identification of important parameters for the reliability assessment of passive systems. Nuclear Engineering and Design, 226, 311-336.

[Zio and Pedroni, 2009a] Zio, E., Pedroni, N., 2009. Building confidence in the reliability assessment of thermal-hydraulic passive systems. Reliability Engineering and System Safety, vol. 94(2), pp. 268-281.

[Zio and Pedroni, 2009b] Zio, E. and Pedroni, N., 2009. Estimation of the functional failure probability of a thermal-hydraulic passive systems by means of Subset Simulation. Nuclear Engineering and Design, 239, pp. 580-599.

[Zio and Pedroni, 2009c] Zio, E. and Pedroni, N., 2009. Functional failure analysis of a thermal-hydraulic passive system by means of Line Sampling. Reliability Engineering and System Safety, 94(11), pp. 1764-1781.

[Zio and Pedroni, 2010] Zio, E., Pedroni, N., 2010. An optimized Line Sampling method for the estimation of the failure probability of nuclear passive systems. Reliability Engineering and System Safety, doi:10.1016/j.ress.2010.06.007.

[Zio et al., 2010] Zio, E., Apostolakis, G. E., Pedroni, N., 2010. Quantitative functional failure analysis of a thermal–hydraulic passive system by means of bootstrapped Artificial Neural Networks. Annals of Nuclear Energy, 37(5): 639-649.