Comparison of bootstrapped artificial neural networks and quadratic response surfaces for the estimation of the functional failure probability of a thermal-hydraulic passive system

Nicola Pedroni, Enrico Zio, George Apostolakis

Comparison of bootstrapped Artificial Neural Networks and quadratic Response Surfaces for the estimation of the functional failure probability of a thermal-hydraulic passive system

N. Pedroni[a], E. Zio[a,]* and G. E. Apostolakis[b]
[a]*Energy Department, Polytechnic of Milan, Via Ponzio 34/3, 20133 Milan, Italy*
**Phone: +39-2-2399-6340; fax: +39-2-2399-6309*
E-mail address: enrico.zio@polimi.it

[b]*Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge (MA) 02139-4307*
Phone: +1-617-252-1570; fax: +1-617-258-8863
E-mail address: apostola@mit.edu

Abstract

In this work, bootstrapped Artificial Neural Network (ANN) and quadratic Response Surface (RS) empirical regression models are used as fast-running surrogates of a thermal-hydraulic (T-H) system code to reduce the computational burden associated with the estimation of the functional failure probability of a T-H passive system.

The ANN and quadratic RS models are built on few data representative of the input/output nonlinear relationships underlying the T-H code. Once built, these models are used for performing, in reasonable computational time, the numerous system response calculations required for failure probability estimation. A bootstrap of the regression models is implemented for quantifying, in terms of confidence intervals, the uncertainties associated with the estimates provided by ANNs and RSs.

The alternative empirical models are compared on a case study of an emergency passive decay heat removal system of a Gas-cooled Fast Reactor (GFR).

1 Introduction

All innovative reactor concepts make use of passive safety features, to a large extent in combination with active safety and operational systems (Mackay et al., 2008). According to the definitions of the International Atomic Energy Agency (IAEA), a passive system does not need external input (especially energy) to operate (IAEA, 1991), while they are expected to contribute significantly to the safety of nuclear power plants thanks to their peculiar characteristics of simplicity, reduction of human interaction and reduction or avoidance of hardware failures (Mathews et al., 2008).

However, the uncertainties involved in the modelling and functioning of passive systems are usually larger than for active systems. This is due to: i) the intrinsically random nature of several of the physical phenomena involved in the functioning of the system (aleatory uncertainty); ii) the incomplete knowledge on the physics of some of these phenomena (epistemic uncertainty) (Apostolakis, 1990; Helton, 2004).

Due to these uncertainties, the physical phenomena involved in the passive system functioning (e.g., natural circulation) might develop in such a way to lead the system to fail its intended function, even if safety margins are present. 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 (Marquès et al., 2005; Patalano et al., 2008).

The problem may be analytically framed by introducing the concept of functional failure, whereby a passive system may fail to perform its function due to deviations from its expected behavior which lead the load imposed on the system to exceed its capacity (Burgazzi, 2003; Burgazzi, 2007). This concept has been exploited in a number of works presented in the literature (Jafari et al., 2003; Marquès et al., 2005; Pagani et al., 2005; Bassi and Marquès, 2008; Fong and Apostolakis, 2008; Mackay et al., 2008; Mathews et al., 2008; Patalano et al., 2008; Zio and Pedroni, 2009a and b), in which the passive system is modeled by a detailed, mechanistic T-H system code and the probability of failing to perform 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 and the numerical values of its parameters/variables.

Since the probabilities of functional failure of passive systems are generally very small (e.g., of the order of 10^{-4}), a large number of samples is necessary for acceptable estimation accuracy (Schueller, 2007); given that the time required for each run of the detailed, mechanistic T-H system code is of the order of several hours (Fong and Apostolakis, 2008), the MC simulation-based procedure typically requires considerable computational efforts.

A viable approach to overcome the computational burden associated to the analysis is that of resorting to fast-running, surrogate regression models, also called response surfaces or meta-models, to substitute the long-running T-H model code. The construction of such regression models entails running the T-H model code a predetermined, reasonably large but feasibly small, number of times (e.g., of the order of 50-100) for specified values of the uncertain input parameters/variables and recording the corresponding values of the output of interest; then, statistical techniques are employed for fitting the response surface of the regression model to the input/output data generated. Several kinds of surrogate meta-models have been recently applied to safety related nuclear, structural and hydrogeological problems, including polynomial Response Surfaces (RSs) (Bucher and Most, 2008; Fong and Apostolakis, 2008; Gavin and Yau, 2008; Liel et al., 2009), Gaussian meta-models (Volkova et al., 2008; Marrel et al., 2009) and learning statistical models such as Artificial Neural Networks (ANNs), Radial Basis Functions (RBFs) and Support Vector Machines (SVMs) (Deng, 2006; Hurtado, 2007; Cardoso et al., 2008; Cheng et al., 2008).

In this work, the possibility of using Artificial Neural Networks (ANNs) and quadratic Response Surfaces (RSs) to reduce the computational burden associated to the functional failure analysis of a natural convection-based decay heat removal system of a Gas-cooled Fast Reactor (GFR) (Pagani et al., 2005) is investigated. To keep the practical applicability in sight, a *small* set of input/output data examples is considered available for constructing the ANN and quadratic RS models: different sizes of the (small) data sets are considered to show the effects of this relevant practical aspect. The comparison of the potentials of the two regression techniques in the case at hand is made with respect to the estimation of the 95^{th} percentile of the naturally circulating coolant temperature and the functional failure probability of the passive system.

Actually, the use of regression models in safety critical applications like nuclear power plants still raises concerns with regards to the control of their accuracy; in this paper, the bootstrap method is used for quantifying, in terms of *confidence intervals*, the uncertainty associated to the estimates provided by the ANNs and quadratic RSs (Efron and Thibshirani, 1993; Zio, 2006; Cadini et al., 2008; Secchi et al., 2008; Storlie et al., 2008).

The paper organization is as follows. In Section 2, a snapshot on the functional failure analysis of T-H passive systems is given. Section 3 is devoted to the detailed presentation of the bootstrap-based method for quantifying, in terms of *confidence intervals*, the model uncertainty associated to

the estimates of safety parameters computed by ANN and quadratic RS regression models. In Section 4, the case study of literature concerning the passive cooling of a GFR is presented. In Section 5, the results of the application of bootstrapped ANNs and quadratic RSs to the percentile and functional failure probability estimations are compared. Finally, conclusions are provided in the last Section.

2 The quantitative steps of functional failure analysis of T-H passive systems

The basic steps of the quantitative phase of the functional failure analysis of a T-H passive system are (Marquès et al., 2005):

1. Detailed modeling of the passive system response by means of a deterministic, best-estimate (typically long-running) T-H code.
2. Identification of the parameters/variables, models and correlations (i.e., the inputs to the T-H code) which contribute to the *uncertainty* in the results (i.e., the outputs) of the best estimate T-H calculations.
3. Propagation of the uncertainties through the deterministic, long-running T-H code in order to estimate the functional failure probability of the passive system.

Step 3. above relies on *multiple* (e.g., many thousands) evaluations of the T-H code for different combinations of system inputs; this can render the associated computing cost prohibitive, when the running time for each T-H code simulation takes several hours (which is often the case for T-H passive systems).

The computational issue may be tackled by replacing the long-running, original T-H model code by a fast-running, surrogate regression model (properly built to approximate the output from the true system model). In this paper, classical three-layered feed-forward Artificial Neural Networks (ANNs) (Bishop, 1995) and quadratic Response Surfaces (RSs) (Liel et al., 2009) are considered for this task. The accuracy of the estimates obtained is analyzed by computing a confidence interval by means of the bootstrap method (Efron and Thibshirani, 1993); a description of this technique is provided in the following Section.

3 The bootstrap method for point value and confidence interval estimation

3.1 Empirical regression modelling

As discussed in the previous Section, the computational burden posed by uncertainty and sensitivity analyses of T-H passive systems can be tackled by replacing the long-running, original T-H model code by a fast-running, surrogate regression model. Because calculations with the surrogate model can be performed quickly, the problem of long simulation times is circumvented.

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), $D_{train} = \left\{\left(\boldsymbol{x}_p, \boldsymbol{y}_p\right), p = 1, 2, ..., N_{train}\right\}$ (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.,

$$\boldsymbol{y}(\boldsymbol{x}) = \boldsymbol{\mu}_y(\boldsymbol{x}) + \boldsymbol{\varepsilon}(\boldsymbol{x}). \quad (1)$$

Notice that 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., 2008); for simplicity, in the following we assume $\boldsymbol{\varepsilon}(\boldsymbol{x}) = \boldsymbol{0}$ (Secchi et al., 2008).

The objective of the regression task is to estimate $\boldsymbol{\mu}_y(\boldsymbol{x})$ in (1) 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, $\boldsymbol{y}_p = \boldsymbol{f}(\boldsymbol{x}_p, \boldsymbol{w}^*)$, $p = 1, 2, ..., N_{train}$; for example, the Root Mean Squared Error (RMSE) is commonly adopted to this purpose (Zio, 2006):

$$RMSE = \frac{1}{N_{train} \cdot n_o} \sum_{p=1}^{N_{train}} \sum_{l=1}^{n_o} \left(y_{p,l} - \hat{y}_{p,l}\right)^2, \quad (2)$$

Once built, the regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ can be used in place of the T-H code to calculate any quantity of interest Q, such as the 95^{th} percentile of a physical variable critical for the system under analysis (e.g., the fuel cladding temperature) or the functional failure probability of the passive system.

In this work, the capabilities of quadratic Response Surface (RS) and three-layered feed-forward Artificial Neural Network (ANN) regression models are compared in the computational tasks involved in the functional failure analysis of a T-H passive system. In extreme synthesis, quadratic RSs are polynomials containing linear terms, squared terms and possibly two-factors interactions of the input variables (Liel et al., 2009); the RS adaptable parameters $\boldsymbol{w}^*$ are usually calibrated by straightforward least squares methods. 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) 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). The particular type of ANN employed in this paper is the classical three-layered feed-forward ANN trained by the error back-propagation algorithm.

The details of these two regression models are not reported here for brevity: the interested reader may refer to the cited references and the copious literature in the field.

3.2 The bootstrap method

The approximation of the system output provided by an empirical regression model introduces an additional source of uncertainty, which needs to be evaluated, particularly in safety critical applications like those related to nuclear power plant technology. One way to do this is by resorting to bootstrapped regression models (Efron and Thibshirani, 1993), i.e., an ensemble of regression models constructed on different data sets bootstrapped from the original one (Zio, 2006; Storlie et al., 2008). 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). From the theory and practice of ensemble 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).

In what follows, the steps of the bootstrap-based technique of evaluation of the so-called Bootstrap Bias Corrected (BBC) point estimate $\hat{Q}_{BBC}$ of a generic quantity Q (e.g., a safety parameter) by a regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$, and the calculation of the associated BBC Confidence Interval (CI) are reported (Zio, 2006; Storlie et al., 2008):

1. Generate a set D_{train} of input/output data examples by sampling N_{train} independent input parameters values $\boldsymbol{x}_p$, p = 1, 2, ..., N_{train}, and calculating the corresponding set of N_{train} output vectors $\boldsymbol{y}_p = \boldsymbol{\mu}_{\boldsymbol{y}}(\boldsymbol{x}_p)$ through the mechanistic T-H system code. Plain random sampling, Latin Hypercube Sampling or other more sophisticated experimental design methods can be adopted to select the input vectors $\boldsymbol{x}_p$, p = 1, 2, ..., N_{train} (Gazut et al., 2008).
2. Build a regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ on the basis of the entire data set $D_{train} = \left\{\left(\boldsymbol{x}_p, \boldsymbol{y}_p\right), p = 1, 2, ..., N_{train}\right\}$ (step 1. above) in order to obtain a fast-running surrogate of the T-H model code represented by the unknown nonlinear deterministic function $\boldsymbol{\mu}_{\boldsymbol{y}}(\boldsymbol{x})$ in (1).
3. Use the regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ (step 2. above), in place of the original T-H model code, to provide a point estimate $\hat{Q}$ of the quantity Q, e.g., the 95th percentile of a system variable of interest or the functional failure probability of the T-H passive system.

   In particular, draw a sample of N_T new input vectors $\boldsymbol{x}_r$, r = 1, 2, …, N_T, from the corresponding epistemic probability distributions and feed the regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ with them; then, use the corresponding output vectors $\boldsymbol{y}_r = \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{w}^*)$, r = 1, 2, …, N_T, to calculate the estimate $\hat{Q}$ for Q (the algorithm for computing $\hat{Q}$ is obviously dependent on the meaning of the quantity Q). Since the regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ can be evaluated quickly, this step is computationally costless even if the number N_T of model estimations is very high (e.g., $N_T = 10^5$ or 10^6).
4. Build an ensemble of B (typically of the order of 500-1000) regression models $\left\{\boldsymbol{f}_b\left(\boldsymbol{x}, \boldsymbol{w}_b^*\right), b = 1, 2, ..., B\right\}$ by random sampling with replacement and 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. In particular, repeat the following steps for b = 1, 2, ..., B:

a. Generate a bootstrap data set $D_{train,b} = \{(\boldsymbol{x}_{p,b}, \boldsymbol{y}_{p,b}), p = 1, 2, ..., N_{train}\}$, b = 1, 2, ..., B, by performing random sampling with replacement from the original data set $D_{train} = \{(\boldsymbol{x}_p, \boldsymbol{y}_p), p = 1, 2, ..., N_{train}\}$ of N_{train} input/output patterns (steps 1. and 2. above). The data set $D_{train,b}$ is thus constituted by the same number N_{train} of input/output patterns drawn among those in D_{train} although, due to the sampling with replacement, some of the patterns in D_{train} will appear more than once in $D_{train,b}$, whereas some will not appear at all.

b. Build a regression model $\boldsymbol{f}_b(\boldsymbol{x}, \boldsymbol{w}_b^*)$, b = 1, 2, ..., B, on the basis of the bootstrap data set $D_{train,b} = \{(\boldsymbol{x}_{p,b}, \boldsymbol{y}_{p,b}), p = 1, 2, ..., N_{train}\}$ (step 3.a. above).

c. Use the regression model $\boldsymbol{f}_b(\boldsymbol{x}, \boldsymbol{w}_b^*)$ (step 4.b. above), in place of the original T-H code, to provide a point estimate $\hat{Q}_b$ of the quantity of interest Q. It is important to note that for a correct quantification of the confidence interval the estimate $\hat{Q}_b$ must be based on the same input and output vectors $\boldsymbol{x}_r$ and $\boldsymbol{y}_r$, r = 1, 2, …, N_T, respectively, obtained in step 3. above.

5. Calculate the so-called Bootstrap Bias Corrected (BBC) point estimate $\hat{Q}_{BBC}$ for Q as

$$\hat{Q}_{BBC} = 2\hat{Q} - \hat{Q}_{boot} \tag{3}$$

where $\hat{Q}$ is the estimate obtained with the regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ trained with the original data set D_{train} (steps 2. and 3. above) and $\hat{Q}_{boot}$ is the average of the B estimates $\hat{Q}_b$ obtained with the B regression models $\boldsymbol{f}_b(\boldsymbol{x}, \boldsymbol{w}_b^*)$, b = 1, 2, ..., B (step 4.c. above), i.e.,

$$Q_{boot} = \frac{1}{B}\sum_{b=1}^{B}\hat{Q}_b \ . \tag{4}$$

The BBC estimate $\hat{Q}_{BBC}$ in (3) is taken as the definitive point estimate for Q.

The explanation for expression (3) is as follows. It can be demonstrated that if there is a bias in the bootstrap average estimate $\hat{Q}_{boot}$ in (4) compared to the estimate $\hat{Q}$ obtained with the single regression model $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ (step 3. above), then the same bias exists in the single estimate $\hat{Q}$ compared to the true value Q of the quantity of interest (Baxt and White, 1995). Thus, in order to obtain an appropriate, i.e. bias-corrected, estimate $\hat{Q}_{BBC}$ for the quantity of interest Q, the estimate $\hat{Q}$ must be adjusted by subtracting the corresponding bias ($\hat{Q}_{boot}$ - $\hat{Q}$): as a consequence, the final, bias-corrected estimate $\hat{Q}_{BBC}$ is $\hat{Q}_{BBC}$ = $\hat{Q}$ - ($\hat{Q}_{boot}$ - $\hat{Q}$) = $2\hat{Q}$ - $\hat{Q}_{boot}$.

6. Calculate the two-sided Bootstrap Bias Corrected (BBC)-100·(1 - α)% Confidence Interval (CI) for the BBC point estimate in (3) by performing the following steps:
    a. Order the bootstrap estimates $\hat{Q}_b$, b = 1, 2, ..., B, (step 4.c. above) by increasing values, such that $\hat{Q}_{(i)} = \hat{Q}_b$ for some b = 1, 2, ..., B, and $\hat{Q}_{(1)} < \hat{Q}_{(2)} < ... < \hat{Q}_{(b)} < ... < \hat{Q}_{(B)}$.
    b. Identify the $100{\cdot}\alpha/2^{th}$ and $100{\cdot}(1 - \alpha/2)^{th}$ quantiles of the bootstrapped empirical probability distribution of Q (step 4. above) as the $[B{\cdot}\alpha/2]^{th}$ and $[B{\cdot}(1 - \alpha/2)]^{th}$ elements $\hat{Q}_{([B\cdot\alpha/2])}$ and $\hat{Q}_{([B\cdot(1-\alpha/2)])}$, respectively, in the ordered list $\hat{Q}_{(1)} < \hat{Q}_{(2)} < ... < \hat{Q}_{(b)} < ... < \hat{Q}_{(B)}$; notice that the symbol [·] stands for "closest integer".
    c. Calculate the two-sided BBC-100·(1 - α)% CI for $\hat{Q}_{BBC}$ as

$$\left[\hat{Q}_{BBC} - \left(\hat{Q}_{boot} - \hat{Q}_{([B\cdot\alpha/2])}\right), \hat{Q}_{BBC} + \left(\hat{Q}_{([B\cdot(1-\alpha/2)])} - \hat{Q}_{boot}\right)\right]. \quad (5)$$

An important advantage of the bootstrap method is that it provides confidence intervals for a given quantity Q without making any model assumptions (e.g., normality); a disadvantage is that the computational cost could be high when the set D_{train} and the number of adaptable parameters $\boldsymbol{w}^*$ in the regression models are large.

4 Case study

The case study considered in this work concerns the natural convection cooling in a Gas-cooled Fast Reactor (GFR) under a post-Loss Of Coolant Accident (LOCA) condition (Pagani et al., 2005). The reactor is a 600-MW GFR cooled by helium flowing through separate channels in a silicon carbide matrix core whose design has been the subject of study in the past several years at the Massachussets Institute of Technology (MIT) (Pagani et al., 2005).

A GFR decay heat removal configuration is shown schematically in Figure 1; in the case of a LOCA, the long-term heat removal is ensured by natural circulation in a given number N_{loops} of identical and parallel loops; only one of the N_{loops} loops is reported for clarity of the picture: the flow path of the cooling helium gas is indicated by the black arrows. The loop has been divided into $N_{sections}$ = 18 sections for numerical calculation; technical details about the geometrical and structural properties of these sections are not reported here for brevity: the interested reader may refer to (Pagani et al., 2005).

In the present analysis, the average core power to be removed is assumed to be 18.7 MW, equivalent to about 3% of full reactor power (600 MW): to guarantee natural circulation cooling at this power level, a pressure of 1650 kPa in the loops is required in nominal conditions. Finally, the secondary side of the heat exchanger (i.e., item 12 in Figure 1) is assumed to have a nominal wall temperature of 90 °C (Pagani et al., 2005).

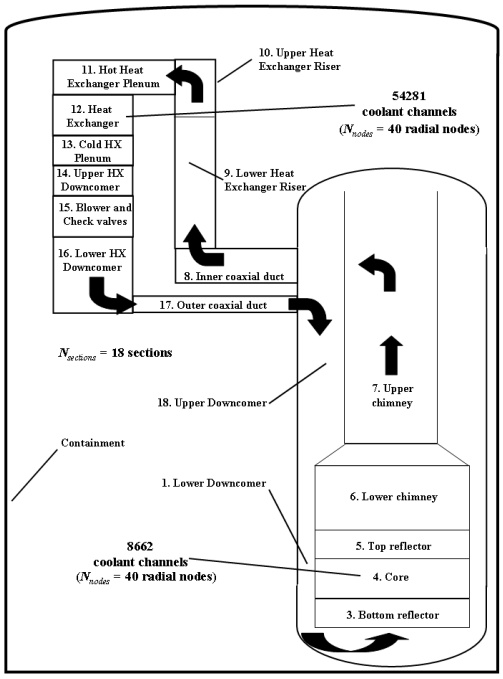

Figure 1. Schematic representation of one loop of the 600-MW GFR passive decay heat removal system (Pagani et al., 2005)

4.1 Uncertainties

Uncertainties affect the modeling of passive systems. There are unexpected events, e.g. the failure of a component or the variation of the geometrical dimensions and material properties, which are random in nature. This kind of uncertainty, often termed aleatory (NUREG-1150, 1990; Helton, 1998; USNCR, 2002), is not considered in this work. There is also incomplete knowledge on the properties of the system and the conditions in which the passive phenomena develop (i.e., natural circulation). This kind of uncertainty, often termed epistemic, affects the model representation of the passive system behaviour, 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).

Only epistemic uncertainties are considered in this work. Epistemic parameter uncertainties are associated to the reactor power level, the pressure in the loops after the LOCA and the cooler wall

temperature; epistemic model uncertainties are associated to the correlations used to calculate the Nusselt numbers and friction factors in the forced, mixed and free convection regimes. The consideration of these uncertainties leads to the definition of a vector $\boldsymbol{x}$ of nine uncertain inputs of the model $\boldsymbol{x} = \{x_j : j = 1,2,...,9\}$, assumed described by normal distributions of known means and standard deviations (Table 1, Pagani et al., 2005).

	Name	Mean, μ	Standard deviation, σ (% of μ)
Parameter uncertainty	Power (MW), x_1	18.7	1%
	Pressure (kPa), x_2	1650	7.5%
	Cooler wall temperature (°C), x_3	90	5%
Model uncertainty	Nusselt number in forced convection, x_4	1	5%
	Nusselt number in mixed convection, x_5	1	15%
	Nusselt number in free convection, x_6	1	7.5%
	Friction factor in forced convection, x_7	1	1%
	Friction factor in mixed convection, x_8	1	10%
	Friction factor in free convection, x_9	1	1.5%

Table 1. Epistemic uncertainties considered for the 600-MW GFR passive decay heat removal system of Figure 1 (Pagani et al., 2005)

4.2 Failure criteria of the T-H passive system

The passive decay heat removal system of Figure 1 is considered failed when the temperature of the coolant helium leaving the core (item 4 in Figure 1) exceeds either 1200 °C in the hot channel or 850 °C in the average channel: these values are expected to limit the fuel temperature to levels which prevent excessive release of fission gases and high thermal stresses in the cooler (item 12 in Figure 1) and in the stainless steel cross ducts connecting the reactor vessel and the cooler (items from 6 to 11 in Figure 1) (Pagani et al., 2005). Denoting by $T_{out,core}^{hot}(\boldsymbol{x})$ and $T_{out,core}^{avg}(\boldsymbol{x})$ the coolant outlet temperatures in the hot and average channels, respectively, the system failure event F can be written as follows:

$$F = \{\boldsymbol{x} : T_{out,core}^{hot}(\boldsymbol{x}) > 1200\} \cup \{\boldsymbol{x} : T_{out,core}^{avg}(\boldsymbol{x}) > 850\}. \quad (6)$$

According to the notation of the preceding Section 3, $T_{out,core}^{hot}(\boldsymbol{x}) = y_1(\boldsymbol{x})$ and $T_{out,core}^{avg}(\boldsymbol{x}) = y_2(\boldsymbol{x})$ are the two target outputs of the T-H model.

5 Functional failure probability estimation by bootstrapped ANNs and quadratic RSs

In this Section, the results of the application of bootstrapped Artificial Neural Networks (ANNs) and quadratic Response Surfaces (RSs) for the estimation of the functional failure probability of the 600-MW GFR passive decay heat removal system in Figure 1 are illustrated. Some details about the construction of the ANN and quadratic RS regression models are given in Section 5.1; their use for estimating the percentiles of the hot-channel and average-channel coolant outlet temperatures is shown in Section 5.2; the estimation of the probability of functional failure of the system is addressed in Section 5.3. The uncertainties associated to the calculated quantities are estimated by bootstrapping of the regression models, as explained in Section 3.

5.1 Building and testing the ANN and quadratic RS regression models

RS and ANN models have been built with training sets $D_{train} = \{(\boldsymbol{x}_p, \boldsymbol{y}_p), p = 1, 2, ..., N_{train}\}$ of input/output data examples of different sizes N_{train} = 20, 30, 50, 70, 100; this has allowed extensive testing of the capability of the regression models to reproduce the outputs of the nonlinear T-H model code, based on different (small) numbers of example data. For each size N_{train} of data set, a Latin Hypercube Sample (LHS) of the 9 uncertain inputs has been drawn, $\boldsymbol{x}_p = \{x_{1,p}, x_{2,p}, \ldots, x_{j,p}, \ldots, x_{9,p}\}$, p = 1, 2, …., N_{train} (Zhang and Foschi, 2004). Then, the T-H model code has been run with each of the input vectors $\boldsymbol{x}_p$, p = 1, 2, …, N_{train}, to obtain the corresponding bidimensional output vectors $\boldsymbol{y}_p = \boldsymbol{\mu}_y(\boldsymbol{x}_p) = \{y_{1,p}, y_{2,p}\}$, p = 1, 2, …, N_{train} (in the present case study, the number n_o of outputs is equal to 2, i.e., the hot- and average-channel coolant outlet temperatures, as explained in Section 4.2). The training data set $D_{train} = \{(\boldsymbol{x}_p, \boldsymbol{y}_p), p = 1, 2, ..., N_{train}\}$ thereby obtained has been used to calibrate the adjustable parameters $\boldsymbol{w}^*$ of the regression models, for best fitting the T-H model code data. More specifically, the straightforward least squares method has been used to find the parameters of the quadratic RSs (Bucher and Most, 2008) and the common error back-propagation algorithm has been applied to *train* the ANNs (Rumelhart et al., 1986). Note that a *single* ANN can be trained to estimate both outputs of the model here of interest, whereas a specific quadratic RS must be developed for each output to be estimated.

The choice of the ANN architecture is critical for the regression accuracy. In particular, the number of neurons in the network determines the number of adjustable parameters available to optimally fit the complicated, nonlinear T-H model code response surface by interpolation of the available training data. The number of neurons in the input layer is n_i = 9, equal to the number of uncertain input parameters; the number n_o of outputs is equal to 2, the outputs of interest; the number n_h of

nodes in the hidden layer is 4 for N_{train} = 20, 30, 70 and 100, whereas it is 5 for N_{train} = 50, determined by trial-and-error. In case of a network with too few neurons (i.e., too few parameters), the regression function $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}^*)$ has insufficient flexibility to adjust its response surface to fit the data adequately: this results in poor generalization properties of interpolation when new input patterns are fed to the network to obtain the corresponding output; on the opposite side, excessively increasing the flexibility of the model by introducing too many parameters, e.g., by adding neurons, may make the network overfit the training data, leading again to poor generalization performance when interpolating new input data. A trade-off is typically sought by controlling the neural model complexity, i.e., the number of parameters, and the training procedure, e.g., by adding a *regularization* term in the error function or by *early stopping* the training, so as to achieve a good fit of the training data with a reasonably smooth regression function which is not over-fit to the data and therefore capable of generalization when interpolating new input data (Bishop, 1995). In the present work, early stopping is adopted: a *validation* input/output data set $D_{val} = \left\{ \left(\boldsymbol{x}_p, \boldsymbol{y}_p \right), p = 1, 2, ..., N_{val} \right\}$ made of patterns different from those of the training set D_{train} is used to monitor the accuracy of the ANN model during the training procedure; in practice, the RMSE (2) is computed on D_{val} at different iterative stages of the training procedure (Figure 2): 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). It is fair to point out that the increased ANN generalization capability typically achieved by early stopping is obtained at the expense of N_{val} additional code simulations, with an increase in the computational cost for the training of the ANN model. In this work, the size N_{val} of the validation set is set to 20 for all sizes N_{train} of the data set D_{train} considered, which means 20 additional runs of the T-H model code.

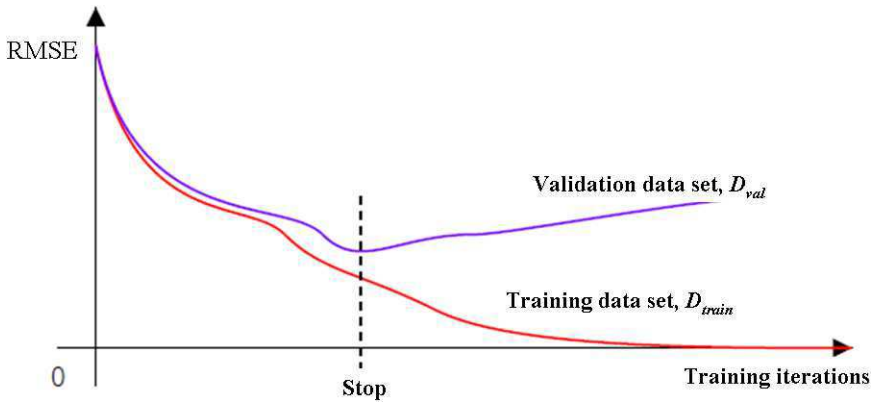

Figure 2. Early stopping the ANN training to avoid overfitting

As measures of the ANN and RS model accuracy, the commonly adopted coefficient of determination R^2 and RMSE have been computed for each output y_l, l = 1, 2, on a new data set $D_{test} = \{(\boldsymbol{x}_p, \boldsymbol{y}_p), p = 1, 2, ..., N_{test}\}$ of size N_{test} = 20, purposely generated for *testing* the regression models built (Marrel et al., 2009), and thus different from those used during training and validation. Table 2 reports the values of the coefficient of determination R^2 and of the RMSE associated to the estimates of the hot- and average- channel coolant outlet temperatures $T_{out,core}^{hot}$ and $T_{out,core}^{avg}$, respectively, computed on the test set D_{test} by the ANN and quadratic RS models built on data sets D_{train} of different sizes N_{train} = 20, 30, 50, 70, 100; the number of adjustable parameters $\boldsymbol{w}^*$ included in the two regression models is also reported for comparison purposes.

Artificial Neural Network (ANN)							
				R^2		**RMSE [°C]**	
N_{train}	N_{val}	N_{test}	**Number of adjustable parameters *w****	$T_{out,core}^{hot}$	$T_{out,core}^{avg}$	$T_{out,core}^{hot}$	$T_{out,core}^{avg}$
20	20	20	50	0.8937	0.8956	38.5	18.8
30	20	20	50	0.9140	0.8982	34.7	18.6
50	20	20	62	0.9822	0.9779	15.8	8.7
70	20	20	50	0.9891	0.9833	12.4	6.8
100	20	20	50	0.9897	0.9866	12.0	6.3
Quadratic Response Surface (RS)							
				R^2		**RMSE [°C]**	
N_{train}	N_{val}	N_{test}	**Number of adjustable parameters *w****	$T_{out,core}^{hot}$	$T_{out,core}^{avg}$	$T_{out,core}^{hot}$	$T_{out,core}^{avg}$
20	0	20	55	0.5971	0.7914	75.0	26.6
30	0	20	55	0.8075	0.9348	51.9	14.8
50	0	20	55	0.9280	0.9353	31.7	14.6
70	0	20	55	0.9293	0.9356	31.4	14.3
100	0	20	55	0.9305	0.9496	31.2	13.1

Table 2. Coefficient of determination R^2 and RMSE associated to the estimates of the hot- and average-channel coolant outlet temperatures $T_{out,core}^{hot}$ and $T_{out,core}^{avg}$, respectively, computed on the test set D_{test} of size N_{test} = 20 by the ANN and quadratic RS models built on data sets D_{train} of different sizes N_{train} = 20, 30, 50, 70, 100; the number of adjustable parameters $\boldsymbol{w}^*$ included in the two regression models is also reported for comparison purposes

The ANN outperforms the RS in all the cases considered: for example, for N_{train} = 100, the coefficients of determination R^2 produced by the ANN and the quadratic RS models for the hot-

channel coolant outlet temperature $T_{out,core}^{hot}$ are 0.9897 and 0.9305, respectively, whereas the corresponding RMSEs are 12.0 °C and 31.2 °C, respectively. This result is due to the higher flexibility in modeling complex nonlinear input/output relationships offered by the ANN with respect to the quadratic RS: the ANN structure made of a large number of adaptable connections (i.e., the synapses) among nonlinear operating units (i.e., the neurons) allows fitting complex nonlinear functions with an accuracy which is superior to that of a plain quadratic regression model. Actually, if the original T-H model is not quadratic (which is often the case in practice), a second-order polynomial RS cannot be a *consistent* estimator, i.e., the quadratic RS estimates may never converge to the true values of the original T-H model outputs, even for a very large number of input/output data examples, in the limit for $N_{train} \rightarrow \infty$. On the contrary, 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.

5.2 Determination of the 95th percentiles of the coolant outlet temperatures

For illustration purposes, a configuration with N_{loops} = 3 loops is considered for the passive system of Figure 1.

The $100 \cdot \alpha^{th}$ percentiles of the hot- and average-channel coolant outlet temperatures $T_{out,core}^{hot}$ and $T_{out,core}^{avg}$ are defined as the values $T_{out,core}^{hot,\alpha}$ and $T_{out,core}^{avg,\alpha}$, respectively, such that

$$P\left(T_{out,core}^{hot} \leq T_{out,core}^{hot,\alpha}\right) = \alpha \tag{7}$$

and

$$P\left(T_{out,core}^{avg} \leq T_{out,core}^{avg,\alpha}\right) = \alpha . \tag{8}$$

Figure 3, left and right, shows the Probability Density Function (PDF) and Cumulative Distribution Function (CDF), respectively, of the hot-channel coolant outlet temperature $T_{out,core}^{hot}$ obtained with N_T = 250000 simulations of the original T-H model code (solid lines); the PDF and CDF of the average-channel coolant outlet temperature $T_{out,core}^{avg}$ are not shown for brevity. The same Figure also shows the PDFs and CDFs constructed with N_T = 250000 estimations from B = 1000 bootstrapped ANNs (dashed lines) and RSs (dot-dashed lines) built on N_{train} = 100 input/output examples.

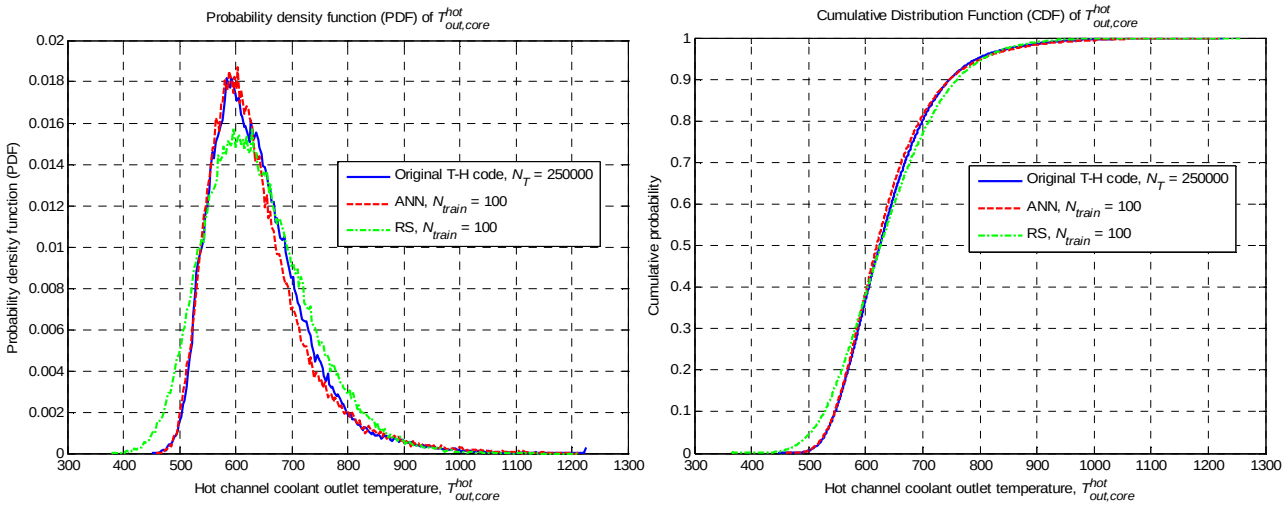

Figure 3. Hot-channel coolant outlet temperature empirical PDFs (left) and CDFs (right) constructed with N_T = 250000 estimations from the original T-H code (solid lines) and from bootstrapped ANNs (dashed lines) and RSs (dot-dashed lines) built on N_{train} = 100 data examples

Notice that the "true" (i.e., reference) PDF and CDF of $T^{hot}_{out,core}$ (Figure 3, solid lines) have been obtained with a very large number N_T (i.e., N_T = 250000) of simulations of the original T-H code, to provide a robust reference for the comparisons. Actually, the T-H code here employed runs fast enough to allow repetitive calculations (one code run lasts on average 3 seconds on a Pentium 4 CPU 3.00GHz): the computational time required by this reference analysis is thus 250000·3 s = 750000 s ≈ 209 h.

The overall good match between the results from the original T-H model code and those from the bootstrapped ANNs and RSs regression models leads us to assert that the accuracy in the estimates can be considered satisfactory for the needs of percentile estimation in the functional failure analysis of the present T-H passive system. Also, it can be seen that the ANN estimates (dashed lines) are much closer to the reference results (solid lines) than the RS estimates (dot-dashed lines).

To quantify the uncertainties associated to the point estimates obtained, bootstrapped ANNs and quadratic RSs have been built to provide Bootstrap Bias Corrected (BBC) point estimates $\hat{T}^{hot,0.95}_{out,core,BBC}$ and $\hat{T}^{hot,0.95}_{out,core,BBC}$ for the 95th percentiles $T^{hot,0.95}_{out,core}$ and $T^{avg,0.95}_{out,core}$ of the hot- and average-channel coolant outlet temperatures $T^{hot}_{out,core}$ and $T^{avg}_{out,core}$, respectively. Figure 4 shows the values (dots) of the BBC point estimates $\hat{T}^{hot,0.95}_{out,core,BBC}$ (top) and $\hat{T}^{hot,0.95}_{out,core,BBC}$ (bottom) obtained with N_T = 250000 estimations from B = 1000 bootstrapped ANNs (left) and quadratic RSs (right) built on N_{train} = 20, 30, 50, 70 and 100 data examples; also the corresponding Bootstrap Bias Corrected (BBC) 95% Confidence Intervals (CIs) (bars) are reported.

Again, notice that the "true" (i.e., reference) values of the 95th percentiles (i.e., $T_{out,core}^{hot,0.95}$ = 796.31 °C and $T_{out,core}^{avg,0.95}$ = 570.22 °C, shown as dashed lines in Figure 4) have been calculated with a very large number N_T (i.e., N_T = 250000) of simulations of the original T-H code, to provide a robust reference for the comparisons: the computational time required by the analysis is 209 h.

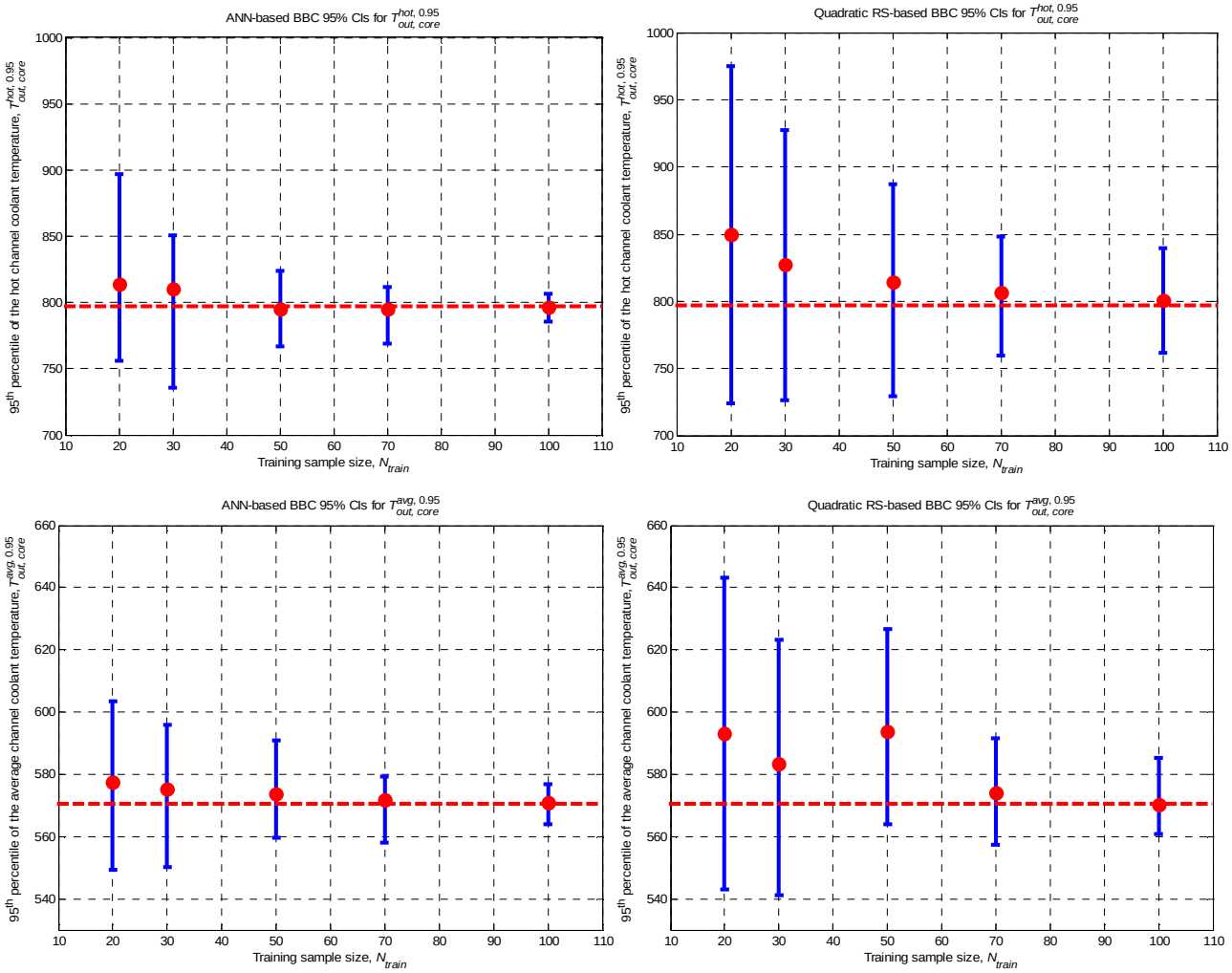

Figure 4. Bootstrap Bias Corrected (BBC) point estimates $\hat{T}_{out,core,BBC}^{hot,0.95}$ *and* $\hat{T}_{out,core,BBC}^{hot,0.95}$ *(dots) and BBC 95% Confidence Intervals (CIs) (bars) for the 95th percentiles* $T_{out,core}^{hot,0.95}$ *and* $T_{out,core}^{avg,0.95}$ *of the hot- (top) and average- (bottom) channel coolant outlet temperatures* $T_{out,core}^{hot}$ *and* $T_{out,core}^{avg}$*, respectively, obtained with* N_T *= 250000 estimations from bootstrapped ANNs (left) and RSs (right) built on* N_{train} *= 20, 30, 50, 70 and 100 data examples; the "true" (i.e., reference) values (i.e.,* $T_{out,core}^{hot,0.95}$ *= 796.31 °C and* $T_{out,core}^{avg,0.95}$ *= 570.22 °C) are shown as dashed lines*

Bootstrapped ANNs turn out to be quite *reliable* and *robust*, providing BBC point estimates very close to the real values in all the cases considered; on the contrary, bootstrapped quadratic RSs provide accurate estimates only for N_{train} = 70 and 100. For example, for N_{train} = 20 the ANN and quadratic RS BBC point estimates $\hat{T}_{out,core,BBC}^{hot,0.95}$ for $T_{out,core}^{hot,0.95}$ = 796.31 °C are 813.50 °C and 849.98 °C, respectively; on the contrary, for N_{train} = 100 the same estimates become 796.70 °C and 800.81 °C, respectively. The superior performance of the ANNs can again be explained by the higher flexibility in nonlinear modeling offered by them with respect to RSs.

Moreover, the uncertainty associated to the bootstrapped ANN estimates is significantly lower than that associated to the quadratic RS estimates, as demonstrated by the width of the corresponding confidence intervals: for example, for N_{train} = 100 the widths of the BBC 95% CIs produced by bootstrapped ANNs and quadratic RSs for $T_{out,core}^{hot,0.95}$ are 21.40 °C and 78.00 °C, respectively. This difference in performance is related to the problem of overfitting (Section 5.1) which can become quite relevant in the bootstrap procedure for the calculation of the BBC 95% CIs (Section 3). The calculation requires that B bootstrap samples $D_{train,b}$, b = 1, 2, …, B, be drawn at random with replacement from the original set D_{train} of input/output data examples: by so doing, some of the patterns in D_{train} will appear more than once in the individual samples $D_{train,b}$, whereas some will not appear at all. As a consequence, the number of *unique* (i.e., *different*) data in each bootstrap sample $D_{train,b}$ will be typically lower than the number N_{train} of "physical" data: this is particularly true if N_{train} is low (e.g., equal to 20 or 30). Since during the bootstrap (step 4. in Section 3) the number of adjustable parameters $\boldsymbol{w}^*$ in each "trained" regression model is fixed, it frequently happens that the number of adaptable parameters $\boldsymbol{w}^*$ is larger than the number of *unique* data in the individual bootstrap sample $D_{train,b}$: this typically causes the regression model to overfit the bootstrap "training" data $D_{train,b}$ with consequent degradation of estimation performance. In the case of ANNs, the early stopping method described in Section 5.1 allows avoiding the overfitting; on the contrary, to the best of the authors' knowledge, no method of this kind is available for polynomial RSs. This explains the higher accuracy of ANN, which within the bootstrap resampling procedure results in a lower "dispersion" of the corresponding bootstrap estimates $\hat{Q}_b$, b = 1, 2, …, B, and in a smaller width of the produced confidence intervals (step 4.c. in Section 3).

Finally, the computational times associated to the calculation of the BBC point estimates $\hat{T}_{out,core,BBC}^{hot,0.95}$ and $\hat{T}_{out,core,BBC}^{hot,0.95}$ for $T_{out,core}^{hot,0.95}$ and $T_{out,core}^{avg,0.95}$, and the corresponding BBC 95% CIs, are compared for the two bootstrapped regression models with reference to the case of N_{train} = 100, by way of example: the overall CPU times required by the use of bootstrapped ANNs and RSs are on average 2.22 h and 0.43 h, respectively. These values include the time required for:

i. generating the N_{train} + N_{val} + N_{test} input/output examples, by running the T-H code: the corresponding CPU times are on average (100 + 20 + 20)·3 = 420 s = 7 min ≈ 0.12 h and (100 + 0 + 20)·3 = 360 s = 6 min ≈ 0.10 h for the ANNs and the RSs, respectively;
ii. training the bootstrapped ensemble of B = 1000 ANN and RS regression models by means of the error back-propagation algorithm and the least squares method, respectively: the corresponding CPU times are on average 2 h and 0.25 h for the ANNs and the RSs, respectively;
iii. performing N_T = 250000 evaluations of *each* of the B = 1000 bootstrapped ANN and RS regression models; the corresponding CPU times are on average 6 min (i.e., 0.1 h) and 4.5 min (i.e., about 0.08 h) for the ANNs and the RSs, respectively.

The overall CPU times required by the use of bootstrapped ANNs (i.e., on average 2.22 h) and quadratic RSs (i.e., on average 0.43 h) is about 90 and 480 times, respectively, lower than that required by the use of the original T-H model code (i.e., on average 209 h). The CPU time required by the ANNs is about 5 times larger than that required by the quadratic RSs, mainly due to the elaborate training algorithm needed to build the structurally complex neural model.

5.3 Functional failure probability estimation

In this Section, the bootstrapped ANNs and quadratic RSs are compared in the task of estimating the functional failure probability of the 600-MW GFR passive decay heat removal system of Figure 1. The previous system configuration with N_{loops} = 3 is analyzed.

Figure 5 shows the values of the Bootstrap Bias Corrected (BBC) point estimates $\hat{P}(F)_{BBC}$ (dots) of the functional failure probability $P(F)$ obtained with N_T = 500000 estimations from the bootstrapped ANNs (left) and quadratic RSs (right) built on N_{train} = 20, 30, 50 ,70 and 100 data examples; the corresponding Bootstrap Bias Corrected (BBC) 95% Confidence Intervals (CIs) are also reported (bars). Notice that the "true" (i.e., reference) value of the functional failure probability $P(F)$ (i.e., $P(F) = 3.34 \cdot 10^{-4}$, shown as dashed lines in Figure 5) has been obtained with a very large number N_T (i.e., N_T = 500000) of simulations of the original T-H code to provide a robust term of comparison: the computational time required by this reference analysis is thus 500000·3 s = 1500000 s ≈ 417 h.

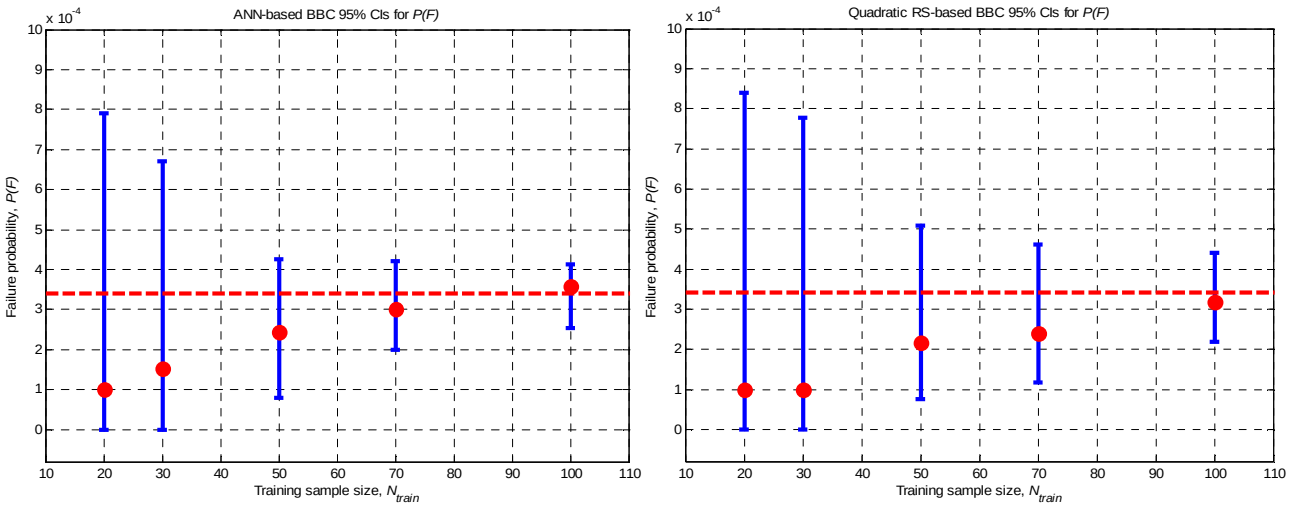

Figure 5. Bootstrap Bias Corrected (BBC) point estimates $\hat{P}(F)_{BBC}$ *(dots) and BBC 95% Confidence Intervals (CIs) (bars) for the functional failure probability P(F) obtained with* N_T *= 500000 estimations from bootstrapped ANNs (left) and RSs (right) built on* N_{train} *= 20, 30, 50, 70 and 100 data examples; the "true" (i.e., reference) value for P(F) (i.e., P(F) =* $3.34 \cdot 10^{-4}$*) is shown as a dashed line*

It can be seen that as the size of the training sample N_{train} increases, both the ANN and quadratic RS provide increasingly accurate estimates of the true functional failure probability $P(F)$, as one would expect. On the other hand, in the cases of small training sets (e.g., N_{train} = 20, 30 and 50) the functional failure probabilities are significantly underestimated by both the bootstrapped ANN and the quadratic RS models (e.g., the BBC point estimates $\hat{P}(F)_{BBC}$ for $P(F)$ lie between $9.81 \cdot 10^{-5}$ and $2.45 \cdot 10^{-4}$) and the associated uncertainties are quite large (e.g., the widths of the corresponding BBC 95% CIs are between $3.47 \cdot 10^{-4}$ and $7.91 \cdot 10^{-4}$). Two considerations seem in order with respect to these results. First, in these cases of small data sets available the analyst would still be able to correctly estimate the order of magnitude of a *small* failure probability (i.e., $P(F) \sim 10^{-4}$), in spite of the *low* number of runs of the T-H code performed to generate the N_{train} = 20, 30 or 50 input/output examples; second, the accuracy of an estimate should be evaluated in relation to the requirements of the specific application; for example, although the confidence interval provided by the bootstrapped ANNs trained with N_{train} = 50 samples ranges from $8.03 \cdot 10^{-5}$ to $4.27 \cdot 10^{-4}$, this variability might be acceptable for demonstrating that the system satisfies the target safety goals.

Finally, it is worth noting that although bootstrapped ANNs provide better estimates and lower model uncertainties than quadratic RSs, the difference in the performances of the two regression models is less evident than in the case of percentile estimation (Section 5.2). This may be due to the fact that estimating the value of the functional failure probability $P(F)$ is a simpler task than

estimating the exact values of the corresponding coolant outlet temperatures. For example, let the true value of the hot channel coolant outlet temperature be 1250 °C and the corresponding estimate by the regression model be 1500 °C: in such a case, the estimate is absolutely *inaccurate* in itself, but "*exact*" for the purpose of functional failure probability estimation with respect to a failure threshold of 1200 °C.

Finally, the computational times required for the estimation of the functional failure probability, and the corresponding confidence interval, in the case of N_{train} = 100 are 2.32 h and 0.50 h for the bootstrapped ANNs and quadratic RSs, respectively.

6 Conclusions

In this paper, Artificial Neural Networks (ANNs) and quadratic Response Surfaces (RSs) have been compared in the task of estimating, in a fast and efficient way, the probability of functional failure of a T-H passive system. A case study involving the natural convection cooling in a Gas-cooled Fast Reactor (GFR) after a Loss of Coolant Accident (LOCA) has been taken as reference. To allow accurate comparison values based on a large number of repeated T-H-model code evaluations, the representation of the system behavior has been limited to a steady-state model.

ANN and quadratic RS models have been constructed on the basis of sets of data of limited, varying sizes, which represent examples of the nonlinear relationships between 9 uncertain inputs and 2 relevant outputs of the T-H model code (i.e., the hot- and average-channel coolant outlet temperatures). Once built, such models have been used, in place of the original T-H model code, to: compute the temperatures 95^{th} percentiles of the hot-channel and average-channel temperatures of the coolant gas leaving the reactor core; estimate the functional failure probability of the system by comparison of the computed values with predefined failure thresholds. In all the cases considered, the results have demonstrated that ANNs outperform quadratic RSs in terms of estimation accuracy: as expected, the difference in the performances of the two regression models is much more evident in the estimation of the 95^{th} percentiles than in the (easier) task of estimating the functional failure probability of the system. Due to their flexibility in nonlinear modelling, ANNs have been shown to provide more reliable estimates than quadratic RSs even when they are trained with very low numbers of data examples (e.g., 20, 30 or 50) from the original T-H model code.

The bootstrap method has been employed to estimate confidence intervals on the quantities computed: this uncertainty quantification is of paramount importance in safety critical applications, in particular when few data examples are used. In this regard, bootstrapped ANNs have been shown to produce *narrower* confidence intervals than bootstrapped quadratic RSs in all the analyses performed.

On the basis of the results obtained, bootstrapped ANNs can be considered more effective than quadratic RSs in the estimation of the functional failure probability of T-H passive systems (while quantifying the uncertainty associated to the results) because they provide more *accurate* (i.e., estimates are closer to the true values) and *precise* (i.e., confidence intervals are narrower) estimates than quadratic RSs; on the other hand, the computational time required by bootstrapped ANNs is somewhat longer than that required by quadratic RSs, due to the elaborate training algorithm for building the structurally complex neural model.

References

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

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

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

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

Burgazzi, L., 2007. State of the art in reliability of thermal-hydraulic passive systems. Reliability Engineering and System Safety, 92(5), pp. 671-675.

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, 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, 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, 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.

Cybenko, G., 1989. Approximation by superposition of sigmoidal functions. Mathematics of Control, Signals and Systems, 2, pp. 303-314.

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

Fong, C. J., Apostolakis, G. E., 2008. The use of response surface methodology to perform uncertainty analyses on passive safety systems. Proceedings of PSA '08, International Topical Meeting on Probabilistic Safety Assessment, Knoxville, Tennessee, September 7-11, 2008, American Nuclear Society, La Grange Park, Illinois.

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.

Gazut, S., Martinez, J. M., Dreyfus, G., Oussar, Y., 2008. Towards the optimal design of numerical experiments. IEEE Transactions on Neural Networks, 19(5), pp. 874-882.

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, J., 2004. Alternative representations of epistemic uncertainties. Reliability Engineering and System Safety, 85 (Special Issue).

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.

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. Safety related terms for advanced nuclear plant. IAEA TECDOC-626.

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

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.

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.

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, 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.

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.

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

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, 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.

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.

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

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.

Storlie, C. B., Swiler, L. P., Helton, J. C., Sallaberry, C. J., 2008. Implementation and evaluation of nonparameteric regression procedures for sensitivity analysis of computationally demanding models. SANDIA Report n. SAND2008-6570.

USNRC, 1998. "An approach for using probabilistic risk assessment in risk-informed decisions on plant-specific changes to the licensing basis." NUREG-1.174, US Nuclear Regulatory Commission, Washington, DC.

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.

Zhang, J. and Foschi, R. O., 2004. Performance-based design and seismic reliability analysis using designed experiments and neural networks. Probabilistic Engineering Mechanics, 19., pp. 259-267.

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, E. and Pedroni, N., 2009a. 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, E. and Pedroni, N., 2009b. Functional failure analysis of a thermal-hydraulic passive system by means of Line Sampling. Reliability Engineering and System Safety, Volume 9, Issue 11, pp. 1764-1781.