A Bayesian framework of inverse uncertainty quantification with principal component analysis and Kriging for the reliability analysis of passive safety systems

Original
A Bayesian framework of inverse uncertainty quantification with principal component analysis and Kriging for the reliability analysis of passive safety systems / Roma, G.; Di Maio, F.; Bersano, A.; Pedroni, N.; Bertani, C.; Mascari, F.; Zio, E.. - In: NUCLEAR ENGINEERING AND DESIGN. - ISSN 0029-5493. - ELETTRONICO. - 379:(2021), p. 111230. [10.1016/j.nucengdes.2021.111230]

Publisher:
Elsevier Ltd

Published
DOI:10.1016/j.nucengdes.2021.111230

A Bayesian framework of inverse uncertainty quantification with principal component analysis and Kriging for the reliability analysis of passive safety systems

Giovanni Roma[1], Francesco Di Maio[2], Andrea Bersano[1], Nicola Pedroni[1*], Cristina Bertani[1], Fulvio Mascari[3], Enrico Zio[2,4,5]

[1]Energy Department, Politecnico di Torino, Corso Duca degli Abruzzi, 24, Torino, 10129, Italy

[2] Energy Department, Politecnico di Milano, Via La Masa 34, Milano, 20156, Italy.

[3]ENEA- BOLOGNA, Via Martiri Monte Sole 4, Bologna, 40129, Italy

[4]MINES ParisTech, PSL Research University, CRC, Sophia Antipolis, France

[5]Department of Nuclear Engineering, College of Engineering, Kyung Hee University, South Korea.

*Corresponding Author: nicola.pedroni@polito.it

Abstract

In this work, we propose an Inverse Uncertainty Quantification (IUQ) approach to assigning Probability Density Functions (PDFs) to uncertain input parameters of Thermal-Hydraulic (T-H) models used to assess the reliability of passive safety systems. The approach uses experimental data within a Bayesian framework. The application to a RELAP5-3D model of the PERSEO (In-Pool Energy Removal System for Emergency Operation) facility located at SIET laboratory (Piacenza, Italy) is demonstrated. Principal Component Analysis (PCA) is applied for output dimensionality reduction and Kriging meta-modeling is used to emulate the reduced set of RELAP5-3D code outputs. This is done to decrease the computational cost of the Markov Chain Monte Carlo (MCMC) posterior sampling of the uncertain input parameters, which requires a large number of model simulations.

Keywords: Passive safety systems, Reliability Analysis, Inverse uncertainty quantification, Surrogate modeling, Kriging, Principal component analysis.

Acronyms

AIC: Akaike's Information Criterion
BCs: Boundary conditions
BE: Best Estimate
BEPU: Best Estimate Plus Uncertainty
CIRCÉ: Calcul des Incertitudes Relatives aux Corrélation Élementaires
CV: Cross Validation
DIPE: Determination of Input parameters Empirical properties

DOE: Design Of Experiment
EM: Expectation Maximization
ENEA: Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile
FFTBM: Fast Fourier Transform Based Method
FOM: Figure of Merit
HX: Heat Exchanger
HXP: Heat Exchanger Pool
ICs: Initial Conditions
IUQ: Inverse Uncertainty Quantification
LHS: Latin Hypercube Sampling
LOO: Leave-One-Out
MAP: Maximum A Posteriori
MCDA: Model Calibration through Data Assimilation
MCMC: Markov Chain Monte Carlo
ML: Maximum Likelihood
MM: Meta-Model
NPPs: Nuclear Power Plants
OP: Overall Pool
PCA: Principal Component Analysis
PCE: polynomial Chaos Expansion
PCs: Principal Component Vectors
PDF: Probability Density Function
PERSEO: in-Pool Energy Removal System for Emergency Operation
PV: Pressure Vessel
QoI: Quantity of interest
RELAP: Reactor Excursion and Leak Analysis Program
SA: Sensitivity Analysis
T-H: Thermal-Hydraulic
TV: Triggering Valve

List of symbols

Symbols	Dimension	Description
$\boldsymbol{d}$	$\mathbf{1} \times \mathbf{1}$	Number of calibration parameters
$\boldsymbol{r}$	$\mathbf{1} \times \mathbf{1}$	Number of design variables
$\boldsymbol{m}$	$\mathbf{1} \times \mathbf{1}$	Number of design points
$\boldsymbol{p}$	$\mathbf{1} \times \mathbf{1}$	Model output dimension
$\boldsymbol{N_{exp}}$	$\mathbf{1} \times \mathbf{1}$	Number of independent experimental measurements
$\boldsymbol{\theta}$	$\boldsymbol{d} \times \mathbf{1}$	Calibration parameters vector
$\boldsymbol{x}$	$\boldsymbol{r} \times \mathbf{1}$	Design variable vector
$\boldsymbol{y^M}$	$\boldsymbol{p} \times \mathbf{1}$	BE computer model output
$\boldsymbol{y^E}$	$\boldsymbol{p} \times \mathbf{1}$	Experimental data
$\boldsymbol{\delta}$	$\boldsymbol{p} \times \mathbf{1}$	Model discrepancy
$\boldsymbol{\epsilon}$	$\boldsymbol{p} \times \mathbf{1}$	Measurement error

$\boldsymbol{I\sigma^2_{exp}}$	$\boldsymbol{p \times p}$	Covariance matrix for the measurement error
$\boldsymbol{\Sigma_{code}}$	$\boldsymbol{p \times p}$	Covariance matrix for the code uncertainty
$\boldsymbol{\Sigma_{Kriging}}$	$\boldsymbol{p^* \times p^*}$	Covariance matrix for the code uncertainty in the PC subspace
$\boldsymbol{R^2_{prediction}}$	$\boldsymbol{1 \times 1}$	CV-derived performance indicator
$\boldsymbol{AIC_c}$	$\boldsymbol{1 \times 1}$	Corrected AIC performance indicator
$\boldsymbol{\widetilde{y}^M}$	$\boldsymbol{p \times 1}$	BE computer model output for Test 7 pt.2
$\boldsymbol{\Theta}$	$\boldsymbol{m \times d}$	Ensemble of $\boldsymbol{m}$ design points
$\boldsymbol{\widetilde{Y}}$	$\boldsymbol{p \times m}$	Ensemble of $\boldsymbol{m}$ BE computer model outputs
$\boldsymbol{\overline{\mu}_{\widetilde{Y}}}$	$\boldsymbol{p \times 1}$	Column vector of the row means of $\boldsymbol{\widetilde{Y}}$
$\boldsymbol{p^*}$	$\boldsymbol{1 \times 1}$	Dimension of the principal component subspace
$\boldsymbol{\Phi}$	$\boldsymbol{p^* \times p}$	Trasformation matrix (PCA)
$\boldsymbol{\tilde{z}}$	$\boldsymbol{p^* \times 1}$	BE computer model output projected in the PCs subspace
$\boldsymbol{\hat{z}}$	$\boldsymbol{p^* \times 1}$	Kriging prediction in the PCs subspace
$\boldsymbol{z^E}$	$\boldsymbol{p^* \times 1}$	Experimental data projected in the PCs subspace
$\boldsymbol{\widetilde{Z}}$	$\boldsymbol{p^* \times m}$	Ensemble of $\boldsymbol{m}$ BE computer model outputs projected in the PCs subspace
$\boldsymbol{\check{y}}$	$\boldsymbol{p \times 1}$	Reconstructed Kriging prediction
$\boldsymbol{\check{y}^E}$	$\boldsymbol{p \times 1}$	Reconstructed experimental data

1 Introduction

The reliability analysis of a passive safety system requires the identification of all the possible accidental scenarios in which the passive system is called to operate (Marquès et al., 2005) and the system capability to accomplish a target mission (e.g., to keep the maximum pressure below a prescribed safety threshold or exchange at least a certain amount of energy in a defined time interval) is verified. To this aim, experiments and Thermal-Hydraulic (T-H) model simulations are performed with different Initial Conditions (ICs), Boundary Conditions (BCs) and different input parameter values. With respect to the simulations, uncertainties, both *aleatory* and *epistemic* (Oberkampf et al., 2004), affect the reliability evaluation results (Burgazzi, 2007, 2004; Pagani et al., 2005).

Aleatory uncertainty refers to the inherent variability associated with random phenomena (Hoffman and Hammonds, 1994). Epistemic uncertainty derives from incomplete information of the phenomena, which reduces the predictiveness of realistic modeling (Apostolakis, 1994; Durga Rao et al., 2007; Ferson and Ginzburg, 1996; Winkler, 1996). The latter uncertainty can be improved by increased knowledge and information of the phenomena (e.g., the availability of new experimental and/or field data). The quantification of the uncertainty (both aleatory and epistemic) in the quantitative description by T-H models of the phenomena pertaining to the function of passive systems typically consists in (Burgazzi, 2004; IAEA, 2014, 2008): (1) the identification of the uncertain input parameters that characterize the phenomena occurring in the system, (2) the quantification of the uncertainties by expert judgment, in terms of nominal values, probability density functions and upper and lower bounds. A particular difficulty in this is the assignment of probability distributions to uncertain parameters (IAEA, 2014). The purpose of the present

work is to determine an innovative method to determine the quantitative description of input parameters affected by epistemic uncertainty. The method implements an inverse uncertainty quantification (IUQ) process to find a characterization of the input parameters' uncertainty that is consistent with the experimental data.
In PREMIUM (Post-BEMUSE Reflood Model Input Uncertainty Methods) (NEA/CSNI/R(2016)9, 2016), a benchmark activity endorsed by OECD/NEA/CSNI/WGAMA, the FFTBM (Fast Fourier Transform Based Method), CIRCÉ (Calcul des Incertitudes Relatives aux Corrélation Élementaires), MCDA (Model Calibration through Data Assimilation) and the DIPE (Determination of Input parameters Empirical properties) methods are compared (de Crecy, 1996; Heo et al., 2014; Kovtonyuk et al., 2012; NEA/CSNI/R(2016)9, 2016). The DIPE method (NEA/CSNI/R(2016)9, 2016) implements an algorithm that compares the T-H code results with experimental data to compute the input parameter distributions. However, this method assumes that: 1) the measurement error is negligible; 2) for each time-instance, there exists a combination of input parameters such that the code output is equal to the experimental value; 3) the T-H code responses are monotonous. The FFTBM method (Kovtonyuk et al., 2012) allows characterizing the ranges of variation of the uncertain input parameters by analyzing a dimensionless figure of merit related to the discrepancy between the code and experimental results. However, this method is incapable of describing dependency and providing a full statistical description of the input parameters (NEA/CSNI/R(2016)9, 2016). The CIRCÉ method (de Crecy, 1996) implements the Expectation-Maximization (EM) algorithm and the Maximum Likelihood Estimate (MLE) method to estimate the input model parameters mean and variance. However, this approach assumes: (1) a linear relationship between model responses and input parameters; (2) the input model parameters to be normally distributed. These two limitations are addressed in the MCDA method, based on Bayesian statistics, where a Markov Chain Monte Carlo (MCMC) algorithm (Gelman et al., 2015) is adopted to determine the input parameters distributions). However, this latter approach results to be cumbersome in the case of a computationally-demanding T-H system model. This issue can be addressed by (Faes et al., 2019) that proposed an IUQ novel multivariate interval approach (i.e., a non-probabilistic methodology) that is very robust in the case of scarce experimental data and computationally affordable, but still incapable to describe dependency and to provide a full statistical description of the input parameters.

In the present work, we address the issues limiting the approaches of (de Crecy, 1996; Faes et al., 2019; Kovtonyuk et al., 2012; Shrestha and Kozlowski, 2016) and propose an IUQ approach based on the Bayesian framework developed by Kennedy and O'Hagan (Kennedy and O'Hagan, 2001). To perform Bayesian inference of the posterior distribution (Gelman et al., 2015), a MCMC sampling technique is adopted, making use of Kriging emulators instead of the original computationally-demanding T-H system model. The proposed IUQ approach is applied to a RELAP5-3D input model (Idaho National Laboratory, 2015) that simulates the behavior of the PERSEO (In-Pool Energy Removal System for Emergency Operation (Ferri et al., 2005)) experimental transients (PERSEO is an experimental facility located at the SIET laboratory (Piacenza, Italy)). Specifically, Principal Component Analysis (PCA) is first carried out to reduce the

dimensionality of the model output; then, Kriging emulators are developed to map the uncertain input parameters on the reduced output Principal Components (PCs). This allows dealing with the large computational cost of MCMC sampling with demanding model codes. The validity of the idea behind the proposed approach is supported by the fact that other works of literature have combined PCA with surrogate modeling for the Bayesian IUQ process to cope with high dimensional computer code outputs (Higdon et al., 2008; Nagel et al., 2020; Wilkinson, 2010; Wu et al., 2018a).
The specific contributions of the present work include: (1) the development of a structured IUQ procedure capable of describing dependency between input parameters and of dealing with time-dependent experimental data; (2) the attainment of insights on the best Kriging options for the development of surrogate models; (3) the observations and reflections on the relationship between the IUQ results and sensitivity analysis.

The remainder of the paper is organized as follows. Section 2 formulates the IUQ problem in mathematical terms. In Section 3, Kriging meta-modeling is briefly recalled. The case study regarding the PERSEO experimental facility and the RELAP5-3D model are introduced in Section 4. Section 5 shows the PCA-based Kriging as multivariate meta-model. Section 6 presents the application of the framework presented in Section 2 on the case study of Section 4. Section 7 concludes the work.

2 The IUQ problem formulation

Let $\boldsymbol{y^M} = \boldsymbol{y^M}(\boldsymbol{x}, \boldsymbol{\theta})$ be the output of a computer code numerically evaluating the outcome of phenomena described by a model, where $\boldsymbol{x} = [x_1, x_2, \dots, x_r]^T$ and $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_d]^T$ are the *design variables* of the system in which the phenomena develop and the model *calibration parameters*, respectively (Kennedy and O'Hagan, 2001; Wu et al., 2018b). Design variables are, for example, Initial Conditions (ICs), Boundary Conditions (BCs), and, in general, all the observable inputs that describe the conditions or scenarios under which the phenomena develop (Wu et al., 2018b). Calibration parameters can be of physical meaning related to the system and phenomenon (e.g., material properties) or without physical meaning (e.g., related to the model nodalization). The calibration parameters $\boldsymbol{\theta}$ are the focus of the IUQ, since the design variables $\boldsymbol{x}$ are defined, known a priori (Wu et al., 2018b).

In general, we deal with different sources of uncertainty (Kennedy and O'Hagan, 2001): *parameter uncertainty* due to the calibration inputs $\boldsymbol{\theta}$, *model discrepancy* (also called *model inadequacy*) due to correlations' inaccuracy or numerical approximations in $\boldsymbol{y^M}(\boldsymbol{x}, \boldsymbol{\theta})$, *measurement error* (also called *observation error*) and *code uncertainty* (or *meta-model uncertainty*, when a Meta-Model (MM) substitutes a computationally intensive Best Estimate (BE) computational model).

An IUQ approach is presented here to infer the uncertainty about $\boldsymbol{\theta}$ using observed experimental data. This requires the definition of the statistical model representing the underlying process that generates the experimental data. Such a statistical model is defined through the formulation of a *model updating equation.*

In mathematical terms, we define the measured experimental data $\boldsymbol{y^E}(\boldsymbol{x})$:

$$\boldsymbol{y^E}(\boldsymbol{x}) = \boldsymbol{y^R}(\boldsymbol{x}) + \boldsymbol{\epsilon} \tag{1}$$

where $\boldsymbol{y^R}(\boldsymbol{x})$ is the real (unknown) value of the measured quantity and $\boldsymbol{\epsilon} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{I}\sigma_{exp}^2)$ is an additive measurement error, here assumed to be Gaussian-distributed.

On the other hand, $\boldsymbol{y^M}(\boldsymbol{x}, \boldsymbol{\theta})$ is affected by the model discrepancy $\boldsymbol{\delta}(\boldsymbol{x})$, so we write:

$$\boldsymbol{y^R}(\boldsymbol{x}) = \boldsymbol{y^M}(\boldsymbol{x}, \boldsymbol{\theta}^*) + \boldsymbol{\delta}(\boldsymbol{x}) \tag{2}$$

Thus, $\boldsymbol{\delta}$ and $\boldsymbol{y^R}$ depend only on $\boldsymbol{x}$, whereas $\boldsymbol{y^M}$ depends on both $\boldsymbol{x}$ and $\boldsymbol{\theta}^*$, where $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ is the vector of the unknown "best" parameter values (i.e., the values that, when plugged into $\boldsymbol{y^M}(\boldsymbol{x}, \boldsymbol{\theta})$, give the most accurate prediction of $\boldsymbol{y^R}(\boldsymbol{x})$ in eq. (2) (Wu et al., 2018b)). Summing equations (1) and (2), we get the model updating equation for the experimental output $\boldsymbol{y^E}(\boldsymbol{x})$, that accounts for *parameter uncertainty*, *model discrepancy* and *measurement error* (Arendt et al., 2012):

$$\boldsymbol{y^E}(\boldsymbol{x}) = \boldsymbol{y^M}(\boldsymbol{x}, \boldsymbol{\theta}^*) + \boldsymbol{\delta}(\boldsymbol{x}) + \boldsymbol{\epsilon} \tag{3}$$

Neglecting $\boldsymbol{\delta}(\boldsymbol{x})$ in the model updating equation (3) is in principle wrong and causes "over-fitting" (Wu et al., 2018b). It means that the vector of "best" input parameters values $\boldsymbol{\theta}^*$ found for the design variables values $\boldsymbol{x}_i$ can be different from that for $\boldsymbol{x}_j$, with $\boldsymbol{i} \neq \boldsymbol{j}$. In many practical cases, such as the one investigated in this work, even though the experimental campaign is exhaustive, enough cases of $\boldsymbol{x}$ to train a statistical model for $\boldsymbol{\delta}(\boldsymbol{x})$ are not available: therefore, a simplified model updating equation is adopted as follows:

$$\boldsymbol{y^E}(\boldsymbol{x}) = \boldsymbol{y^M}(\boldsymbol{x}, \boldsymbol{\theta}^*) + \boldsymbol{\epsilon} \tag{4}$$

2.1 Bayesian formulation of the IUQ problem

The purpose of IUQ is to quantify $p(\boldsymbol{\theta}^*|\boldsymbol{y^E})$, which can be obtained by Bayes rule:

$$p(\boldsymbol{\theta}^*|\boldsymbol{y^E}) = \frac{p(\boldsymbol{y^E}|\boldsymbol{\theta}^*)p(\boldsymbol{\theta}^*)}{\int p(\boldsymbol{y^E}|\boldsymbol{\theta}^*)p(\boldsymbol{\theta}^*)d\boldsymbol{\theta}^*} \tag{5}$$

where $p(\boldsymbol{\theta}^*)$ is the prior Probability Density Function (PDF) and $p(\boldsymbol{\theta}^*|\boldsymbol{y^E})$ is the posterior PDF of the vector of calibration parameters $\boldsymbol{\theta}^*$ (i.e., the PDFs before and after the experimental data are observed, respectively). The likelihood function $p(\boldsymbol{y^E}|\boldsymbol{\theta}^*)$ is the probability of observing the experimental data $\boldsymbol{y^E}$ given a particular value of $\boldsymbol{\theta}^*$. Notice that in the Bayes rule (5), an equal "weighting" is implicitly assigned to the likelihood function and the prior distributions to obtain the posterior probability density function of the uncertain parameters.

In general terms, if we assume $\boldsymbol{y}^{\boldsymbol{E}}(\boldsymbol{x}) = [y_1^E(\boldsymbol{x}), \dots, y_p^E(\boldsymbol{x})]$ and $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{x}, \boldsymbol{\theta}) = [y_1^M(\boldsymbol{x}, \boldsymbol{\theta}), \dots, y_p^M(\boldsymbol{x}, \boldsymbol{\theta})]$ to be p-dimensional vectors, and the measurement error to be zero-mean Gaussian-distributed (i.e., $\boldsymbol{\epsilon} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{exp})$), from equation (4), we get:

$$p(\boldsymbol{y}^E|\boldsymbol{\theta}^*) = \prod_{i=1}^{N_{exp}} \frac{1}{\left(\sqrt{2\pi}\right)^p \sqrt{|\boldsymbol{\Sigma}_{exp}|}} \exp\left[-\frac{1}{2}[\boldsymbol{y}^{\boldsymbol{E}}(\boldsymbol{x}_i) - \boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{x}_i, \boldsymbol{\theta}^*)]^T \left(\boldsymbol{I}\sigma_{exp}^2\right)^{-1} [\boldsymbol{y}^{\boldsymbol{E}}(\boldsymbol{x}_i) - \boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{x}_i, \boldsymbol{\theta}^*)]\right] \tag{6}$$

where N_{exp} is the number of independent experimental measurements and $\boldsymbol{I}\sigma_{exp}^2$ is the $p \times p$ covariance matrix for the measurement error. It should be noted that each experimental measurement corresponds to a particular vector of values of the design variable of $\boldsymbol{x}_{\boldsymbol{i}}, i = 1, \dots, N_{exp}$.

The denominator of the right-hand side of expression (5) is usually analytically intractable (Gelman et al., 2015). A Markov Chain Monte Carlo (MCMC) algorithm can be adopted (Gelman et al., 2015) to address this problem, as it allows sampling from distributions known only up to a normalization constant (i.e., $\int p\left(\boldsymbol{y}^{\boldsymbol{E}}\middle|\boldsymbol{\theta}^*\right)p(\boldsymbol{\theta}^*)d\boldsymbol{\theta}^*$ in our case). The MCMC samples can, then, be used to reconstruct the posterior PDF. The main drawback of MCMC is that it requires a considerable number (sometimes hundreds of thousands) of code executions (i.e., $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{x}, \boldsymbol{\theta})$) to evaluate $p\left(\boldsymbol{\theta}^*\middle|\boldsymbol{y}^{\boldsymbol{E}}\right)$, each of which can be computationally intensive. To address this issue, Kriging meta-modeling is a solution very popular by adopted IUQ process, since it provides also an estimation of the meta-model uncertainty. If a Kriging meta-model $\boldsymbol{y}^{\boldsymbol{MM}}(\boldsymbol{x}, \boldsymbol{\theta})$ is used to emulate $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{x}, \boldsymbol{\theta})$ (i.e., $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{x}, \boldsymbol{\theta}) \cong \boldsymbol{y}^{\boldsymbol{MM}}(\boldsymbol{x}, \boldsymbol{\theta})$), then $\boldsymbol{y}^{\boldsymbol{MM}}(\boldsymbol{x}, \boldsymbol{\theta})$ turns out to be Gaussian distributed. Assuming that $\boldsymbol{y}^{\boldsymbol{MM}}(\boldsymbol{x}, \boldsymbol{\theta})$ has mean value $\widehat{\boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{\theta})$ and covariance matrix $\boldsymbol{\Sigma}_{MM}$ (i.e., the matrix that provides the estimation of the meta-model uncertainty), the posterior PDF becomes:

$$p(\boldsymbol{\theta}^*|\boldsymbol{y}^E) \propto p(\boldsymbol{\theta}^*) \cdot \prod_{i=1}^{N_{exp}} \frac{1}{\left(\sqrt{2\pi}\right)^p \sqrt{|\boldsymbol{\Sigma}|}} \exp\left[-\frac{1}{2}[\boldsymbol{y}^{\boldsymbol{E}}(\boldsymbol{x}_i) - \widehat{\boldsymbol{y}}(\boldsymbol{x}_i, \boldsymbol{\theta}^*)]^T \boldsymbol{\Sigma}^{-1} [\boldsymbol{y}^{\boldsymbol{E}}(\boldsymbol{x}_i) - \widehat{\boldsymbol{y}}(\boldsymbol{x}_i, \boldsymbol{\theta}^*)]\right] \tag{7}$$

where $\boldsymbol{\Sigma} = \boldsymbol{I}\sigma_{exp}^2 + \boldsymbol{\Sigma}_{MM}$ is the likelihood covariance matrix that is given by the sum of the covariance matrix for the measurement error $\boldsymbol{I}\sigma_{exp}^2$ and the covariance matrix for the meta-model uncertainty.

3 Meta-Modeling

A Meta-Model (MM) is a functional approximation of the input/output relations of the original model that is faster to be evaluated. It is constructed starting from a set of input parameters and their corresponding model responses, obtained by running the original time-demanding model. In this work, we adopt Kriging MM because it provides a direct estimation of the meta-model uncertainty. Kriging is a stochastic algorithm which assumes that the model output $z(\boldsymbol{\theta})$ is a realization of a Gaussian process indexed by $\boldsymbol{\theta} \in \boldsymbol{D}_{\boldsymbol{\theta}} \subset \mathbb{R}^d$. The mathematical form of a Kriging is given by:

$$z(\boldsymbol{\theta}) = \sum_{j=1}^{n} \beta_j f_j(\boldsymbol{\theta}) + G(\boldsymbol{\theta}) = \boldsymbol{\beta}^{\boldsymbol{T}} \boldsymbol{f}(\boldsymbol{\theta}) + G(\boldsymbol{\theta}) \tag{8}$$

where $\boldsymbol{\beta}^{\boldsymbol{T}} \boldsymbol{f}(\boldsymbol{\theta})$, is the *mean* value of the Gaussian process, also called the “trend”; it is given by n arbitrary functions $[f_1(\boldsymbol{\theta}), \dots, f_n(\boldsymbol{\theta})]$ and the related coefficients $[\beta_1, \dots, \beta_n]$. The second term, $G(\boldsymbol{\theta})$, is a Gaussian process with zero mean and covariance:

$$Cov\left[G\left(\boldsymbol{\theta}^{(i)}\right), G\left(\boldsymbol{\theta}^{(j)}\right)\right] = \sigma^2 R\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) \tag{9}$$

where $\sigma^2$is the *process variance* and $R(\cdot,\cdot)$ is the *correlation function* (also called *correlation kernel*), defined for any two points in the input domain $\boldsymbol{D}_{\boldsymbol{\theta}}$, that is a function of the distance $h(\cdot,\cdot)$:

$$R\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) = R\left(h(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)})\right) \tag{10}$$

with :

$$h\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) = \left[\sum_{k=1}^{d} \left(\frac{\theta_k^{(i)} - \theta_k^{(j)}}{\omega_k}\right)^2\right]^{0.5} \tag{11}$$

where the parameters $\boldsymbol{\omega} = [\omega_1, \dots, \omega_d]$ are called *length scale parameters*. Estimation methods such as M*aximum likelihood estimation* (ML) or *Cross-Validation estimation* (CV) are usually applied to estimate $\boldsymbol{\beta}, \sigma^2$ and $\boldsymbol{\omega}$ (Bachoc, 2013). In this work, we use the software package UQLab (Lataniotis et al., 2019b) to implement the Kriging meta-modeling. Further details on Kriging meta-modeling are given in Appendix A.

4 Case Study

The proposed IUQ approach is applied to a TH input model for the RELAP5-3D code (Idaho National Laboratory, 2015) developed by Politecnico di Torino (Bersano et al., 2020) for the PERSEO (In-Pool Energy Removal System for Emergency Operation) facility. The PERSEO equipment is not designed to simulate a specific passive system of a reactor, but rather to assess the performance and the efficiency of a new in-pool heat exchanger for decay heat removal by natural circulation (Mascari et al., 2019). Experimental data from PERSEO test 7 have been used for IUQ. PERSEO test 7 experimental data have been originally distributed to Politecnico di Torino by ENEA for the participation at the PERSEO benchmark, conducted in the framework of OECD/NEA/CSNI/WGAMA “Status report on thermal-hydraulic passive systems design and safety assessment” (NEA/CSNI/R(2021)2, 2021). The PERSEO test facility, located at SIET laboratories, is sketched in Figure 1. The system is made of a primary side and a pool side. The primary side consists of (Bersano et al., 2020; Ferri et al., 2005; Mascari et al., 2019):

- The pressure vessel (PV) ($43\ m^3$ volume and $13\ m$ height);
- The steam line from the PV to the heat exchanger (HX) upper header;
- The condensate line from the HX lower header to the PV;
- The HX (composed of two cylindrical headers and 120 vertical straight pipes).

The pool side includes (Bandini et al., 2011; Ferri et al., 2005; Mascari et al., 2019):

- The Heat Exchanger Pool (HXP) ($29\ m^3$ volume and $5.7\ m$ height);
- The Overall Pool (OP) ($173\ m^3$ volume and $5.8\ m$ height);
- The line connecting the OP and the HXP with the triggering valve (TV);
- The steam duct connecting the two pools.

The PV is maintained in saturation conditions with steam from a nearby power plant. The system is activated by opening the TV, which causes the flooding of the HXP and the steam condensation inside the HX tubes. In this way, the power is transferred from the primary side to the pool side. The OP represents the water reservoir of the system. When the HXP water starts boiling, the steam produced in the HXP is driven to the OP through an injector that, conveying the steam below the water level, promotes the direct condensation. The condensation of steam inside the OP increases the OP water temperature up to the boiling point. The steam produced in the OP flows outside through the boil-off pipe.

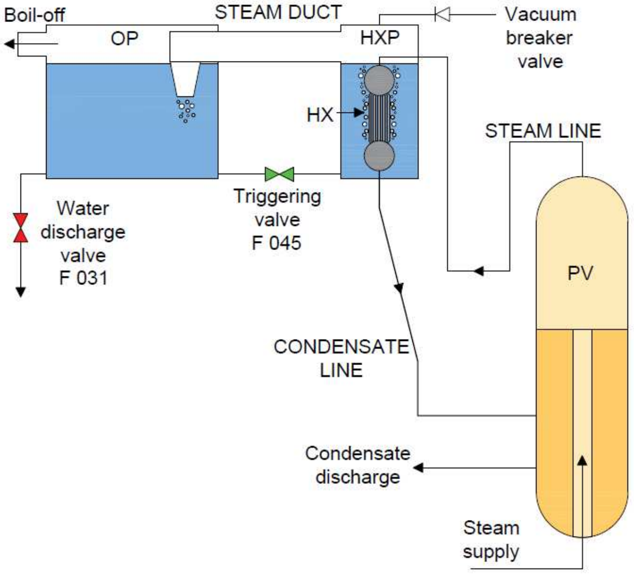


Figure 1. Scheme of the PERSEO facility (Bersano et al., 2020).

The PERSEO experimental campaign consists of nine tests (Ferri et al., 2005; Mascari et al., 2019). Test 7 consists of a first part aimed at verifying the behavior of the system with two different water levels and a second part aimed at characterizing the long-term cooling capability of the system (Mascari et al., 2019). For our analysis, the measured experimental data of Test 7 part 2 (Mascari et al., 2019) have been used. With reference to the classification of input parameters made in Section 2, the vector of the design variables $\boldsymbol{x}$ contains ICs, BCs and the set of scenario-description inputs (e.g., TV opening/closure) (Bandini et al., 2011; Mascari et al., 2019). Concerning the calibration variables $\boldsymbol{\theta}$, the RELAP5-3D code contains various model parameters. In the present analysis, only some are considered for the IUQ, as the objective of the present study is not to carry out a complete uncertainty analysis but rather only to show how the IUQ is performed in the proposed framework.

Uniform priors with wide ranges are used so as to not inject any preference to some input values over others. Table 1 reports the upper and lower bound of the eight calibration parameters selected for the IUQ (normalized with respect to their prior nominal values), being these expected to play a crucial role in natural circulation (IAEA, 2012; Marque`s et al., 2002; Pierro et al., 2009).

Table 1. Uncertain input parameters selected for the RELAP5-3D model

$\boldsymbol{\theta_i}$	Parameter (multiplication factor)	Parameter Name	Lower bound	Upper bound
$\boldsymbol{\theta_1}$	Inner fouling factor	Inner_FF	0.5	1.5
$\boldsymbol{\theta_2}$	Outer fouling factor	Outer_FF	1.0	1.5
$\boldsymbol{\theta_3}$	Injector K factor	K_injector	0.5	1.5
$\boldsymbol{\theta_4}$	Sum of the steam line's K factors	K_sum_steam	0.5	1.5
$\boldsymbol{\theta_5}$	Sum of condensate line's K factors	K_sum_condensate	0.5	1.5
$\boldsymbol{\theta_6}$	Diaphragm K factor	K_ diaphragm	0.5	1.5
$\boldsymbol{\theta_7}$	Rockwool thermal conductivity	k_rockwool	1.0	1.5
$\boldsymbol{\theta_8}$	HXP first pipe flow area	A_effective	0.5	1.5

K_injector is the localized pressure drop coefficient assumed in the model at the injector. K_sum_steam is the weighted sum of the localized pressure drop coefficients assumed in the model on the steam line. The weighted sum is such that the sum of localized pressure drops of the steam line is conserved:

$$\Delta p_{tot,loc} = \sum_i k_i \frac{1}{2}\rho v_i^2 = \frac{W^2}{2\rho}\sum_i \frac{k_i}{A_i^2} = \frac{W^2}{2\rho A_{ref}^2}\left(\sum_i \frac{k_i A_{ref}^2}{A_i^2}\right) \tag{12}$$

where A_{ref} is the area of the junction adopted as reference for *K_sum_steam*, A_i is the i^{th} junction area of the steam line, ρ is the average steam density in the steam line, W is the mass flow rate, k_i is the nominal value of the i^{th} junction's K factor. It follows that the sum of the steam line's K factors is:

$$K_{steam\ line} = \sum_i \frac{k_i A_{ref}^2}{A_i^2} \tag{13}$$

K_sum_condensate has been obtained with the same procedure. *K_diaphragm* is the localized pressure drop coefficient assumed in the model for the diaphragm located in the HX pool (Bassenghi, 2013). Concerning the Inner_FF and Outer_FF, they are not fouling factors as usually defined to account for the fouling thermal resistances of the tubes; indeed, in RELAP5-3D they are (dimensionless) multiplicative factors applied to the heat transfer coefficient of the inner (Inner_FF) and the outer (Outer_FF) side of the HX tubes, respectively. Both the Inner_FF and Outer_FF are selected as calibration parameters to avoid underestimation of the heat transfer from the HX to the surrounding pool (Bersano et al., 2020). Pressure drops are taken into account by including K_injector, K_sum_steam, K_sum_condensate and K_diaphragm into the calibration parameters. Heat losses have been considered by including k_rockwool (i.e., rockwool thermal conductivity adopted as an insulator for some of the steam line's pipes), and geometrical nodalization by including A_effective (i.e., the flow area of pipe 130 (thermally connected to the HX) that is not a fixed geometrical quantity, unlike the sum of pipe 130 and pipe 140 flow areas (that is equal to the HXP cross-section)). For more details about the facility's nodalization, refer to Figure 2, that shows the RELAP5-3D nodalization of the PERSEO facility adopted in the present work, and (Bersano et al., 2020, 2019).

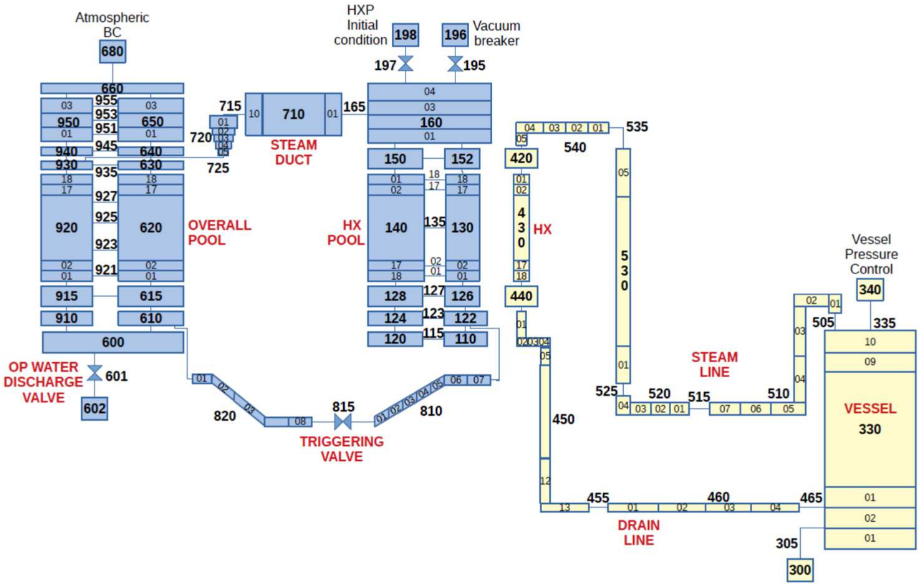


Figure 2. RELAP5-3D nodalization of the PERSEO facility (Bersano et al., 2020).

For each parameter, the prior PDF is set as uniform: $p(\theta) = \frac{1}{U-L}$ for $\theta \in [L, U]$ and $p(\theta) = 0$ otherwise, where U and L are the upper and lower bounds, respectively (listed in Table1).

The range of variation of each parameter, except for the Outer_FF and k_rockwool, varies from -50% to $+50\%$ of the parameter nominal value. This is done to consider a relatively large range. The Outer_FF lower bound is set equal to the nominal value to avoid underestimation of the heat transfer from the HX to the pool (Bersano et al., 2020). This is not the case for the Inner_FF, since the nominal value that reproduces at best the experimental data of Test 7 is already set higher than 1 (3.54) following the results of (Bersano et al., 2020). Concerning k_rockwool, the upper and lower bounds are set considering the thermal conductivity values for the range of temperature experienced by rockwool during this Test.

Figure 3 shows a comparison between the RELAP5-3D HX exchanged power *(Test* 7 part 2) computed obtained using: (1) the nominal value of *Inner_FF* (i.e., that one found in (Bersano et al., 2020)) and (2) the RELAP5-3D default value for the *Inner_FF* (i.e., 1.0); in both cases, the other parameters are set to their nominal values. One can notice that in the case of *Inner_FF*$= 1.0$, the code underestimates the HX exchanged power. As pointed out in (Bersano et al., 2020, 2019), a possible explanation can be that some correlations adopted in RELAP5-3D for condensation are applied in conditions outside their validity ranges. The introduction of $\delta(x)$ in equation (3) could partially account for such model discrepancy, whereas, neglecting $\delta(x)$, the model discrepancy is “dumped” in the uncertainty of $\boldsymbol{\theta}^*$. In general, a direct application of the results obtained in the IUQ process to a full-scale prototype must be carefully evaluated. Since PERSEO is a full-scale Separate Effect Test Facility (SETF) (Bestion et al., 2017; Mascari et al., 2015), thus, the PDFs obtained in the IUQ process using the experimental data can give some realistic insights of a full-scale prototype that includes such a passive heat removal system.

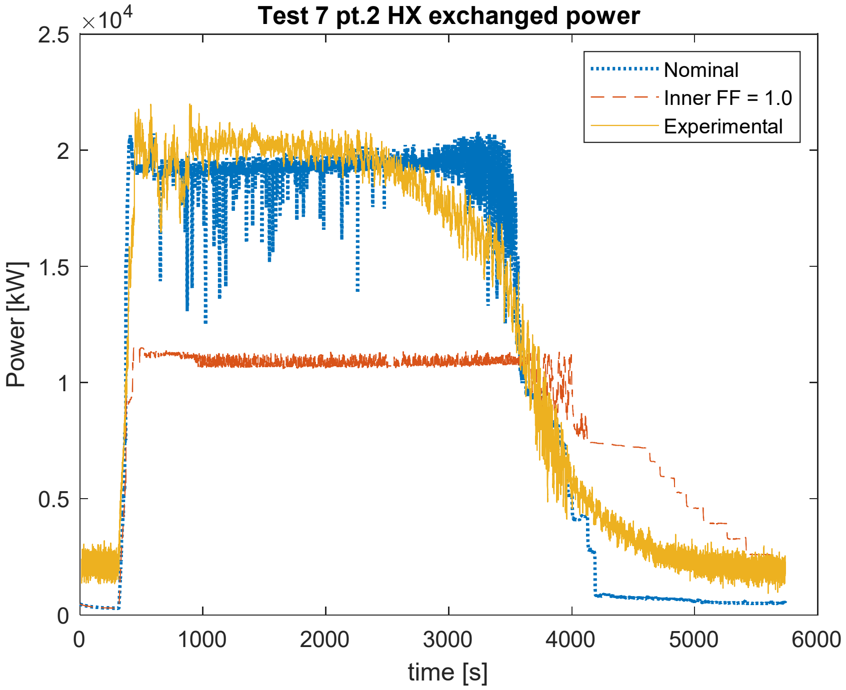


Figure 3. HX exchanged power (Test 7, part 2). In solid yellow line, the experimental measurement; in dotted blue line, the RELAP5-3D simulation for the nominal values; in dashed orange line, the RELAP5-3D simulation in case the Inner_FF is set to the RELAP5-3D default value (i.e., 1.0).

The HX exchanged power is selected as experimental data to perform the IUQ, since it is one of the most representative outputs of the transient. Indeed, the HX exchanged power level is affected by many thermal-hydraulic phenomena taking place during the Test, such as (1) the heat exchange between the HX tubes and the HXP, and (2) the HXP water boiling (with consequent HXP water level decrease). Thus, the HX exchanged power conveys a considerable amount of information about the transient. It is worth mentioning that the proposed methodology is quite general so that if a different parameter were selected as FOM, only some fine-tuning would be needed of the input parameters (such as those reported in Table 1), whereas no major modification would be needed to the overall approach. Since we consider only a single experiment (i.e., Test 7 part 2), we absorb the dependence on $\boldsymbol{x}$ of the forward model $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{x},\boldsymbol{\theta})$ into the definition of $\tilde{\boldsymbol{y}}^{\boldsymbol{M}} = \tilde{\boldsymbol{y}}^{\boldsymbol{M}}(\boldsymbol{\theta})$, that represents the RELAP5-3D model. Simulation performed using a dual-core processor with a clock speed of 2.7 GHz and 3.5 GHz in normal and TurboBoost modes, respectively, takes around 2 hours.

5 Dimensionality reduction

The HX exchanged power is a vectorial quantity (i.e., a time-series), which introduces somewhat challenging additional complexity in the use of Kriging to approximate the RELAP5-3D results. Indeed, the vectorial output entries are usually highly correlated; thus, building a separate independent Kriging for each entry would be computationally unfeasible and, in principle, wrong. Two approaches can be adopted to tackle this problem (McFarland et al., 2008):

1. performing a dimensionality reduction (e.g., Principal Component Analysis, PCA) to extract from the vectorial output relevant features that are used to represent the entire output. In this case, a possible solution would be to build a Kriging independent model for each feature;
2. considering the variables that index the output spectrum (e.g., the time in our case) as additional labeling inputs. In this case, the output can be treated as a scalar quantity. This approach may introduce relevant difficulties in the case of very long time-series. For instance, if the RELAP5-3D code outputs the response quantity at 1000 time instances, considering a DOE based on 100 samples and treating time as an additional input, the total number of training points for the Kriging would be $1000 \cdot 100 = 100000$, which is too cumbersome.

In the present work, the first option is adopted by implementing PCA for the dimensionality reduction of the output. PCA is a space transformation from a high dimensional space to a lower-dimensional space such that transformed variables are uncorrelated and retain as much as possible of the variation present in the data set (Jolliffe, 2002). To carry out the PCA, Singular Value Decomposition (SVD) is adopted (Wall et al., 2003). Further details on PCA are given in Appendix B.

6 Results

The entire procedure for the IUQ can be subdivided into two main steps:

1. Surrogate modeling;
2. Bayesian inference (MCMC sampling).

6.1 Surrogate Modeling

Let $\boldsymbol{\Theta} = [\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m)}]$ be the Design Of Experiments (DOE) and $\widetilde{\boldsymbol{Y}} = [\widetilde{\boldsymbol{y}}^{(1)}, \dots, \widetilde{\boldsymbol{y}}^{(m)}]$ the $p \times m$ matrix that contains the m corresponding p-dimensional BE model responses. PCA is applied to $\widetilde{\boldsymbol{Y}}$ that is transformed into a p^*-dimensional ($p^* \ll p$)obtaining the $p^* \times m$ features matrix $\widetilde{\boldsymbol{Z}} = \left[\tilde{\boldsymbol{z}}^{(1)}, \dots, \tilde{\boldsymbol{z}}^{(m)}\right]$, whose rows are called *PC scores* and columns represent, in this case, the RELAP5-3D results projected into the features space. An independent MM (i.e., one for each feature) is trained to emulate each feature using the m input-output training patterns (i.e., $\boldsymbol{\Theta} = [\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m)}]$ and the respective m transformed model responses $\left[z_j^{(1)}, \dots, z_j^{(m)}\right]$, with $j = 1,2, \dots, p^*$). The Kriging prediction for a specific input $\boldsymbol{\theta}^{(i)}$ will be the $p^* \times 1$ vector $\hat{\boldsymbol{z}}$. To antitransform a generic vector $\boldsymbol{z}$ from the features space to the original space:

$$\breve{\boldsymbol{y}} = \overline{\boldsymbol{\mu}}_{\tilde{Y}} + \Phi^T \boldsymbol{z} \tag{14}$$

where $\overline{\boldsymbol{\mu}}_{\tilde{Y}}$ is the simulations means vector and Φ is the PCA transformation matrix (see expression (B.2) and (B.6) of Appendix B).

This solution usually performs as good as using a fully multivariate meta-model (Wilkinson, 2010), even though the PC scores are not independent. Figure 4 shows a schematic diagram of the meta-modeling approach adopted.

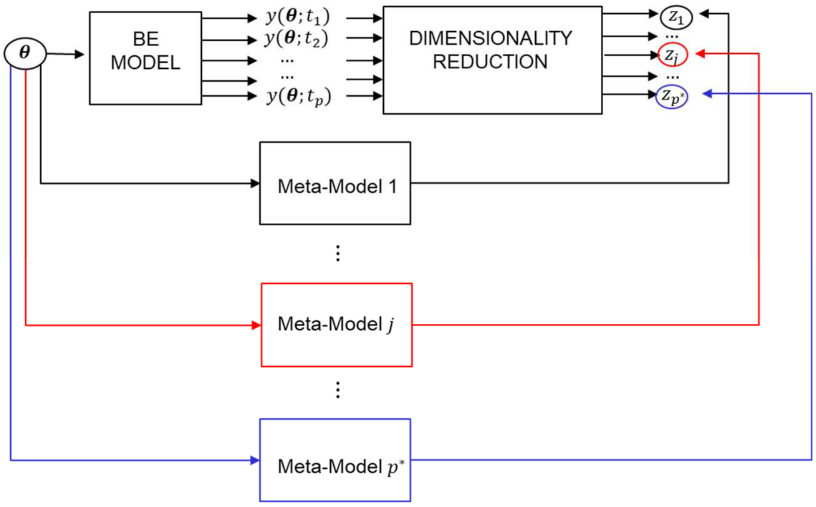


Figure 4. A schematic diagram of the Meta-Model approach adopted.

Kriging, unlike other meta-modeling techniques, provides the uncertainty associated with each prediction, making such type of MMs particularly suitable for the Bayesian IUQ, since this uncertainty directly enters into the likelihood formulation (see Eq. (17) and (18) in Section 6.2).

The very first step we have to deal with is the selection of the DOE points $\boldsymbol{\Theta} = [\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m)}]$. LHS is adopted to build $\boldsymbol{\Theta}$ according to the prior distributions reported in Table 1. To analyze the effect of m on the Kriging performances, we decide to work with two different DOEs, the first one (i.e., $\boldsymbol{\Theta}_1$) characterized by $m_1 = 90$ points, the second one (i.e., $\boldsymbol{\Theta}_2$) characterized by $m_2 = 180$ points[1]. Both $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$ are simulated through the RELAP5-3D code. Let $\boldsymbol{\Theta}_1 = [\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m_1)}]$ and $\boldsymbol{\Theta}_2 = [\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m_2)}]$ be the ensembles of design points and $\tilde{\boldsymbol{Y}}_1 = [\tilde{\boldsymbol{y}}^{(1)}, \dots, \tilde{\boldsymbol{y}}^{(m_1)}]$ and $\tilde{\boldsymbol{Y}}_2 = [\tilde{\boldsymbol{y}}^{(1)}, \dots, \tilde{\boldsymbol{y}}^{(m_2)}]$ their respective RELAP5-

---

[1] Actually, the m_2 design points are obtained adding 90 LHS-sampled points to the first experimental design $\boldsymbol{\Theta}_1$.

3D simulated output. Both $\widetilde{\boldsymbol{Y}}_1$ and $\widetilde{\boldsymbol{Y}}_2$ are pre-processed by applying a moving median filter (by the "*movmedian*" MATLAB function) to reduce the noise by which they are affected. Once $\widetilde{\boldsymbol{Y}}_1$ and $\widetilde{\boldsymbol{Y}}_2$ are pre-processed, PCA is applied. The percentage of variance explained (i.e., $\frac{\sum_{j=1}^{p^*} \lambda_j}{\sum_{j=1}^{p} \lambda_j} \cdot 100$) is set equal to $95\ \%$. According to (Wilkinson, 2010), even though PCA is a linear technique for dimensionality reduction, it can be used for nonlinear models. Indeed, what makes PCA a linear approach is the linear projection into the reduced space, whereas the mapping from the input space $\Theta \subset \mathbb{R}^d$ to the features space $Z \subset \mathbb{R}^{p^*}$ is performed through the Kriging MMs and, thus, can be nonlinear. Figure 5 displays the cumulative percentage of variation explained by the first 50 PCs. We find that $p^* = 4$ guarantees a percentage of variation explained at least equal to 95% for both $\widetilde{\boldsymbol{Y}}_1$ and $\widetilde{\boldsymbol{Y}}_2$. Choosing $p^* > 4$ would decrease the Kriging's accuracy for a higher order for PCs, as shown later.

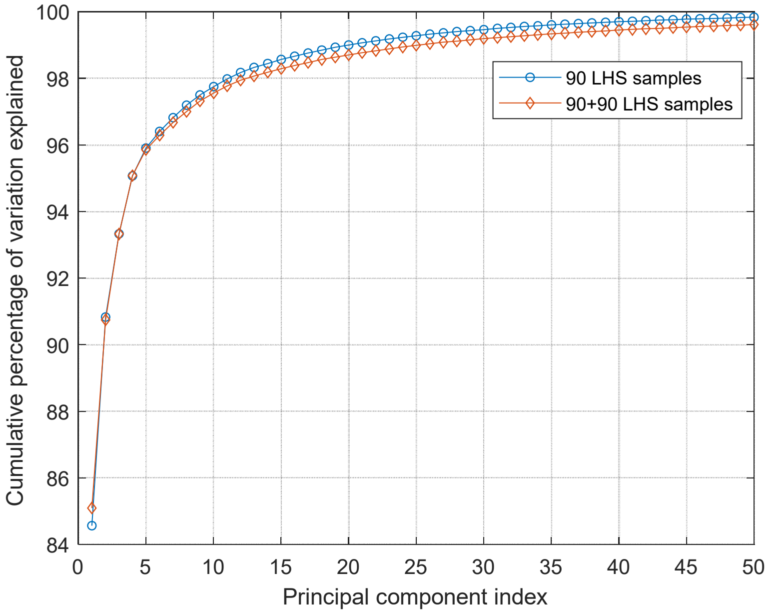


Figure 5. Cumulative percentage of variation explained by the first fifty PC dimensions.

In the current research, for each principal component, we study which set of Kriging options (i.e., which combination of trend type, correlation family, and estimation method) provides the best predictive performances. We explore:

- Three different trend types: Constant, Linear, Quadratic;
- Four different correlation families types: Exponential, Gaussian, Matérn-3/2, Matérn-5/2;
- Two different estimation methods types: CV and ML;

for a total of twenty-four different combinations (i.e., $3 \times 4 \times 2$). In Table 2, each combination of Kriging options is labeled with a number (i.e., $1,2, \dots ,24$). For each combination $R^2_{prediction}$ and AIC_c are computed to assess Kriging performance. For further details on Kriging options, $R^2_{prediction}$ and AIC_c, refer to Appendix A. Figure 6 and Figure 7 show, for each PCs, the plots of $R^2_{prediction}$ vs AIC_c in the case of $m = 90$ and $m = 180$, respectively. The Kriging options leading to the best-performances are those that occupy the top-left region of these plots; indeed, we want to maximize $R^2_{prediction}$ and to minimize AIC_c. Thus, the search for the best Kriging option reduces to a multi-objective optimization problem. For $m = 90$, combination 13 (see Table 2) seems to be non-dominated by other solutions (i.e., it is a Pareto optimal solution) for all PCs. With $m = 180$, combination 9 results to be non-dominated for the first three PCs; instead, for the fourth principal component, one of the best options is given by combination 11. In Figure 8, the performances of the 24 Kriging combinations are compared for $m = 90$ and $m = 180$, using only the $R^2_{prediction}$ factor. On the y-axis the $R^2_{prediction}$ factor is plotted, whereas the x-axis reports the identification number of each combination.

It is worth noticing that:

- the accuracy decreases for PCs of higher-order; thus, we decide to take into account only the first $p^* = 4$ PCs (i.e., the number that guarantees a cumulative 95% of variation explained) since for PCs of higher-order $R^2_{prediction}$ would be too low;
- as expected, the accuracy for $m = 180$ is generally higher than that for $m = 90$.

Given the observations above, the following Kriging MMs are adopted for each principal component to carry out the MCMC sampling:

- PC1, PC2, PC3: Kriging combination 9, trained with $m = 180$;
- PC4: Kriging combination 11 trained with $m = 180$.

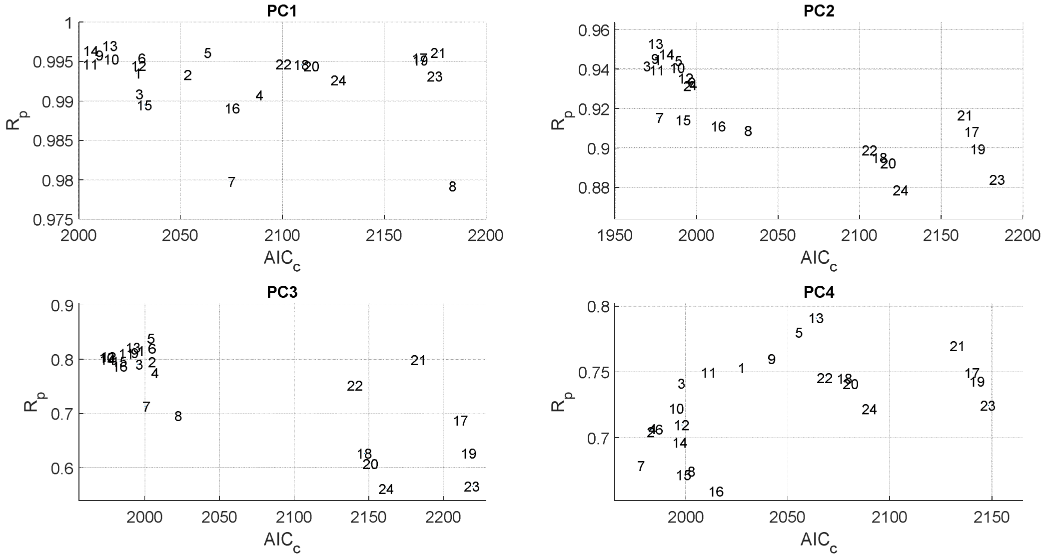


Figure 6. Plots of $R^2_{prediction}$ vs AIC_c factors of the 24 Kriging option sets (see Table 2) for the four PCs ($m = 90$).

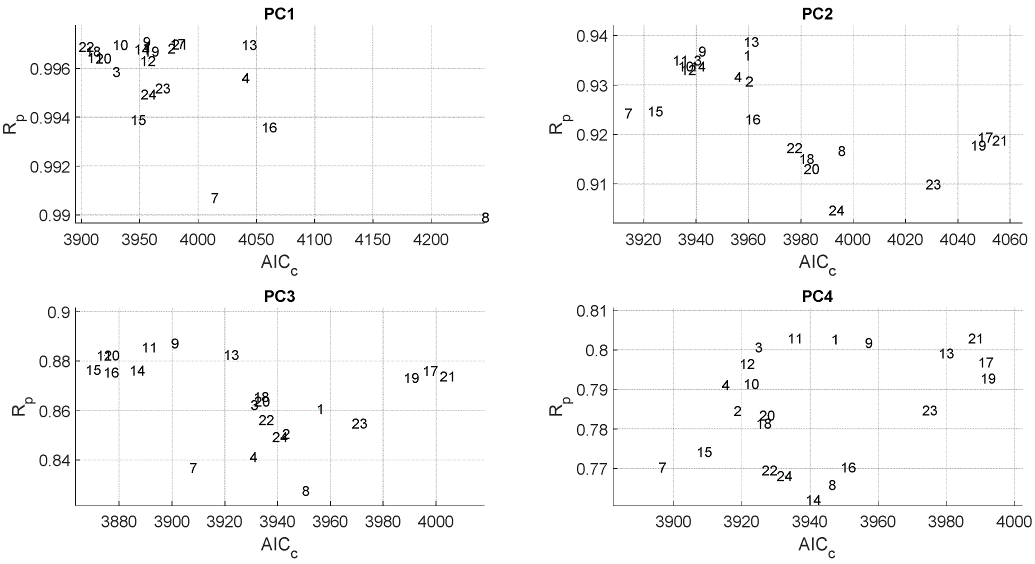


Figure 7. Plots of $R^2_{prediction}$vs AIC_c factors of the 24 Kriging option sets (see Table 2) for the four PCs ($m = 180$).

Table 2. Correspondence between combination numbers and the respective Kriging options

Combination	Trend type	Correlation kernel	Estimation method	Combination	Trend type	Correlation kernel	Estimation method
1	Constant	Matérn-5/2	CV	13	Linear	Gaussian	CV
2	Constant	Matérn-5/2	ML	14	Linear	Gaussian	ML
3	Constant	Matérn-3/2	CV	15	Linear	Exponential	CV
4	Constant	Matérn-3/2	ML	16	Linear	Exponential	ML
5	Constant	Gaussian	CV	17	Quadratic	Matérn-5/2	CV
6	Constant	Gaussian	ML	18	Quadratic	Matérn-5/2	ML
7	Constant	Exponential	CV	19	Quadratic	Matérn-3/2	CV
8	Constant	Exponential	ML	20	Quadratic	Matérn-3/2	ML
9	Linear	Matérn-5/2	CV	21	Quadratic	Gaussian	CV
10	Linear	Matérn-5/2	ML	22	Quadratic	Gaussian	ML
11	Linear	Matérn-3/2	CV	23	Quadratic	Exponential	CV
12	Linear	Matérn-3/2	ML	24	Quadratic	Exponential	ML

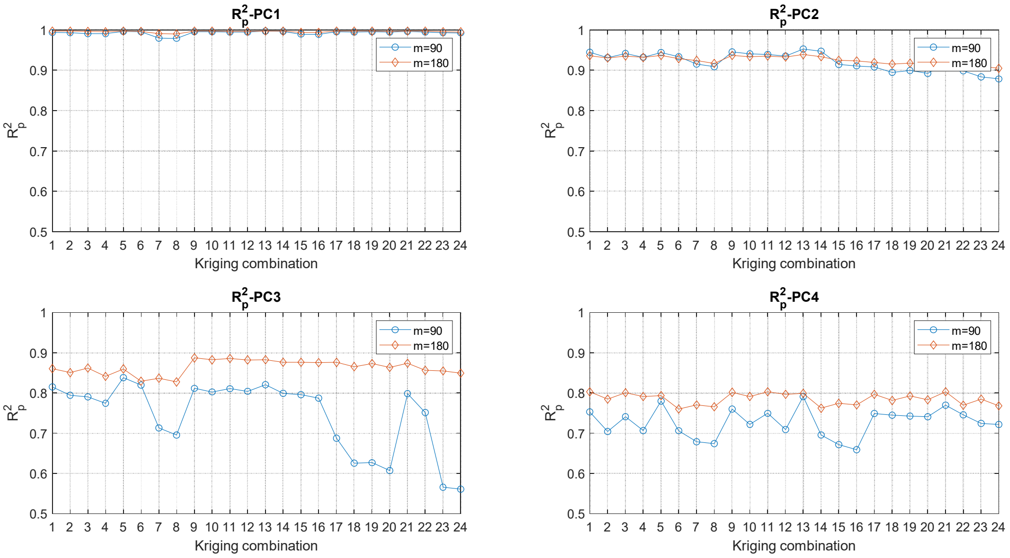


Figure 8. $\boldsymbol{R^2_{prediction}}$ values of different combinations (see Table 2) for each principal component.

6.2 Bayesian inference (MCMC sampling)

Before using the experimental time-series $\boldsymbol{y^E}$ for the IUQ process, data smoothing is performed since experimental data are noisy. Thus, we fit the measured HX exchanged power using the MATLAB function "Smoothing Splines," and we project $\boldsymbol{y^E}$ onto the feature space through:

$$\boldsymbol{z}^E = \boldsymbol{\Phi}\left(\boldsymbol{y^E} - \bar{\boldsymbol{\mu}}_{\tilde{\boldsymbol{Y}}}\right) = \left[z_1^E, \dots z_{p^*}^E\right]^T \tag{15}$$

where $z_1^E, \dots z_{p^*}^E$ are the PC scores of the experimental response, $\bar{\boldsymbol{\mu}}_{\tilde{\boldsymbol{Y}}}$ is the simulations means vector (see expression (B.2)) and $\boldsymbol{\Phi}$ is the PCA transformation matrix. In Figure 9, except for the initial and the last part of the transient, a relatively good agreement between $\boldsymbol{y^E}$ and the reconstructed experimental time-series $\breve{\boldsymbol{y}}^E$ can be appreciated:

$$\breve{\boldsymbol{y}}^E = \bar{\boldsymbol{\mu}}_{\tilde{\boldsymbol{Y}}} + \boldsymbol{\Phi}^T \, \boldsymbol{z^E} \tag{16}$$

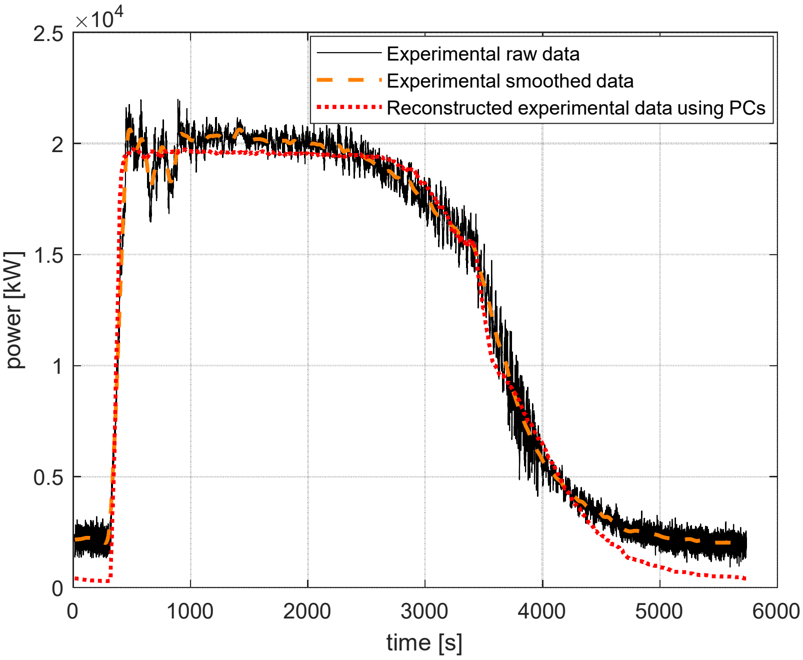


Figure 9. HX exchanged power: a comparison between experimental raw data, experimental smoothed data (i.e., $\boldsymbol{y}^{\boldsymbol{E}}$) and reconstructed experimental data $\breve{\boldsymbol{y}}^{\boldsymbol{E}}$.

The discrepancy between $\boldsymbol{y}^{\boldsymbol{E}}$ and $\breve{\boldsymbol{y}}^{E}$ (at both the beginning and the last part of the transient, where $\breve{\boldsymbol{y}}^{E}$ approaches zero) is due to the fact that the PCA transformation matrix $\boldsymbol{\Phi}$ is obtained using only RELAP5-3D time-series. These time-series, unlike $\boldsymbol{y}^{\boldsymbol{E}}$, approach zero at both the beginning and the last part of the transient. According to (Wu et al., 2018a), performing the IUQ in the original space (i.e., $\Upsilon \subset \mathbb{R}^{p}$), applying a back-transformation of each Kiging's prediction $\hat{\boldsymbol{z}}(\boldsymbol{\theta}^{*})$, may cause convergence issues in the MCMC sampling. Indeed, if for any combination of $\boldsymbol{\theta}$, $\breve{\boldsymbol{y}}^{\boldsymbol{M}}$(i.e., the reconstructed Kriging prediction) does not agree with $\boldsymbol{y}^{\boldsymbol{E}}$, MCMC sampling does not reach the equilibrium distribution. For this reason, in the present work $\boldsymbol{y}^{\boldsymbol{E}}$ is transformed into the PC subspace (obtaining $\boldsymbol{z}^{E}$) and the IUQ is carried out in the PC subspace.

The prior $p(\boldsymbol{\theta}^{*})$ is a multivariate distribution of independent uniformly distributed variables. The posterior $p\left(\boldsymbol{\theta}^{*} \middle| \boldsymbol{z}^{\boldsymbol{E}}\right)$ is proportional to $p(\boldsymbol{\theta}^{*})$ multiplied by the likelihood $p\left(\boldsymbol{z}^{\boldsymbol{E}} \middle| \boldsymbol{\theta}^{*}\right)$. Since in our case $N_{exp} = 1$, the posterior expression reduces to:

$$p\left(\boldsymbol{\theta}^{*} \middle| \boldsymbol{z}^{\boldsymbol{E}}\right) \propto p(\boldsymbol{\theta}^{*}) \cdot \frac{1}{\left(\sqrt{2\pi}\right)^{p} \sqrt{|\boldsymbol{\Sigma}|}} \exp\left[-\frac{1}{2}\left[\boldsymbol{z}^{E} - \hat{\boldsymbol{z}}(\boldsymbol{\theta}^{*})\right]^{T} \boldsymbol{\Sigma}^{-1} \left[\boldsymbol{z}^{E} - \hat{\boldsymbol{z}}(\boldsymbol{\theta}^{*})\right]\right] \quad (17)$$

where $\hat{\boldsymbol{z}}(\boldsymbol{\theta}^{*})$ is the mean value of the Kriging prediction and $\boldsymbol{\Sigma}$ is the likelihood covariance matrix:

$$\boldsymbol{\Sigma} = \boldsymbol{\Phi} \boldsymbol{I} \sigma_{exp}^{2} \boldsymbol{\Phi}^{T} + \boldsymbol{\Sigma}_{\boldsymbol{Kriging}} \quad (18)$$

where $\boldsymbol{\Phi} \boldsymbol{I} \sigma_{exp}^{2} \boldsymbol{\Phi}^{T}$ represents the transformation of the measurement error covariance matrix from the original p −dimensional space onto the features p^{*} −dimensional space. In the present work, σ_{exp} is set equal to $500\,kW$ (Ferri et al., 2005). $\boldsymbol{\Sigma}_{\boldsymbol{Kriging}}$ is the covariance matrix associated with the Kriging prediction uncertainty and it is a $p^* \times p^*$ matrix having the mean square errors of each PC score prediction as diagonal entries:

$$\boldsymbol{\Sigma}_{\boldsymbol{Kriging}} = \begin{bmatrix} \sigma^2_{z_1(\theta^*)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma^2_{z_{p^*}(\theta^*)} \end{bmatrix} \tag{19}$$

An adaptive Metropolis algorithm (Haario et al., 2001) is applied to generate eight parallel chains with $5 \cdot 10^5$ iterations. We adopt the UQLab MATLAB package (Wagner et al., 2019) to implement the MCMC algorithm. It took almost 2.5 hours to compute the posterior using a dual-core processor having a clock speed of 2.7 GHz and 3.5 GHz in normal and TurboBoost modes, respectively. On the same processor, we can run at most four parallel RELAP5-3D simulations together, lasting two hours; thus, the MCMC sampling would take about 228 years. For parameter samples that go outside the bounds of a priori distribution during the MCMC simulation (that have a zero prior probability), the likelihood function is not evaluated. This is done for efficiency because the resulting realizations would be rejected in any case. We post-process the samples discarding from each chain the first half of the samples for burn-in. To assess the MCMC simulation convergence, we examine the trace plots, the Gelman-Rubin potential scale reduction factor (Gelman et al., 2015) (i.e., a quantitative approach used to assess MCMC convergence) reported in Table 3 and the autocorrelation functions in Figure 11. For clarity sake, in Figure 10, only the first 10^4 iterations of the trace plots (out of 500×10^3) are plotted: in all cases, chains have good mixing and stationary properties (i.e., cover a common distribution).

The Gelman-Rubin potential scale reduction factors are close to 1 for each θ, indicating that the chains have achieved convergence(Gelman et al., 2015). From the autocorrelation functions, it is evident that samples within each chain are correlated and, therefore, not independent. A common practice adopted to reduce the autocorrelation is *thinning*. It consists of keeping every k^{th} sample from each sequence. We do not perform thinning to compute the posterior distribution. Indeed, according to Gelman et al. (Gelman et al., 2015), even though we do not perform thinning, we can use samples for inference about θ^*if the chains have achieved convergence.

Table 3. Gelman-Rubin potential scale reduction factors.

$\boldsymbol{\theta_i^*}$	$\boldsymbol{\theta_1^*}$	$\boldsymbol{\theta_2^*}$	$\boldsymbol{\theta_3^*}$	$\boldsymbol{\theta_4^*}$	$\boldsymbol{\theta_5^*}$	$\boldsymbol{\theta_6^*}$	$\boldsymbol{\theta_7^*}$	$\boldsymbol{\theta_8^*}$
$\boldsymbol{\hat{R}}$	$\mathbf{1.0003}$	$\mathbf{1.0002}$	$\mathbf{1.0003}$	$\mathbf{1.0005}$	$\mathbf{1.0002}$	$\mathbf{1.0002}$	$\mathbf{1.0003}$	$\mathbf{1.0002}$

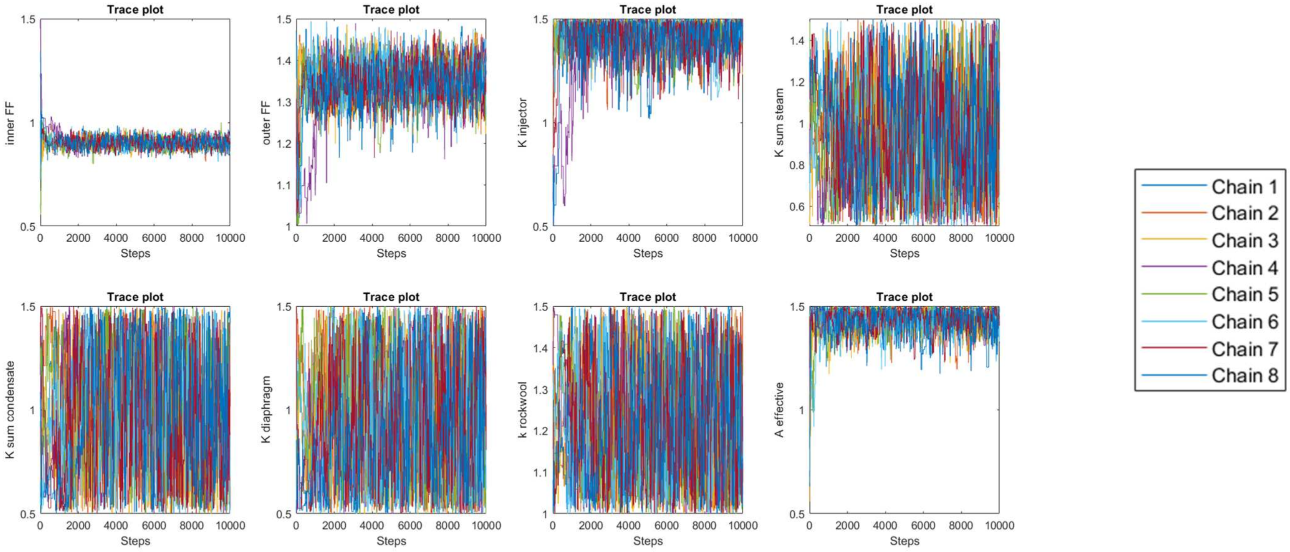


Figure 10. Trace plots for the first eight calibration parameters. To make the Figures more readable, only the first $\mathbf{10^4}$ iterations (out of $\mathbf{500} \times \mathbf{10^3}$) are shown. For the remaining MCMC iterations, trace plots result to be essentially unchanged.

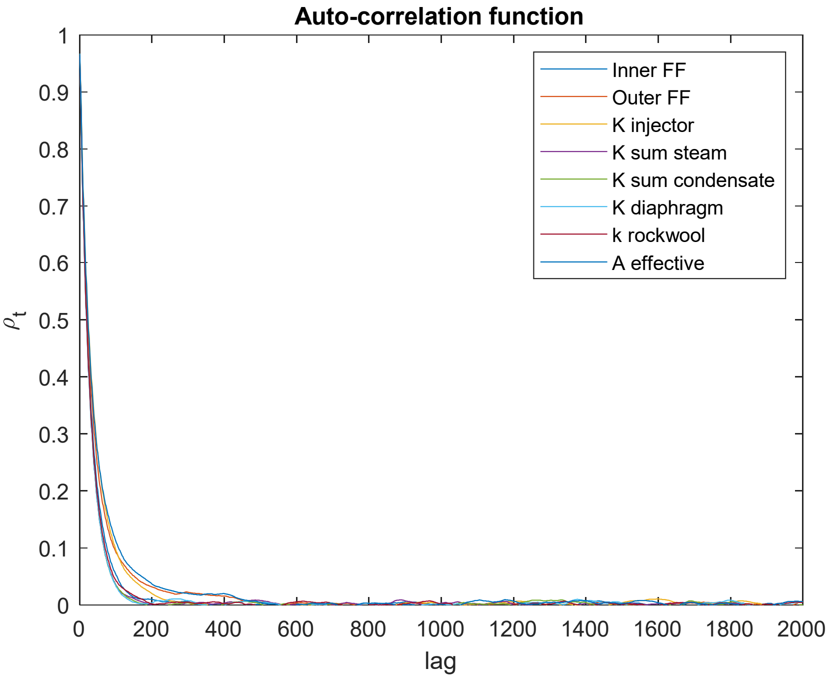


Figure 11. Autocorrelation functions of the MCMC samples.

Figure 12 reports the posterior marginal PDFs of the uncertain calibration parameters $\boldsymbol{\theta}^*$ that are obtained through a kernel density estimation using a reflection boundary correction (Cox et al., 1986).

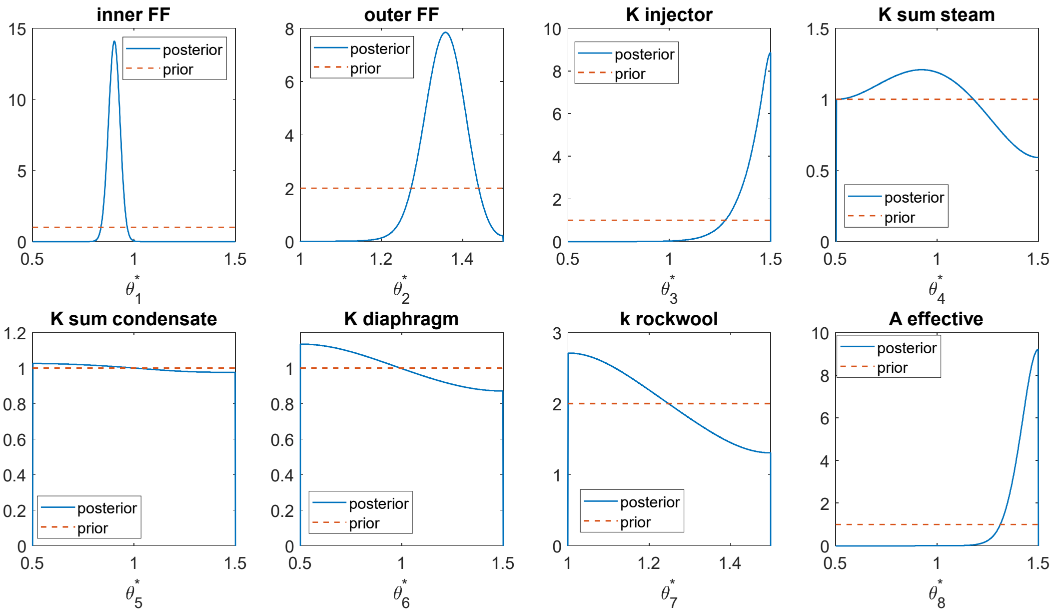


Figure 12. Prior distributions and posteriors' kernel density estimations.

The marginal posteriors of $\theta_1^*, \theta_2^*, \theta_3^*, \theta_8^*$ display a significant change with respect to their priors; in contrast, the other marginal posteriors are only slightly different and maintain the same support of the priors. For θ_1^* and θ_2^*, the posterior PDFs appear to be bounded by the prior supports, which is not the case for the other parameters. In particular, for θ_3^* and θ_8^* the distributions are peaked at their upper bounds, meaning that the IUQ approach is very effective in shaping posterior distributions largely different from the prior distributions. Future works will further investigate the effect of different prior ranges of input on the posterior distribution results, eventually embedding some pieces of experimental evidence. Also, it is worth noticing that different prior PDFs of θ_3^* and θ_8^* would impact on the results of the Sobol indices Sensitivity Analysis (discussed in Section 6.2.1).

Some summary statistics of the posterior distribution (i.e., mean values, modes, 5^{th} and 95^{th} percentiles) are reported in Table 4.

Table 4. Posterior summaries.

$\boldsymbol{\theta_i^*}$	Parameter	Mean values	Modes	5th percentiles	95th percentiles
$\boldsymbol{\theta_1^*}$	Inner_FF	0.90	0.90	0.86	0.94
$\boldsymbol{\theta_2^*}$	Outer_FF	1.36	1.40	1.28	1.43
$\boldsymbol{\theta_3^*}$	K_injector	1.41	1.49	1.25	1.49
$\boldsymbol{\theta_4^*}$	K_sum_steam	0.96	0.88	0.56	1.40
$\boldsymbol{\theta_5^*}$	K_sum_condensate	0.99	1.06	0.55	1.45
$\boldsymbol{\theta_6^*}$	K_ diaphragm	0.96	1.30	0.54	1.44
$\boldsymbol{\theta_7^*}$	k_rockwool	1.21	1.04	1.02	1.45
$\boldsymbol{\theta_8^*}$	A_effective	1.43	1.48	1.33	1.50

The marginal posterior distributions in Figure 12 cannot be used to draw samples because the calibration parameters are not independent. Indeed, Table 5 shows that some of the calibration parameters are strongly correlated.

Table 5. Correlation Matrix computed using the MCMC samples

	$\boldsymbol{\theta_1^*}$	$\boldsymbol{\theta_2^*}$	$\boldsymbol{\theta_3^*}$	$\boldsymbol{\theta_4^*}$	$\boldsymbol{\theta_5^*}$	$\boldsymbol{\theta_6^*}$	$\boldsymbol{\theta_7^*}$	$\boldsymbol{\theta_8^*}$
$\boldsymbol{\theta_1^*}$	1.0000	-	-	-	-	-	-	-
$\boldsymbol{\theta_2^*}$	-0.3984	1.0000	-	-	-	-	-	-
$\boldsymbol{\theta_3^*}$	-0.2768	0.4478	1.0000	-	-	-	-	-
$\boldsymbol{\theta_4^*}$	0.2807	0.0370	0.0493	1.0000	-	-	-	-
$\boldsymbol{\theta_5^*}$	0.0500	0.0209	0.0421	-0.0738	1.0000	-	-	-
$\boldsymbol{\theta_6^*}$	0.0446	0.0916	0.0375	-0.0489	0.0115	1.0000	-	-
$\boldsymbol{\theta_7^*}$	0.0404	-0.2789	0.0477	-0.0301	0.0634	-0.0386	1.0000	-
$\boldsymbol{\theta_8^*}$	0.0267	0.0144	-0.2345	-0.1393	0.0868	0.0473	-0.0628	1.0000

6.2.1 Sensitivity Analysis

In Figure 12, we notice that the posterior distributions of $\theta_4^*, \theta_5^*, \theta_6^*, \theta_7^*$ are slightly different and maintain the same support of the priors. In this regard, employing a global Sensitivity Analysis (SA), we want to determine to what extent the uncertainty of these parameters contributes to the output uncertainty. Global SA is the study of how the output uncertainty of a mathematical model is apportioned in the uncertainty of its inputs considering their entire range of interest (Iooss and Lemaître, 2015). In this work, we adopt Sobol' indices (Saltelli et al., 2010), i.e., a global SA method that decomposes the variance of the model output into fractions assigned to inputs or sets of inputs. If the model output is approximated through a Polynomial Chaos Expansions (PCE) (Sudret, 2015), Sobol' indices can be directly computed by postprocessing the PCE expansion coefficients (Sudret, 2008). Conventional approaches to global SA assume that the model output is scalar (Garcia-Cabrejo and Valocchi, 2014). In the case of vectorial output, (Campbell et al., 2006) proposes expanding such output into a basis (e.g., using PCA) and, then, determining the Sobol' indices for the expansion coefficients with respect to the inputs. In this regard, we firstly approximate each PC through a PCE using a Least Angle Regression algorithm (Lataniotis et al., 2019a); then, following (Sudret, 2008), we post-process the PCE expansion coefficients to compute the first-order Sobol' indices. Further details about Sobol' indices sensitivity analysis are given in (Marelli et al., 2019). Figure 13 reports the first-order Sobol' indices of the four principal components with respect to the eight calibration parameters. While the parameters $\theta_1^*, \theta_2^*, \theta_3^*, \theta_8^*$ are revealed to be the most important, $\theta_4^*, \theta_5^*, \theta_6^*, \theta_7^*$ result to be less influential. Interestingly, the proposed IUQ approach provides posterior PDFs that hardly differ from the prior for input parameters characterized by a low first-order Sobol' index.

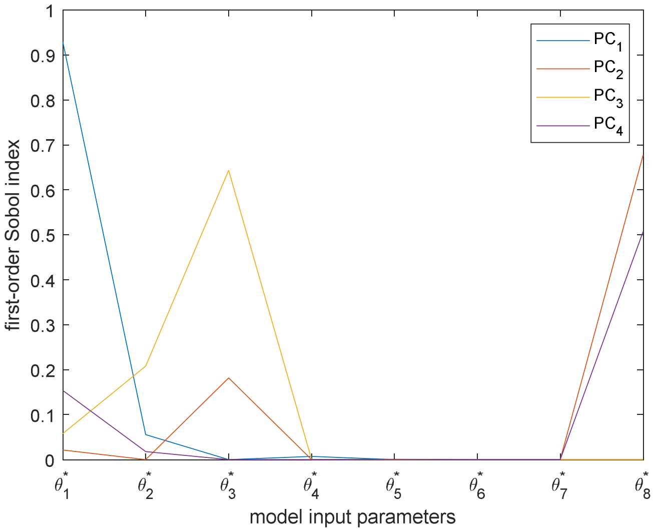


Figure 13. First-order Sobol' indices of the four principal components with respect to the eight calibration parameters.

6.3 Forward Uncertainty Propagation

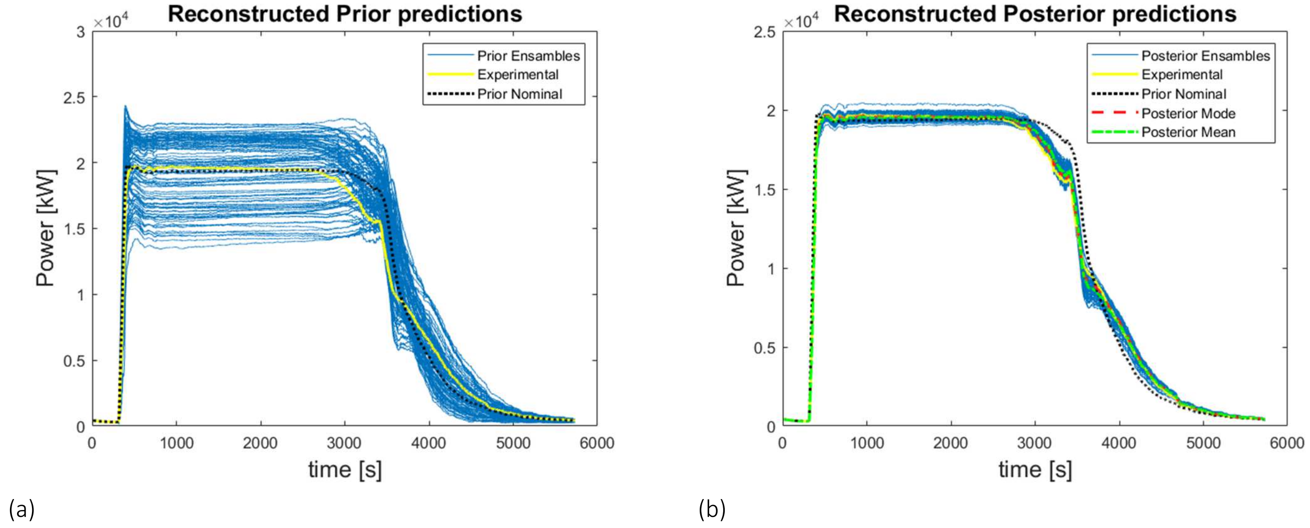
It is not possible to assess the quality of the IUQ results only through the posterior PDFs obtained. Thus, we perform forward uncertainty propagation (by simulating the posterior samples obtained from the MCMC after an appropriate thinning[2]) to check that the updated PDFs are consistent with the Test 7 part 2 experimental data. The forward uncertainty propagation is carried out using both the meta-model (reconstructing the reduced-space predictions through expression (16)) and the RELAP5-3D BE model.

Concerning the forward uncertainty propagation performed through the meta-model, Figure 14 compares the reconstructed Kriging predictions for the prior (Figure 14a) with the reconstructed Kriging predictions for the posterior (Figure 14b). In particular, Figure 14a shows the reconstructed Kriging prediction $\breve{\boldsymbol{y}}$ for 100 input values randomly sampled from the prior compared to the reconstructed experimental data $\breve{\boldsymbol{y}}^E$,

---

[2] For the thinning we keep every 2000th sample draws from each chain and, then, we randomly chose 100 samples from the thinned chains.

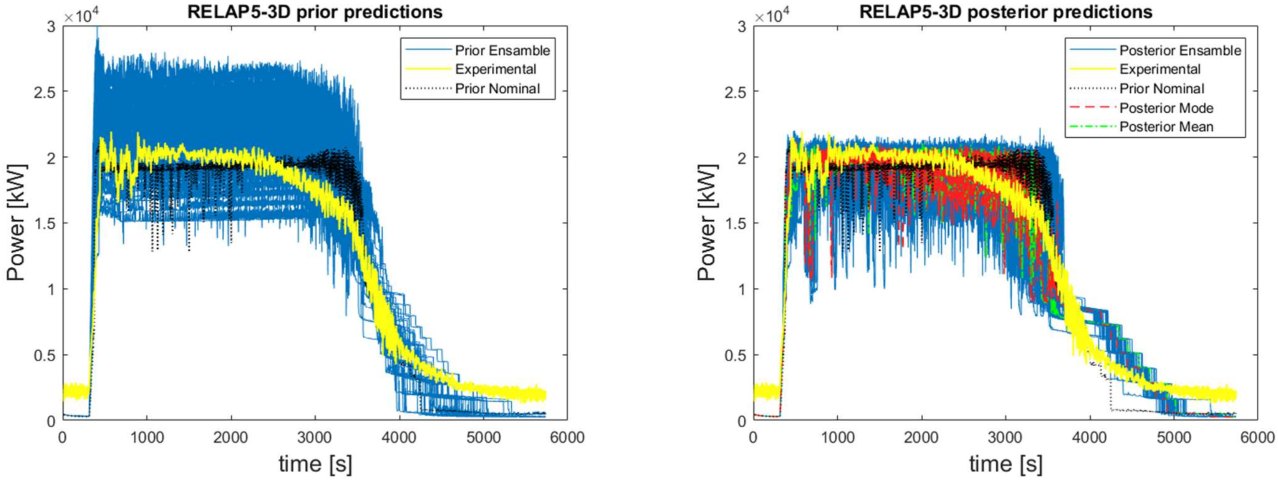
whereas Figure 14b compares $\breve{\boldsymbol{y}}^{E}$ with respect to the reconstructed Kriging predictions $\breve{\boldsymbol{y}}$ of: (1) 100 posterior samples, (2) the posterior mode, (3) the posterior mean value and (4) the prior nominal value. As expected, the simulated posterior samples give a smaller spreading of the results with respect to the prior samples; moreover, the posterior samples predictions envelop the experimental data. Comparing the reconstructed Kriging prediction for the posterior mode and the posterior mean value with respect to that for the prior nominal value, we can notice that the posterior mode and the posterior mean value allow reproducing better $\breve{\boldsymbol{y}}^{E}$.



(a) (b)

Figure 14. Reconstructed Kriging predictions of 100 prior samples compared with respect to the reconstructed experimental data and the reconstructed Kriging prediction of the prior nominal value (Figure 14a). Reconstructed Kriging predictions for: 100 posterior samples, the prior nominal value, the posterior mode and posterior mean value with respect to the reconstructed experimental data (Figure 14b).

The good agreement that we observe for the forward uncertainty propagation performed through the meta-model must be further investigated when the RELAP5-3D BE model is adopted. In this regard, Figure 15 compares the results of $N = 59$ prior and posterior samples simulated through RELAP5-3D. In particular, Figure 15a shows the RELAP5-3D simulations for $N = 59$ input values randomly sampled from the prior compared to $\boldsymbol{y}^{E}$, whereas Figure 15b compares $\boldsymbol{y}^{E}$ with respect to the RELAP5-3D simulations of: (1) $N = 59$ posterior samples, (2) the posterior mode $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{\theta}_{\boldsymbol{posterior}}^{\boldsymbol{mode}})$, (3) the posterior mean value $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{\theta}_{\boldsymbol{posterior}}^{\boldsymbol{mean}})$ and (4) the prior nominal value $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{\theta}_{\boldsymbol{prior}}^{\boldsymbol{nominal}})$. The ensemble of the posterior simulations in Figure 15b is coherent with experimental data (i.e., it roughly envelopes $\boldsymbol{y}^{E}$) and it displays less spreading of the results compared to the prior; however, one can notice that all the RELAP5-3D simulated posterior samples, in general, show much wider oscillations than $\boldsymbol{y}^{\boldsymbol{M}}(\boldsymbol{\theta}_{\boldsymbol{prior}}^{\boldsymbol{nominal}})$. A possible reason for such significant oscillations can be the data filtering performed on the RELAP5-3D simulated HX exchanged power before the dimensionality reduction (i.e., before meta-modeling). Indeed, because of data filtering, the meta-model is no longer able to distinguish input parameters $\boldsymbol{\theta}$ that generate noisy time-series, and, eventually, this can affect the IUQ results.



(a)

(b)

Figure 15. RELAP5-3D predictions of $\boldsymbol{N = 59}$ prior samples compared with respect to experimental data and the RELAP5-3D prediction of the prior nominal value (Figure 15a). RELAP5-3D predictions for: $\boldsymbol{N = 59}$ posterior samples, the prior nominal value, the posterior mode and posterior mean value with respect to experimental data (Figure 15b).

6.3.1 Safety margin calculation

Best Estimate Plus Uncertainty (BEPU) methodologies (D'Auria et al., 2006; IAEA, 2008) typically provide the results in terms of uncertainty ranges for the calculated Figure of Merit (FOM); this allows us to compute the safety margin with respect to a threshold value. Following the GRS method (Glaeser, 2008), order statistic, and in particular Wilk's formula (Wilks, 1941), can be employed to determine the number of simulations N that guarantee the confidence level β and the probability content γ for p FOMs (Guba et al., 2003). In the case of one-sided confidence level with one (i.e., $p = 1$) FOM, the Wilk's formula reduces to:

$$1 - \gamma^N \geq \beta \tag{20}$$

The integral of the HX exchanged power $P(t)$ over the mission time, $T_{mission} = 5736\ s$ (i.e., the duration of Test 7 part 2), is selected as FOM, since it is one of the most relevant parameters of the experiment:

$$E = \int_0^{T_{mission}} P(t)dt \tag{21}$$

In case of a safety review process, other parameters may be selected as FOM, such as the minimum and/or maximum HX power or the energy exchanged over certain intervals. The acceptance criterion for the selected FOM is, for demonstration purposes, set to be $0.9E_{nominal}$; that is, the system is considered to be failed if $E < 0.9 \int_0^{T_{mission}} P_{nominal}(t)dt$, where $P_{nominal}$ is the RELAP5-3D simulated HX exchanged power using as input the prior nominal values. $N = 59$ RELAP5-3D simulations are carried out for calculating the one-sided 95% confidence interval with 95% probability content. According to Wilks's formula, the lowest simulated value of E (i.e., $E_{lowest} = 6.716 \times 10^7\ kJ$) is within the lower 5% range with at least 95% confidence; thus, the margin obtained in this demonstrative application, with respect to the assumed threshold value (i.e., $0.9E_{nominal} = 6.004 \times 10^7 kJ$), is $M = 7.120 \times 10^6 kJ$. Note that the uncertainty analysis carried out in this Section to compute the safety margin is performed only for demonstrative purposes; indeed, it involves the propagation of only some epistemic input uncertainty (i.e., that found through the IUQ process). Additional epistemic input uncertain parameters could be taken into account considering (e.g., model input parameters within the code). A realistic calculation of the safety margin must include both epistemic and aleatory uncertainty. At least, in Figure 16, it can be noticed that the safety margin is higher than the results spreading: this is encouraging considering that the propagation of aleatory uncertainty can introduce further uncertainty in the output, reducing in this way the safety margin.

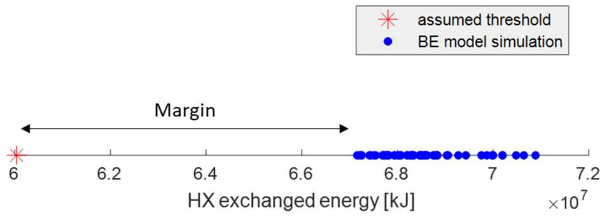


Figure 16. Safety margin with respect to the assumed threshold value (i.e., $\mathbf{0.9}\boldsymbol{E_{nominal}}$).

7 Conclusions

In the present paper, we have proposed an approach for the uncertainty quantification of T-H model codes used for passive safety systems reliability assessment. In particular, a Bayesian IUQ approach has been applied to a RELAP5-3D model of the PERSEO facility, adopting a single time-series measurement data of the HX exchanged power. Principal Component Analysis has been implemented for output dimensionality reduction, and Kriging meta-modeling has been adopted for the emulation of the original T-H model code. In this way, the computational burden related to the MCMC posterior sampling for Bayesian posterior evaluation has been lowered by several orders of magnitudes. Since Kriging accuracy decreases with the principal components order, only four PCs have been retained for a cumulative percentage of variance explained equal to $95\ \%$. The posterior marginal distributions have been obtained from the MCMC samples through a kernel density estimation. The experimental data have been projected onto the PC subspace, thus performing the entire IUQ on such reduced space.

Uniform priors with relatively wide ranges have been adopted to represent the hypothetical case of scarce knowledge about the BE model input parameters. After IUQ, some of the marginal posterior PDFs obtained show supports significantly narrower than the priors. From this perspective, credibly, the proposed IUQ process might reduce the epistemic uncertainty in problems affected by a scarce knowledge of the input parameter PDFs.

A forward uncertainty propagation based on the posterior samples shows agreement between the Test 7 part 2 experimental data of PERSEO and the RELAP5-3D simulations. However, significant oscillations characterize such simulations; thus, further solutions that allow avoiding such oscillations should be considered.

We computed the first-order Sobol’ index of the uncertain input parameters, and we found that the proposed IUQ approach, for input parameters characterized by low first-order Sobol indexes, finds posterior PDFs that hardly differ from the prior. Thus, the IUQ method applied in this paper does not calibrate parameters that are not significant to the output uncertainty.

There are some limitations and drawback in the proposed approach that should be further investigated:

1. Experimental data of a single test (Test 7 part 2) have been considered, which is typically insufficient to train a Kriging for the discrepancy term $\boldsymbol{\delta}(\boldsymbol{x})$. Neglecting this term, the discrepancies in the code results are partially taken into account by the uncertainty of $\boldsymbol{\theta}^*$ (that does not depend on the experimental conditions). This means that the calibrated TH model may perform poorly when applied to other experiments. In spite of this, the current work serves as a demonstration of the methodology and, in future works, additional experimental parameters and tests will be considered to improve the model calibration for developing the a posteriori parameter distribution.

2. PC scores are uncorrelated by definition, but linear uncorrelation of two random variables does not, in general, imply their independence. In this paper, assuming that PC scores are mutually independent allows us to build p^* independent Kriging meta-models for each principal component. As a result, $\boldsymbol{\Sigma_{Kriging}}$ is a diagonal matrix. On the other hand, without the assumption of mutually independent PC scores, we cannot state that they are distributed according to a multivariate gaussian distribution: in that case, one should thoroughly review the formulation of the likelihood function.
3. We implemented PCA, which is a linear dimensionality reduction technique. In general, real data is likely to form a highly nonlinear manifold, and nonlinear dimensionality reduction techniques may offer an advantage (Van Der Maaten et al., 2009). Future works will investigate new approaches for constructing Kriging emulators of time-dependent quantities using nonlinear dimensionality reduction (e.g., autoencoders).
4. The higher the input dimensionality, the higher the size of the training dataset needed to obtain the same meta-model accuracy (Loeppky et al., 2009). In this work, the selection of some input calibration parameters was expert-based. A more rigorous approach should 1) identify and include all the possible model parameters affecting the code output and, eventually, 2) propose a method (e.g., variance-based sensitivity analysis) for selecting the most relevant ones. This last point would reduce the input dimensionality and, consequently, increase Kriging performance (Loeppky et al., 2009).
5. For some of the input parameters (i.e., θ_3^* and θ_8^*), the posterior distributions are peaked at the upper boundary, meaning that if the model is not constrained by the ranges reported in Table 1, the posterior might be different. Future works will further investigate the effect of the range of the prior distribution on the posterior distributions.
6. Filtering of data (both the experimental data and the T-H model predictions) may affect the IUQ results. Although filtering noisy data affected by numerical oscillation may appear reasonable, justifying a particular filtering method rather than another is a non-trivial task. Future works should explore new approaches that allow dealing with noisy raw data without resorting to filtering techniques.

Appendix A

A.1 Kriging Theory

Let us consider a general mathematical model $z = z(\boldsymbol{\theta})$, where z is a scalar and $\boldsymbol{\theta} = [\theta_1, \dots, \theta_d]$ is a d-dimensional vector. Assume that $\boldsymbol{\Theta} = \left[\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m)}\right]^T$ is the Design Of Experiment (DOE) and $\boldsymbol{z} = [z_1, \dots, z_m]^T$ is the vector of the corresponding m model responses. Kriging is a method of interpolation, which assumes that the model output $z(\boldsymbol{\theta})$ is a realization of a Gaussian process indexed by $\boldsymbol{\theta} \in \boldsymbol{D}_{\boldsymbol{\theta}} \subset \mathbb{R}^d$. The mathematical form of a Kriging is given by:

$$z(\boldsymbol{\theta}) = \sum_{j=1}^{n} \beta_j f_j(\boldsymbol{\theta}) + G(\boldsymbol{\theta}) = \boldsymbol{\beta}^{\boldsymbol{T}} \boldsymbol{f}(\boldsymbol{\theta}) + G(\boldsymbol{\theta}) \tag{A.1}$$

The first term in Eq. (8), $\boldsymbol{\beta}^{\boldsymbol{T}} \boldsymbol{f}(\boldsymbol{\theta})$, is the *mean* value of the Gaussian process, also called the "trend"; it is given by n arbitrary functions $[f_1(\boldsymbol{\theta}), \dots, f_n(\boldsymbol{\theta})]$ and the related coefficients $[\beta_1, \dots, \beta_n]$. The second term, $G(\boldsymbol{\theta})$, is a Gaussian process with zero mean and covariance:

$$Cov\left[G\left(\boldsymbol{\theta}^{(i)}\right), G\left(\boldsymbol{\theta}^{(j)}\right)\right] = \sigma^2 R\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) \tag{A.2}$$

where $\sigma^2$is the *process variance* and $R(\cdot,\cdot)$ is the *correlation function* (also called *correlation kernel*), defined for any two points in the input domain $\boldsymbol{D}_{\boldsymbol{\theta}}$.

$R(\cdot,\cdot)$ is a function of the distance $h(\cdot,\cdot)$:

$$R\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) = R\left(h(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)})\right) \tag{A.3}$$

According to (Lataniotis et al., 2019b), $h(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)})$ has the following expression:

$$h\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) = \left[\sum_{k=1}^{d} \left(\frac{\theta_k^{(i)} - \theta_k^{(j)}}{\omega_k}\right)^2\right]^{0.5} \tag{A.4}$$

where the parameters $\boldsymbol{\omega} = [\omega_1, \dots, \omega_d]$ are called *length scale parameters*. Table 6 and Table 7 show, respectively, some common correlation kernels and trend types that are implemented and compared in this work to find the best Kriging options.

Table 6. Common correlation kernels

Correlation Kernel	Expression
Exponential	$\boldsymbol{R}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) = \mathbf{exp}\left[-\boldsymbol{h}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right)\right]$
Gaussian	$\boldsymbol{R}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) = \mathbf{exp}\left[-\frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{h}^2\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right)\right]$
Matérn-3/2	$\boldsymbol{R}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) = \left(\mathbf{1} + \sqrt{\mathbf{3}} \cdot \boldsymbol{h}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right)\right) \mathbf{exp}\left[-\sqrt{\mathbf{3}} \cdot \boldsymbol{h}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right)\right]$
Matérn-5/2	$\boldsymbol{R}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) = \left(\mathbf{1} + \sqrt{\mathbf{5}} \cdot \boldsymbol{h}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right) + \frac{\mathbf{5}}{\mathbf{3}} \cdot \boldsymbol{h}^2\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right)\right) \mathbf{exp}\left[-\sqrt{\mathbf{5}} \cdot \boldsymbol{h}\left(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^{(j)}\right)\right]$

Table 7. Common Trend Types

Trend type	Expression
Constant (ordinary Kriging)	$\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{f}(\boldsymbol{\theta}) = \boldsymbol{\beta}_{\boldsymbol{0}}$
Linear	$\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{f}(\boldsymbol{\theta}) = \boldsymbol{\beta}_{\boldsymbol{0}} + \sum_{i=1}^{d} \boldsymbol{\beta}_i\boldsymbol{\theta}_i$
Quadratic	$\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{f}(\boldsymbol{\theta}) = \boldsymbol{\beta}_{\boldsymbol{0}} + \sum_{i=1}^{d} \boldsymbol{\beta}_i\boldsymbol{\theta}_i + \sum_{i=1}^{d}\sum_{j=1}^{d} \boldsymbol{\beta}_{ij}\boldsymbol{\theta}_i\boldsymbol{\theta}_j$

In line with the Kriging definition, $\boldsymbol{z}$ and a generic model prediction $z(\boldsymbol{\theta})$ at the input location $\boldsymbol{\theta}$ are jointly distributed according to a multivariate Gaussian distribution:

$$\begin{Bmatrix} z(\boldsymbol{\theta}) \\ \boldsymbol{z} \end{Bmatrix} \sim N\left(\begin{Bmatrix} \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{f}(\boldsymbol{\theta}) \\ \boldsymbol{F\beta} \end{Bmatrix}, \sigma^2 \begin{Bmatrix} 1 & \boldsymbol{r}^{\boldsymbol{T}}(\boldsymbol{\theta}) \\ \boldsymbol{r}(\boldsymbol{\theta}) & \boldsymbol{R} \end{Bmatrix}\right) \tag{A.5}$$

where $\boldsymbol{F}$ is the matrix of regression functions $\boldsymbol{f}(\boldsymbol{\theta})$ evaluated in correspondence of the m DOE points (i.e., $F_{ij} = f_j(\boldsymbol{\theta}_{\boldsymbol{i}})$, $i = 1, \dots, m;\ \ j = 1, \dots, n$); $\boldsymbol{r}(\boldsymbol{\theta})$ is the vector of correlations between $\boldsymbol{\theta}$ and the DOE points (i.e., $r_i = R\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)}\right)\ \ i = 1, \dots, m$); $\boldsymbol{R}$ is the correlation matrix (i.e., $R_{ij} = R\left(\boldsymbol{\theta}^{(j)}, \boldsymbol{\theta}^{(i)}\right)\ \ i, j = 1, \dots, m$).

The Kriging prediction (i.e., the mean value) $\hat{z}(\boldsymbol{\theta})$ and the variance $\sigma^2_{z(\boldsymbol{\theta})}$ of the Gaussian random variable $z(\boldsymbol{\theta})$ are given by (Dubourg, 2011):

$$\hat{z}(\boldsymbol{\theta}) = \boldsymbol{f}(\boldsymbol{\theta})^T\widehat{\boldsymbol{\beta}} + \boldsymbol{r}(\boldsymbol{\theta})^T\boldsymbol{R}^{-1}(\boldsymbol{z} - \boldsymbol{F}\widehat{\boldsymbol{\beta}}) \tag{A.6}$$

$$\sigma^2_{z(\boldsymbol{\theta})} = \sigma^2\left(1 - \boldsymbol{r}^T(\boldsymbol{\theta})\boldsymbol{R}^{-1}\boldsymbol{r}(\boldsymbol{\theta}) + \left(\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{R}^{-\boldsymbol{1}}\boldsymbol{r}(\boldsymbol{\theta}) - \boldsymbol{f}(\boldsymbol{\theta})\right)^{\boldsymbol{T}}(\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{R}^{-\boldsymbol{1}}\boldsymbol{F})^{-1}\left(\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{R}^{-\boldsymbol{1}}\boldsymbol{r}(\boldsymbol{\theta}) - \boldsymbol{f}(\boldsymbol{\theta})\right)\right) \tag{A.7}$$

where $\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{R}^{-\boldsymbol{1}}\boldsymbol{F}\right)^{-1}\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{R}^{-\boldsymbol{1}}\boldsymbol{z}$ is the least square estimate of the regression coefficients.

A crucial consequence of the Gaussian assumption is that:

$$z(\boldsymbol{\theta}) \sim N\left(\hat{z}(\boldsymbol{\theta}), \sigma^2_{z(\boldsymbol{\theta})}\right). \tag{A.8}$$

A property of the Kriging predictor is that it always interpolates the design sites. The mean square error of the prediction (i.e., $\sigma^2_{z(\boldsymbol{\theta}^*)}$) collapses to zero as the new point $\boldsymbol{\theta}$ gets close to the design sites and increases when $\boldsymbol{\theta}$ moves away from the design sites. Intuitively, a richer DOE will give a Kriging with a better prediction capability. However, increasing the DOE size is generally computationally intensive; therefore, the selection of design sites through an efficient DOE approach is a crucial issue to be addressed. Latin Hypercube Sampling (LHS) is one of the most popular DOE techniques (Helton and Davis, 2003), and it is adopted in this work.

To formulate the general theory of Kriging, we introduced $(n + d + 1)$ hyperparameters (i.e., n regression coefficients $\boldsymbol{\beta}$, d length scale parameters $\boldsymbol{\omega}$, one process variance σ^2) that are unknown and must be estimated. Estimation methods such as M*aximum likelihood estimation* (ML) or *Cross-Validation estimation*

(CV) are usually applied to estimate them. In this work, we use the software package UQLab (Lataniotis et al., 2019b) to implement the Kriging meta-modeling. Further details regarding the comparison between ML and CV can be found in (Bachoc, 2013).

A.2 Kriging performance indicators

The quality assessment for a Kriging model is usually performed considering the accuracy in reproducing the original model output at unobserved locations $\boldsymbol{\theta}$ (Martin and Simpson, 2005). This is typically done by measuring the error in Kriging's prediction using a *validation dataset* (i.e., a set of model simulations different from $\boldsymbol{\Theta}$). Such errors are collected in the R^2_{actual} indicator, which is defined as:

$$R^2_{actual} = 1 - \frac{\sum_{i=1}^{N_{val}} \left(z(\boldsymbol{\theta}^{(i)}) - \hat{z}(\boldsymbol{\theta}^{(i)}) \right)^2}{\sum_{i=1}^{N_{val}} (z(\boldsymbol{\theta}^{(i)}) - \bar{z})^2} \tag{A.9}$$

where N_{val} is the size of the validation dataset, $z\left(\boldsymbol{\theta}^{(i)}\right)$ and $\hat{z}(\boldsymbol{\theta}^{(i)})$ are the model and the Kriging predictions for $\boldsymbol{\theta}^{(i)}$ (i.e., the i^{th} DOE input point of the validation dataset), respectively, $\bar{z} = \frac{1}{N_{val}} \sum_{i=1}^{N_{val}} z(\boldsymbol{\theta}^{(i)})$. However, this approach requires a considerable number of additional simulations ($N_{val} \gg m$) that likely are computationally intensive.

Computationally cheaper alternatives, namely the *cross-validation* (CV) and *Akaike's information criterion* (AIC), can be used to estimate the meta-model quality without any additional simulation beyond those used to train the Kriging (Martin and Simpson, 2005).

Concerning the indicator base on CV, the meta-model is trained leaving out a fixed number of simulations (typically one) from the training dataset. This new model is used to predict the remaining simulation outcomes and the respective errors. Such errors are collected in the $R^2_{prediction}$ indicator, which is defined as:

$$R^2_{prediction} = 1 - \frac{\sum_{i=1}^{m} \left(z(\boldsymbol{\theta}^{(i)}) - \hat{z}_{(-i)}(\boldsymbol{\theta}^{(i)}) \right)^2}{\sum_{i=1}^{m} \left(z(\boldsymbol{\theta}^{(i)}) - \bar{z} \right)^2} \tag{A.10}$$

where m is the training dataset size, $\hat{z}_{(-i)}(\boldsymbol{\theta}^{(i)})$ is the prediction of the Kriging that is trained using all the points of $\boldsymbol{\Theta}$, except $\boldsymbol{\theta}^{(i)}$, $\bar{z} = \frac{1}{m} \sum_{i=1}^{m} z(\boldsymbol{\theta}^{(i)})$.

Concerning the indicator based on the AIC, which is a function of the number of hyperparameters q (i.e., $n + d + 1$) and the likelihood function $L(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\omega}|\boldsymbol{z})$:

$$AIC = -2 \ln\left(L(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\omega}|\boldsymbol{z}) \right) + 2q \tag{A.11}$$

where:

$$L(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\omega}|\boldsymbol{z}) = \frac{(\det(\boldsymbol{R}))^{-1/2}}{(2\pi\sigma^2)^{m/2}} \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{z} - \boldsymbol{F\beta})^T \boldsymbol{R}^{-1}(\boldsymbol{z} - \boldsymbol{F\beta})\right] \tag{A.12}$$

If m is small (i.e., $\frac{m}{q} < 40$) the AIC should be corrected through:

$$AIC_c = AIC + 2q(q+1)/(m-q-1). \tag{A.13}$$

$R^2_{prediction}$ and AIC_c will be used in this work as indicators of the quality of the Kriging models. In agreement with their definitions, we want $R^2_{prediction}$ to be as close to one as possible and AIC_c as small as possible. Moreover, we expect that Kriging, having hyperparameters estimated through CV, will result in a better $R^2_{prediction}$ than that obtained by ML estimation, since the CV estimation method computes $\{\boldsymbol{\beta}, \boldsymbol{\omega}, \sigma^2\}$ such that $R^2_{prediction}$is maximized. As concluded in (Martin and Simpson, 2005):

1. AIC_c provides the best criterion to select between competing Kriging model forms, but it cannot be used to estimate the actual model quality (i.e., R^2_{actual});
2. The CV-based factor $R^2_{prediction}$ seems to be a reasonable indicator of the actual model quality.

Appendix B

Let us suppose that that $\boldsymbol{\Theta} = \left[\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m)}\right]$ is the DOE and $\widetilde{\boldsymbol{Y}} = \left[\widetilde{\boldsymbol{y}}^{(1)}, \dots, \widetilde{\boldsymbol{y}}^{(m)}\right]$ is the $p \times m$ *data matrix* containing the m corresponding p-dimensional time-series (also called *observations*) (where $\widetilde{\boldsymbol{y}}^{(i)} = \left[\tilde{y}_1{}^{(i)}, \dots, \tilde{y}_p{}^{(i)}\right]^T$). Before applying SVD, $\widetilde{\boldsymbol{Y}}$ is centered with respect to the simulations mean vector:

$$\widetilde{\boldsymbol{Y}}_{centered} = \left[(\widetilde{\boldsymbol{y}}^{(1)} - \overline{\boldsymbol{\mu}}_{\widetilde{Y}}), \dots, (\widetilde{\boldsymbol{y}}^{(m)} - \overline{\boldsymbol{\mu}}_{\widetilde{Y}})\right] \tag{B.1}$$

where

$$\overline{\boldsymbol{\mu}}_{\widetilde{Y}} = \frac{1}{m}\sum_{i=1}^{m} \widetilde{\boldsymbol{y}}^{(i)} \tag{B.2}$$

The SVD decomposition of $\widetilde{\boldsymbol{Y}}_{centered}$ is (Wall et al., 2003):

$$\widetilde{\boldsymbol{Y}}_{centered} = U \Lambda V^T \tag{B.3}$$

where U is a $p \times p$ unitary matrix whose columns $\boldsymbol{u}_j$ are called *left singular vectors*, Λ is a $p \times m$ diagonal matrix whose entries $\sqrt{\lambda_j}$ are called *singular values*, V^T is the conjugate transpose of V that is a $m \times m$ unitary matrix. The columns of V^T are called *right singular vectors*. There is a direct relation between PCA and SVD (Wall et al., 2003):

- the entries of Λ (i.e., $\sqrt{\lambda_j}$) are the square roots of the eigenvalues (arranged in descending order) of $\widetilde{\boldsymbol{Y}}_{centered}\widetilde{\boldsymbol{Y}}^{\boldsymbol{T}}_{centered}$, thus they are proportional to the covariance matrix eigenvalues;

- $\boldsymbol{u}_j$ are the *principal components vectors* (PCs).

To establish the dimension of the principal subspace p^* a heuristic method is often used: one finds the smallest p^* such that the cumulative percentage of variation explained $\left(i.e., \frac{\sum_{j=1}^{p^*} \lambda_j}{\sum_{j=1}^{p} \lambda_j} \cdot 100\right)$ is at least equal to a threshold value (usually chosen between 95% and 99%) (Wilkinson, 2010). The first p^* PCs $\boldsymbol{u}_j$ form the $p^* \times p$ transformation matrix Φ (with $p^* \ll p$):

$$\Phi = \begin{bmatrix} \boldsymbol{u}_1^{\boldsymbol{T}} \\ \vdots \\ \boldsymbol{u}_{p^*}^{\boldsymbol{T}} \end{bmatrix} \tag{B.4}$$

The transformation of the data matrix into the feature space is given by:

$$\tilde{\boldsymbol{Z}} = \Phi(\tilde{\boldsymbol{Y}} - \bar{\boldsymbol{\mu}}_{\tilde{Y}}) = [\tilde{\boldsymbol{z}}^{(1)}, \dots, \tilde{\boldsymbol{z}}^{(m)}] = \begin{bmatrix} \tilde{z}_1^{(1)} & \cdots & \tilde{z}_1^{(m)} \\ \vdots & \ddots & \vdots \\ \tilde{z}_{p^*}^{(1)} & \cdots & \tilde{z}_{p^*}^{(m)} \end{bmatrix} \tag{B.5}$$

where $\tilde{\boldsymbol{Z}}$ is a $p^* \times m$ matrix whose rows are called *PC scores* and columns represent, in this case, the RELAP5-3D results projected into the features space. As training examples for the surrogate model, the DOE points $[\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(m)}]$ and the respective p^* −dimensional ($p^* \ll p$) projected model responses $\tilde{\boldsymbol{Z}} = [\tilde{\boldsymbol{z}}^{(1)}, \dots, \tilde{\boldsymbol{z}}^{(m)}]$ are used. The Kriging prediction for a specific input $\boldsymbol{\theta}^{(i)}$ will be the $p^* \times 1$ vector $\hat{\boldsymbol{z}}$. To antitransform a generic vector $\boldsymbol{z}$ from the features space to the original space:

$$\breve{\boldsymbol{y}} = \bar{\boldsymbol{\mu}}_{\tilde{Y}} + \Phi^T \boldsymbol{z} \tag{B.6}$$

However, no antitransformation is performed in this work for the IUQ since both the likelihood and the experimental are transformed into the reduced subspace.

References

Apostolakis, G., 1994. A commentary on model uncertainty.

Arendt, P.D., Apley, D.W., Chen, W., 2012. Quantification of model uncertainty: Calibration, model discrepancy, and identifiability. J. Mech. Des. Trans. ASME 134, 1–12. https://doi.org/10.1115/1.4007390

Bachoc, F., 2013. Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecification. Comput. Stat. Data Anal. 66, 55–69. https://doi.org/10.1016/j.csda.2013.03.016

Bandini, G., Meloni, P., Polidori, M., Lombardo, C., 2011. Validation of CATHARE V2.5 thermal-hydraulic code against full-scale PERSEO tests for decay heat removal in LWRs. Nucl. Eng. Des. 241, 4662–4671. https://doi.org/10.1016/j.nucengdes.2011.02.034

Bassenghi, F., 2013. Validation of the CFD code NEPTUNE for a full scale simulator for decay heat removal systems with in-pool heat exchangers. https://doi.org/10.6092/unibo/amsdottorato/5405

Bersano, A., Bertani, C., Falcone, N., De Salve, M., Mascari, F., Meloni, P., 2020. Qualification of RELAP5-3D code against the in-pool passive energy removal system PERSEO data, in: Baraldi, P., Di Maio, F., Zio, E. (Eds.), Proceedings of the 30th European Reliability Conference and the 15th Probability Safety Assessment and Management Conference. pp. 1–8.

Bersano, A., Falcone, N., Bertani, C., De Salve, M., Mascari, F., Lombardo, C., 2019. Validation of RELAP5-3D thermal-hydraulic code against full- scale PERSEO Test 9, in: International Congress on Advances in Nuclear Power Plants.

Bestion, D., D'Auria, F., Lien, P., Nakamura, H., 2017. A state-of-the-art report on scaling in system thermal-hydraulics applications to nuclear reactor safety and design, NEA/CSNI/R(2016)14.

Burgazzi, L., 2007. Addressing the uncertainties related to passive system reliability. Prog. Nucl. Energy 49, 93–102. https://doi.org/10.1016/j.pnucene.2006.10.003

Burgazzi, L., 2004. Evaluation of uncertainties related to passive systems performance. Nucl. Eng. Des. 230, 93–106. https://doi.org/10.1016/j.nucengdes.2003.10.011

Campbell, K., McKay, M.D., Williams, B.J., 2006. Sensitivity analysis when model outputs are functions. Reliab. Eng. Syst. Saf. 91, 1468–1472. https://doi.org/10.1016/j.ress.2005.11.049

Cox, D.R., Hinkley, D. V, Rubin, D., Silverman, B.W., Bartlett, M.S., Beard, R.E., Pentikäinen, T., Pesonen, E., Berry, D.A., Fristedt, B., Cook, R.D., Weisberg, S., Isham, V., Lewis, P.A.W., Oakes, D., Cox, Q.D.R., Smith, W.L., Davis, M.H.A., Vinter, R., 1986. Density Estimation for Statistics and Data Analysis.

D'Auria, F., Bousbia-salah, A., Petruzzi, A., del Nevo, A., 2006. State of the Art in Using Best Estimate Calculation Tools in Nuclear Technology. Nucl. Eng. Technol. 38, 11–32.

de Crecy, A., 1996. Determination of the uncertainties of the constitutive relationships in the Cathare 2 code. Proc. ASME/JSME Int. Conf. Nucl. Eng. ICONE.

Dubourg, V., 2011. Méta-modèles adaptatifs pour l'analyse de fiabilité et l'optimisation sous contrainte fiabiliste 308.

Durga Rao, K., Kushwaha, H.S., Verma, A.K., Srividya, A., 2007. Quantification of epistemic and aleatory uncertainties in level-1 probabilistic safety assessment studies. Reliab. Eng. Syst. Saf. 92, 947–956. https://doi.org/10.1016/j.ress.2006.07.002

Faes, M., Broggi, M., Patelli, E., Govers, Y., Mottershead, J., Beer, M., Moens, D., 2019. A multivariate interval approach for inverse uncertainty quantification with limited experimental data. Mech. Syst. Signal Process. 118, 534–548. https://doi.org/10.1016/j.ymssp.2018.08.050

Ferri, R., Achilli, A., Cattadori, G., Bianchi, F., Meloni, P., 2005. Design, experiments and Relap5 code calculations for the perseo facility. Nucl. Eng. Des. 235, 1201–1214. https://doi.org/10.1016/j.nucengdes.2005.02.011

Ferson, S., Ginzburg, L.R., 1996. Different methods are needed to propagate ignorance and variability. Reliab. Eng. Syst. Saf. 54, 133–144. https://doi.org/10.1016/S0951-8320(96)00071-3

Garcia-Cabrejo, O., Valocchi, A., 2014. Global Sensitivity Analysis for multivariate output using Polynomial Chaos Expansion. Reliab. Eng. Syst. Saf. 126, 25–36. https://doi.org/10.1016/j.ress.2014.01.005

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2015. Bayesian data analysis 3rd ed, CEUR Workshop Proceedings. https://doi.org/10.1017/CBO9781107415324.004

Glaeser, H., 2008. GRS method for uncertainty and sensitivity evaluation of code results and applications. Sci. Technol. Nucl. Install. 2008. https://doi.org/10.1155/2008/798901

Guba, A., Makai, M., Pál, L., 2003. Statistical aspects of best estimate method - I. Reliab. Eng. Syst. Saf. 80, 217–232. https://doi.org/10.1016/S0951-8320(03)00022-X

Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. Bernoulli 7, 223–242. https://doi.org/10.2307/3318737

Helton, J.C., Davis, F.J., 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab. Eng. Syst. Saf. 81, 23–69. https://doi.org/10.1016/S0951-8320(03)00058-9

Heo, J., Lee, S.W., Kim, K.D., 2014. Implementation of data assimilation methodology for physical model uncertainty evaluation using post-chf experimental data. Nucl. Eng. Technol. 46, 619–632. https://doi.org/10.5516/NET.02.2013.083

Higdon, D., Gattiker, J., Williams, B., Rightley, M., 2008. Computer model calibration using high-dimensional output. J. Am. Stat. Assoc. 103, 570–583. https://doi.org/10.1198/016214507000000888

Hoffman, F.O., Hammonds, J.S., 1994. Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability. Risk Anal. 14, 707–712. https://doi.org/10.1111/j.1539-6924.1994.tb00281.x

IAEA, 2014. Progress in Methodologies for the Assessment of Passive Safety System Reliability in Advanced Reactors.

IAEA, 2012. Natural Circulation Phenomena and Modelling for Advanced Water Cooled Reactors, Iaea-Tecdoc-1677.

IAEA, 2008. Best Estimate Safety Analysis for Nuclear Power Plants: Uncertainty Evaluation; Safety Reports Series 52 1–211.

Idaho National Laboratory, 2015. RELAP5-3D Code Manual Volume I: Code Structure, System Models and Solution Methods, 2015.

Iooss, B., Lemaître, P., 2015. A Review on Global Sensitivity Analysis Methods, in: Uncertainty Management in Simulation-Optimization of Complex Systems: Algorithms and Applications. pp. 101–122.

Jolliffe, I.T., 2002. Principal Component Analysis, 2nd ed. Springer-Verlag New York.

Kennedy, M.C., O'Hagan, A., 2001. Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B (Statistical Methodol. 63, 425–464. https://doi.org/10.1111/1467-9868.00294

Kovtonyuk, A., Petruzzi, A., D'Auria, F., 2012. A procedure for characterizing the range of input uncertainty parameters by the use of the FFTBM. Int. Conf. Nucl. Eng. Proceedings, ICONE 3, 9–19. https://doi.org/10.1115/ICONE20-POWER2012-54025

Lataniotis, C., Marelli, S., Sudret, B., 2019a. UQLab user manual – Polynomial chaos expansions, Report # UQLab-V1.3-104, Chair of Risk, Safety and Uncertainty Quantification. ETH Zurich, Switzerland.

Lataniotis, C., Wicaksono, D., Marelli, S., Sudret, B., 2019b. UQLab User Manual: Kriging (Gaussian Process Modeling) Report # UQLab-V1.3-105. Chair Risk, Saf. Uncertain. Quantif. ETH Zurich, Switz. 1–18.

Loeppky, J.L., Sacks, J., Welch, W.J., 2009. Choosing the sample size of a computer experiment: A practical guide. Technometrics 51, 366–376. https://doi.org/10.1198/TECH.2009.08040

Marelli, S., Lamas, C., Konakli, K., Mylonas, C., Wiederkehr, P., Sudret, B., 2019. UQLab user manual - Sensitivity analysis 86.

Marque`s, M., Pignatel, J.F., D’Auria, F., Burgazzi, L., Mu¨ller, C., Cojazzi, G., La Lumia, V., 2002. Reliability Methods for Passive Safety Functions. https://doi.org/10.1115/ICONE10-22274

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. Nucl. Eng. Des. 235, 2612–2631. https://doi.org/10.1016/j.nucengdes.2005.06.008

Martin, J.D., Simpson, T.W., 2005. Use of kriging models to approximate deterministic computer models. AIAA J. 43, 853–863. https://doi.org/10.2514/1.8650

Mascari, F., Lombardo, C., De Salve, M., Bertani, C., Bersano, A., Falcone, N., Panella, B., 2019. Description of PERSEO Test n. 7 for International Open Benchmark Excercise, ADPFISS-LP1-126.

Mascari, F., Nakamura, H., Ummlnger, K., De Rosa, F., D’Auria, F., 2015. Scaling issues for the experimental characterization of reactor coolant system in integral test facilities and role of system code as extrapolation tool. Proc. 16th Int. Top. Meet. Nucl. React. Therm. Hydraul. 6, 4921–4934.

McFarland, J., Mahadevan, S., Romero, V., Swileir, L., 2008. Calibration and uncertainty analysis for computer simulations with multivariate output. AIAA J. 46, 1253–1265. https://doi.org/10.2514/1.35288

Nagel, J.B., Rieckermann, J., Sudret, B., 2020. Principal component analysis and sparse polynomial chaos expansions for global sensitivity analysis and model calibration: Application to urban drainage simulation. Reliab. Eng. Syst. Saf. 195, 106737. https://doi.org/10.1016/j.ress.2019.106737

NEA/CSNI/R(2016)9, 2016. PREMIUM, a benchmark on the quantification of the uncertainty of the physical models in the system thermal-hydraulic codes: methodologies and data review - csni-r2016-9.pdf, CSNI report.

NEA/CSNI/R(2021)2, 2021. Status Report on Reliability of Thermal-Hydraulic Passive Systems (Addendum: PERSEO Benchmark Report).

Oberkampf, W.L., Helton, J.C., Joslyn, C.A., Wojtkiewicz, S.F., Ferson, S., 2004. Challenge problems: Uncertainty in system response given uncertain parameters. Reliab. Eng. Syst. Saf. 85, 11–19. https://doi.org/10.1016/j.ress.2004.03.002

Pagani, L.P., Apostolakis, G.E., Hejzlar, P., 2005. The impact of uncertainties on the performance of passive systems. Nucl. Technol. 149, 129–140. https://doi.org/10.13182/NT149-129

Pierro, F., Araneo, D., Galassi, G., D’Auria, F., 2009. Application of REPAS methodology to assess the reliability of passive safety systems. Sci. Technol. Nucl. Install. 2009. https://doi.org/10.1155/2009/768947

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput. Phys. Commun. 181, 259–270. https://doi.org/10.1016/j.cpc.2009.09.018

Shrestha, R., Kozlowski, T., 2016. Inverse uncertainty quantification of input model parameters for thermal-hydraulics simulations using expectation–maximization under Bayesian framework. J. Appl. Stat. 43, 1011–1026. https://doi.org/10.1080/02664763.2015.1089220

Sudret, B., 2015. Polynomial chaos expansions and stochastic finite element methods, in: Risk and Reliability in Geotechnical Engineering. pp. 265–300.

Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliab. Eng. Syst. Saf. 93, 964–979. https://doi.org/10.1016/j.ress.2007.04.002

Van Der Maaten, L.J.P., Postma, E.O., Van Den Herik, H.J., 2009. Dimensionality Reduction: A Comparative Review. J. Mach. Learn. Res. 10, 1–41. https://doi.org/10.1080/13506280444000102

Wagner, P.-R., Nagel, J., Marelli, S., Sudret, B., 2019. UQLab user manual: Bayesian inference for model calibration and validation.

Wall, M.E., Rechtsteiner, A., Rocha, L.M., 2003. Singular Value Decomposition and Principal Component Analysis, in: A Practical Approach to Microarray Data Analysis. pp. 91–109. https://doi.org/10.1007/0-306-47815-3_5

Wilkinson, R.D., 2010. Bayesian Calibration of Expensive Multivariate Computer Experiments. Large-Scale Inverse Probl. Quantif. Uncertain. 195–215. https://doi.org/10.1002/9780470685853.ch10

Wilks, S.., 1941. Determination of sample sizes for setting tolerance limits. Ann. Math. Stat. 12, 91–96.

Winkler, R.L., 1996. Uncertainty in probabilistic risk assessment. Reliab. Eng. Syst. Saf. 54, 127–132. https://doi.org/10.1016/S0951-8320(96)00070-1

Wu, X., Kozlowski, T., Meidani, H., 2018a. Kriging-based inverse uncertainty quantification of nuclear fuel performance code BISON fission gas release model using time series measurement data. Reliab. Eng. Syst. Saf. 169, 422–436. https://doi.org/10.1016/j.ress.2017.09.029

Wu, X., Kozlowski, T., Meidani, H., Shirvan, K., 2018b. Inverse uncertainty quantification using the modular Bayesian approach based on Gaussian process, Part 1: Theory. Nucl. Eng. Des. 335, 339–355. https://doi.org/10.1016/j.nucengdes.2018.06.004