POLITECNICO DI TORINO
Repository ISTITUZIONALE

Global Sensitivity Analysis for Segmented Inverse Uncertainty Quantification in the Safety Analysis of Nuclear Power Plants

Original
Global Sensitivity Analysis for Segmented Inverse Uncertainty Quantification in the Safety Analysis of Nuclear Power Plants / Di Maio, F., Matteo Coscia, T., Pedroni, N., Bersano, A., Mascari, F., Zio, E.. - In: ANNALS OF NUCLEAR ENERGY. - ISSN 0306-4549. - ELETTRONICO. - 208:(2024). [10.1016/j.anucene.2024.110791]

Availability:
This version is available at: 11583/2993658 since: 2024-10-24T14:45:10Z

Publisher:
Elsevier Ltd

Published
DOI:10.1016/j.anucene.2024.110791

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

16 November 2026

Annals of Nuclear Energy 208 (2024) 110791



Contents lists available at ScienceDirect

Annals of Nuclear Energy

journal homepage: www.elsevier.com/locate/anucene





Global Sensitivity Analysis for Segmented Inverse Uncertainty Quantification in the Safety Analysis of Nuclear Power Plants

Francesco Di Maio [a,*], Thomas Matteo Coscia [a], Nicola Pedroni [b], Andrea Bersano [c], Fulvio Mascari [c], Enrico Zio [a,d]

[a] Energy Department, Politecnico di Milano, Milano, Italy
[b] Energy Department, Politecnico di Torino, Torino, Italy
[c] ENEA-BOLOGNA, Bologna, Italy
[d] MINES Paris-PSL, Centre de Recherche sur les Risques et les Crises (CRC), Sophia Antipolis, France

ARTICLE INFO

Keywords:
Nuclear Power Plants (NPPs)
Safety analysis
Best Estimate Plus Uncertainty (BEPU)
Inverse Uncertainty Quantification (IUQ)
Sensitivity Analysis (SA)

ABSTRACT

Within the Best Estimate Plus Uncertainty framework for the safety analysis of Nuclear Power Plants, the quantification of the uncertainties affecting the Thermal-Hydraulics (T-H) codes used is crucial. For this, Inverse Uncertainty Quantification (IUQ) methodologies are being developed for determining the probability density functions of relevant T-H codes input parameters, based on experimental data from Separate Effect Tests (SETs) experimental facilities. In practice, IUQ is challenged by the large range of variability of the experimental data in terms of Initial and Boundary Conditions (ICs & BCs), because the experimental campaigns are designed to cover the widest possible domain of conditions with the smallest number of experiments, so that same or similar ICs and BCs are seldomly repeated. To address this issue, we propose to use global sensitivity analysis, to tailor the IUQ on specific sub-regions described by segmented ICs & BCs domains. The methodology proposed is exemplified on two SETs, namely Sozzi-Sutherland and Super Moby Dick, whose experimental databases have been made available in the ATRIUM (Application Tests for Realization of Inverse Uncertainty quantification and validation Methodologies in thermal hydraulics) project promoted by the OECD/NEA/CSNI. The results obtained are superior to those of traditional IUQ methodologies for models highly sensitive to ICs & BCs.

1. Introduction

In the last decades, the Best Estimate Plus Uncertainty (BEPU) approach has been increasingly used in nuclear Thermal-Hydraulic (T-H) for performing the safety analysis of Nuclear Power Plants (NPPs) (D'Auria et al., 2012; Agnello et al., 2022; D'Auria et al., 2022). The objective is the verification that enough safety margin exists for some safety relevant Quantity of Interest (QoI) (e.g., fuel element Peak Cladding Temperature (PCT), hydrogen generation or core level), to demonstrate the capability of NPPs to withstand even the most demanding accidental conditions (Marquès et al., 2005; Unal et al., 2011; Sánchez et al., 2012). In general terms, safety margins quantification in the BEPU approach requires the assessment of the sources of uncertainty affecting the T-H code that can be due to the inherent variability of the phenomena (*aleatory uncertainty*) or to the lack of knowledge on its nature (*epistemic uncertainty*), and their forward propagation through simulation models so as to estimate the uncertainties of the QoI. Thanks to the research activity carried out in the last years (NEA, 2011, 2016), such forward Uncertainty Quantification (UQ) is being increasingly adopted in the nuclear sector. On the contrary, as pointed out in (IAEA, 2008, 2014), the characterization and quantification of the uncertainty in the model input parameters still raises concerns. In the current practice, the distributions of the input parameters of T-H codes are generally guessed by expert judgment or derived by literature analysis (Bersano et al., 2020; Hou et al., 2020). To limit the inherent subjectivity, Inverse Uncertainty Quantification (IUQ) methods relying on experimental data are being considered to find the input distributions that, propagated through the simulation model, have generated the data, to name a few (Wu et al., 2021):

- *Empirical (design-of-experiments):* the IUQ is based on heuristics and trial-and-error procedures, typically exploiting Monte Carlo sampling from plausible input distributions, running the model and infer ranges able to envelop the experimental data (Vinai et al., 2007). Examples of this class of methods are IPREM (Input Parameter Range

* Corresponding author.
E-mail address: francesco.dimaio@polimi.it (F. Di Maio).

https://doi.org/10.1016/j.anucene.2024.110791
Received 21 December 2023; Received in revised form 20 May 2024; Accepted 13 July 2024
Available online 25 July 2024
0306-4549/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Nomenclature

Acronyms

ATRIUM	Application Tests for Realization of Inverse Uncertainty quantification and validation Methodologies in thermal-hydraulics
BCs	Boundary Conditions
BE	Best Estimate
BEPU	Best Estimate Plus Uncertainty
CASUALIDAD	Code with the capability of Adjoint Sensitivity and Uncertainty AnaLysis by Internal Data ADjustment and assimilation
CIRCE	Calcul des Incertitudes Relatives aux Corrélation Elementaires
CSNI	Committee on the Safety of Nuclear Installations
DAA	Data Adjustment and Assimilation
ENEA	Agenzia nazionale per le nuove tecnologie l'energia e lo sviluppo economico sostenibile
FFTBM	Fast Fourier Transform Based Method
IBLOCA	Intermediate Break Loss Of Coolant Accident
ICs	Initial Conditions
IETs	Integral-Effect Tests
IPREM	Input Parameter Range Evaluation Methodology
IUQ	Inverse Uncertainty Quantification
KL	Kullback-Leibler divergence
LWR	Light Water Reactor
MAP	Maximum A Posteriori
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimation
MRAE	Mean Relative Absolute Error
NEA	Nuclear Energy Agency
NPPs	Nuclear Power Plants
OECD	Organization for Economic Cooperation and Development
PDF	Probability Density Function
PREMIUM	Post-BEMUSE Reflood Models Input Uncertainty Methods
QoI	Quantity of Interest
SA	Sensitivity Analysis
SAPIUM	Systematic Approach for Input Uncertainty quantification Methodology
SETs	Separate-Effect Tests
SMD	Super Moby Dick experiment
SNAP	Symbolic Nuclear Analysis Package
S-S	Sozzi-Sutherland experiment
T-H	Thermal Hydraulics
TI	Tolerance Interval
TRACE	TRAC/RELAP Advanced Computational Engine
UQ	Uncertainty Quantification
V&V	Verification and Validation
WGAMA	Working Group on the Analysis and Management of Accidents

Latin symbols

$\mathbb{E}[\cdot]$	Expectation operator
$\mathscr{G}(\cdot)$	Model Function
H	Hellinger sensitivity measure
$\mathscr{J}(\cdot)$	Posterior distribution
KL	Kullback-Leibler sensitivity measure
$\mathscr{L}(\cdot)$	Likelihood function
L/D	Length over diameter ratio
N	Dimension of quantification database
$\mathscr{N}(\cdot)$	Normal distribution
N_s	Number of forward simulations
$\mathscr{P}(\cdot)$	Prior distribution
P_0	Pressure at stagnation point
P	Number of model outputs/observations
$Q(\cdot)$	Quantile of a PDF
S	First order Sobol index
S^T	Total Sobol index
T_0	Temperature at stagnation point
V	Dimension of validation database
$Var(\cdot)$	Variance
X_0	Quality at stagnation point
Y_m	Model predictions database
Y_{obs}	Experimental observations database
Y_{val}	Experimental observations database (validation domain)
$det(\cdot)$	Matrix determinant
$exp[\cdot]$	Exponential function
$f(\cdot)$	Probability density function
h_0	Enthalpy at stagnation point
i	Calibration parameters index
j	ICs & BCs index
k	Sub-regions index
n	Observations points index
p	Model output/observations index
r	Number of observations exceeding the γ coverage extent
s	Forward simulation index
v	Validation database index
$\overline{x}$	ICs & BCs domain
$\overline{y}_m$	Model predictions (QoI)
$\overline{y}_m^{High}$	Model predictions (High values cluster)
$\overline{y}_m^{Low}$	Model predictions (Low values cluster)
$\overline{y}_m^{Medium}$	Model predictions (Medium values cluster)
$\overline{y}_{obs}$	Experimental observations
$\overline{y}_{val}$	Experimental observations (validation domain)

Greek symbols

$\dot{\Gamma}$	Steady state critical mass flux at the break section
Σ_{obs}	Covariance matrix of experimental observation error
β	Confidence level
γ	Coverage value
$\overline{\delta}$	Discrepancy model
δ_i	Borgonovo Delta sensitivity measure
ε_{obs}	Experimental observations noise
$\overline{\theta}$	Calibration parameters
ρ	Two-phase mixture density
$\rho_{g,sat}$	Saturated vapour density
$\rho_{l,sat}$	Saturated liquid density
σ_{obs}	Experimental observation error

Evaluation Methodology) (Kovtonyuk et al., 2017) based on sensitivity calculations performed through the Fast Fourier Transform Based Method (FFTBM) and the approach developed by (Domitr and Włostowski, 2021), in which UQ results are used to build the training database for a classification algorithm, that distinguishes the prediction capabilities of the uncertain parameters.

- *Frequentist (deterministic):* the IUQ is formulated as an optimization problem, in which Maximum Likelihood Estimation (MLE) and the Expectation Maximization (EM) algorithm are employed to find "best-fit" estimators of the model uncertain parameters. For example, relevant frequentist approaches are CIRCE (Calcul des Incertitudes Relatives aux Corrélation Elementaires) (De Crécy, 2001)

based on MLE and EM and CASUALIDAD (Code with the capability of Adjoint Sensitivity and Uncertainty AnaLysis by Internal Data ADjustment and assimilation) (Petruzzi, 2019) based on Data Adjustment and Assimilation (DAA).

- *Bayesian (probabilistic):* The IUQ is based on the application of the Bayes theorem and the Markov Chain Monte Carlo (MCMC) algorithm (Brooks et al., 2011); the input distributions can be obtained through the marginalization of the joint posterior distribution estimated with MCMC (Stuart, 2010). Among the Bayesian methods, we mention the Modular Bayesian (Wu and Kozlowski, 2017) and the Full Bayesian (Kennedy and O'Hagan, 2001) methods, which basically differ on the treatment of the model uncertainty (discrepancy) term (Wu et al., 2021). Usually, Bayesian approaches are quite computationally expensive because of MCMC: to address this issue they have been coupled with metamodelling techniques (e.g., Kriging, Polynomial Chaos Expansion) and dimensionality reduction methods like (e.g., Principal Component Analysis, Autoencoders) (Roma et al., 2021, 2022).

With regards to the IUQ of nuclear T-H Best Estimate (BE) models, the PREMIUM (Post-BEMUSE Reflood Models Input Uncertainty Methods) project (NEA, 2017) and the SAPIUM (Systematic Approach for Input Uncertainty quantification Methodology) project (Baccou et al., 2019, 2020) have been promoted by OECD/NEA/CSNI for benchmarking the different IUQ methods. In the following, we present a novel approach for IUQ, tailored on the data made available within the most recent ATRIUM (Application Tests for Realization of Inverse Uncertainty quantification and validation Methodologies in thermal hydraulics) project, in which a set of experimental databases related to two Separate Effect Test (SETs), i.e., Super Moby Dick (SMD) (Rousseau, 1987) and Sozzi-Sutherland (S-S) (Sozzi and Sutherland, 1975), have been identified in (Di Maio et al., 2024) as the most suitable for the IUQ of a T-H code for simulating the break flow in an Intermediate Break Loss Of Coolant Accident (IBLOCA) of an Integral Effect Test experiment (IETs), i.e., the LSTF IB-HL-01 (NEA, 2017). The T-H codes have been developed in TRAC/RELAP Advanced Computational Engine (TRACE) code (v5 patch 6), developed by (U.S. Nuclear Regulatory Commission, 2020) and the nodalization using the Symbolic Nuclear Analysis Package (SNAP) (Applied Programming Technology Inc., 2021).

The rationale behind the methodological development here presented is that SETs experimental facilities are an abundant source of data for the study of safety relevant T-H phenomena, hence performing the IUQ on SETs models can greatly increase the knowledge on the input parameters of nuclear T-H systems codes. Nevertheless, performing a satisfactory IUQ analysis on SETs is often challenged by:

- the large variability of the experimental campaigns in terms of Initial and Boundary Conditions (ICs & BCs);
- the non linearity of the phenomenon reproduced by the SET experiment, i.e., the model QoI is greatly affected by the variability of ICs & BCs;
- the sparsity of the experimental information, e.g., specific subregions of the ICs & BCs domain are better investigated than others;
- the sensitivity of the model response to small portions of the ICs & BCs domain;

To capture the localized behaviour of the model response (driven by specific ICs & BCs sub-regions), we propose a Bayesian IUQ approach (Kennedy and O'Hagan, 2001) driven by a global SA (Puppo et al., 2021), inspired by (Wu et al., 2019) where a clear link between SA and IUQ is conjectured: intuitively, only sensitive parameters to the QoI can be properly characterized by IUQ and, thus, SA can be used for dimensionality reduction. In addition to this, we claim that SA can enable the space segmentation of such inputs to a priori characterize the input domain in the sub-regions of the ICs & BCs that mostly affect the QoI. Without loss of generality, for the SA we adopt the Kullback-Leibler (KL) divergence (Park and Ahn, 1994) to guide a localized IUQ, which objective is obtaining parameter distributions conditioned only on the relevant sub-regions of the SET ICs & BCs domain, to ultimately better capture QoI behaviour in safety relevant regions.

The results obtained with the proposed method are compared with those provided by a traditional IUQ procedure, as suggested in the work of (Porter and Mousseau, 2019; Porter et al., 2019). It is shown that:

1) The results of the proposed approach are comparable with respect to the ones obtained with a traditional IUQ analysis for QoI not sensitive to the ICs & BCs domain.
2) The proposed approach provides better results when the QoI is sensitive to the ICs & BCs domain.

The structure of the paper is organized as follows: Section 2 presents the adopted SA and IUQ approaches; Section 3 introduces the case study and the developed TRACE BE models of SMD and S-S; Section 4 shows the results of the application of the approach to the models; also, a comparison of the results with those obtained with a traditional IUQ is described. Finally, in Section 5 conclusions are drawn.

2. The integrated SA-IUQ framework

In this work, we propose an integrated SA-IUQ methodology whose flowchart is sketched in Fig. 1:

The following steps are to be performed:

1) Collection of the SETs databases: the available experimental data regarding the phenomena of interest (i.e., the SETs experimental databases related to the QoI) are collected and organized in databases (Ghione, 2023).
2) Data adequacy evaluation: a preliminary analysis of the available experimental data is conducted to identify which SETs are the most suitable to be used within the following IUQ. This, for example and without loss of generality, can be formulated as a multi criteria decision making problem where the available experimental databases are evaluated accordingly to the representativeness and completeness criteria (Di Maio et al., 2024; Baccou et al., 2024). The outcome of this step is a ranking of the most adequate SETs experimental databases, that are, then partitioned into two independent databases: the quantification database (that will be used within the steps 3-4-5) and the validation database (that will be used in step 6).
3) T-H model development: the T-H BE models of the adequate SETs experiments are to be built (D'Auria et al., 2012; D'Auria et al., 2022).
4) Sensitivity Analysis (SA): distribution-based global SA approaches (Borgonovo, 2007; Di Maio et al., 2015; Puppo et al., 2021) are adopted to deal with the complexity and strong nonlinearities of the T-H BE models to evaluate whether the developed BE T-H models show any localized behaviour due to the dependency on ICs and BCs and to perform the input space segmentation.
5) IUQ: depending on the SA results, if a low sensitivity of the model to ICs and BCs is observed, a traditional IUQ analysis (step 5.a) is performed on the whole quantification database. Otherwise, when a high sensitivity of the model to ICs and BCs is observed, a local IUQ (step 5.b) is performed by exploiting the data according to the segmented input domain.
6) Validation: the goodness of the obtained IUQ results is verified through the comparison of the models predictions after the IUQ with the experimental data retained for the validation database.

For the sake of brevity, the interested reader may refer to (Baccou et al., 2020), (Di Maio et al., 2024), (D'Auria et al., 2012), for further insights on steps 1 to 3, respectively; instead, for the sake of clarity, in the following the developed SA methodology will be firstly presented in Section 2.1 along with the corresponding notation adopted throughout

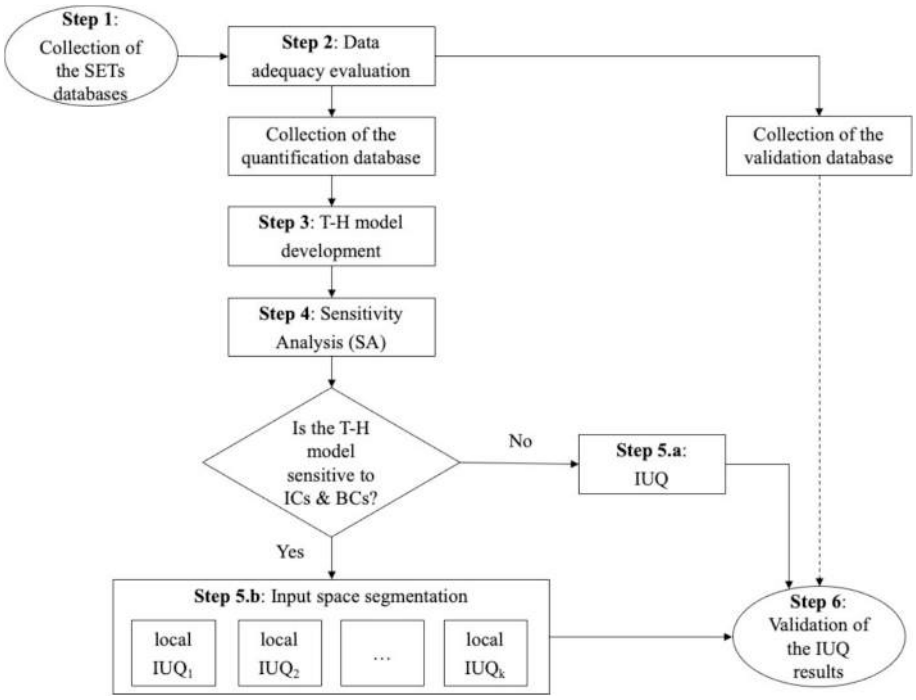


Fig. 1. The flowchart of the proposed SA-IUQ approach.

the paper. Then, the employed traditional IUQ framework will be illustrated in Section 2.2 alongside with the deduction of the developed (novel) local IUQ approach. Finally, the validation step is presented in Section 2.3. (Di Maio et al., 2015; Di Maio et al., 2024).

2.1. Sensitivity analysis (SA)

Let us define $\overline{y}_m(\overline{x},\overline{\theta}) = \mathscr{G}(\overline{x},\overline{\theta})$ as the output of a T-H BE code $\mathscr{G}$ simulating the physical behaviour of a SET experimental facility whose experimental observations are $\overline{y}_{obs}(\overline{x}) = \left[y_{obs}(\overline{x})_1, \cdots, y_{obs}(\overline{x})_p, \cdots, y_{obs}(\overline{x})_P\right]^T$, where $p = 1, 2, \cdots, P$, is the index of the measured variables. We can describe the code behaviour as a *black box* function $\mathscr{G}: \mathbb{R}^J \times \mathbb{R}^I \rightarrow \mathbb{R}^P$ linking the prediction of a vector of QoI $\overline{y}_m(\overline{x},\overline{\theta})$ $(P \times 1)$, with respect to a set of *design variables* $\overline{x} = \left[x_1, \cdots, x_j, \cdots, x_J\right]^T$ and *calibration parameters* $\overline{\theta} = [\theta_1, \cdots, \theta_i, \cdots, \theta_I]^T$, with $j = 1, 2, \cdots, J$ and $i = 1, 2, \cdots, I$, respectively. Sensitivity Analysis (SA) can be employed to rank the inputs $\overline{x}$ in terms of their relative contribution to the uncertainty of the model output $\overline{y}_m$ (Saltelli et al., 2004). In literature, it is customary to make the distinction between local SA (or one-at-a-time), in which just one input is varied typically in the neighborhood of its nominal values, and global SA, where the interactions between the inputs are considered along their whole support. In this work, we apply global SA methods with respect to the design variables $\overline{x}$ and the calibration parameters $\overline{\theta}$ to highlight and distinguish their peculiar roles in shaping the model response: intuitively, the QoI is expected to depend mainly on $\overline{x}$ (because the physics is driven by $\overline{x}$), whereas the dependency on $\overline{\theta}$ becomes relevant just when $\overline{x}$ is fixed (because of the $\overline{y}_m$ fine tuning role of $\overline{\theta}$). In other words, the model is evaluated at fixed combinations of $\overline{x}$ (i.e., leading to $Y_{obs}(\overline{x}) = \left[\overline{y}_{obs}(\overline{x}^1), \cdots, \overline{y}_{obs}(\overline{x}^n), \cdots, \overline{y}_{obs}(\overline{x}^N)\right]$ where $n = 1, 2, \cdots, N$ are different sets of ICs & BCs): thus, to evaluate the impact of $\overline{\theta}$ on the model response it is possible to fix $\overline{x}^n$ for each n-th experimental condition. In what follows we will, therefore, first define the SA procedure to rank the contribution of $\overline{x}$ to the QoI (localized behaviour) and then the SA procedure for ranking $\overline{\theta}$, conditioned on the segmented domain of $\overline{x}$. In particular, to perform the SA of the input $\overline{x}$, distribution-based SA techniques are adopted for ranking the sensitive inputs of multimodal model output distributions (Borgonovo and Plischke, 2016), as in the case for the T-H codes outputs f_{y_m}, due to the changes in the underlying physics and the respective implemented modelling solutions (e.g., empirical correlations, closure relationships) at varying values of the input $\overline{x}$. Two different distribution-based metrics, i.e., Hellinger (H) and Kullback-Leibler (KL), are compared to add robustness to the ranking results. Instead, for the SA of the model parameters $\overline{\theta}$, since the variability of the input $\overline{x}$ is removed (i.e., $\overline{x}$ is kept fixed), a unimodal output distribution is expected and, hence, in this case the application of a variance-based technique (i.e., Sobol Indexes) is appropriate and less computationally demanding. Nevertheless, the parameters $\overline{\theta}$ ranking results are compared with a moment independent measure (i.e., Borgonovo Delta) to validate the assumptions on the output unimodality.

2.1.1. Input space segmentation

Distribution-based (or moment independent) SA methods (Borgonovo and Plischke, 2016; Xiong et al., 2020b; Xiong et al., 2020a; Alibrandi et al., 2022) are particularly suitable to investigate the local (and global) behaviour of the model responses $\overline{y}_m(\overline{x}) = \mathscr{G}(\overline{x},\overline{\theta}^*)$, where $\overline{\theta}^*$ are the calibration parameters values frozen at their default value (e.g., prior mean), with respect to the variability of the design variables domain $\overline{x}$. Indeed, T-H codes usually exhibit non smooth and multimodal output Probability Density Function (PDF) f_{y_m}, because of the changes in the underlying physics reproduced by the simulation model at varying inputs $\overline{x}$ (e.g., ICs and BCs). Among the distribution-based methods, without loss of generality we adopt that developed in (Di Maio et al., 2015) that, first, clusters the QoI into K different response regions and, then, quantifies the sensitivity of the j-th input $\overline{x}$ in shaping the k-th response region of the output, i.e., $\overline{y}_m^k$, with $k = 1, 2, \cdots, K$, by measuring the distance between the unconditional input distribution f_{x_j} and the same input distribution conditioned on the k-th region of the output distribution $f_{x_j|\overline{y}_m^k}$, for example by H distance (Gibbs and Su, 2002):

$$H_j^k = \left[\frac{1}{2}\int\left(\sqrt{f_{x_j}} - \sqrt{f_{x_j|y_m^k}}\right)^2 dx_j\right]^{\frac{1}{2}} \quad (1)$$

and KL divergence (Park and Ahn, 1994):

$$KL_j^k\left(f_{x_j}\middle\|f_{x_j|y_m^k}\right) = \int f_{x_j}\log\left(\frac{f_{x_j}}{f_{x_j|y_m^k}}\right)dx_j \quad (2)$$

an example of a multimodal distribution with a scalar output y_m and $K = 3$ is plotted in Fig. 2.

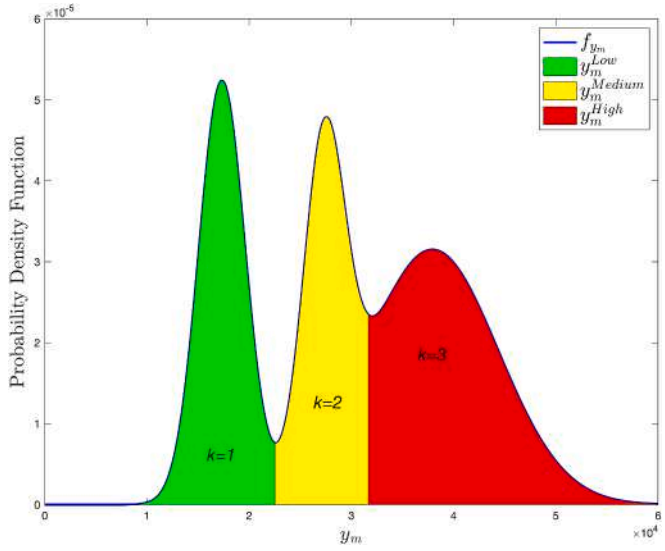


Fig. 2. Example of output PDF clustering with $K = 3$ (i.e., a low ($k = 1$), medium ($k = 2$) and high ($k = 3$) values region).

Each $f_{x_j|y_m^k}$ carries relevant information on the role of $\overline{x}$ on the model responsiveness to the ICs & BCs domain: the support of $f_{x_j|y_m^k}$ generates a segmented input space $\overline{x}^k$ that drives the model output in the k-th region, with $k = 1, 2, \cdots, K$, which maps the relationship between the inputs $\overline{x}^k$ and the output y_m^k. An example, for $K = 3$, two ICs inputs x_1 and x_2 and one output y_m is shown in Fig. 3.

The identified segmentation of x_1 and x_2 with respect to $k = 1, 2, 3$ can, thus, be used to identify the domain of the experimental data used both for the IUQ and the validation, i.e., IUQ for the $Y_{obs}^k(\overline{x}^k)$ and the validation using the database $Y_{val}^k(\overline{x}^k)$.

2.1.2. Verification of calibration of parameters

As customary (Perret et al., 2019, 2022; Wu et al., 2019; Roma et al., 2021) to verify the consistency of the IUQ results, an additional global SA analysis has to be performed on $\overline{\theta}$ prior PDF support in correspondence of each model response $Y_m(\overline{x}, \overline{\theta}) = \left[\mathcal{G}(\overline{x}^1, \overline{\theta}), \cdots, \mathcal{G}(\overline{x}^n, \overline{\theta}), \cdots, \mathcal{G}(\overline{x}^N, \overline{\theta})\right]$, where $n = 1, 2, \cdots, N$ is the index of varying ICs and BCs in the experimental database. The δ_i sensitivity measure (Borgonovo 2007) can be adopted for each i-th parameter of $\overline{\theta} = [\theta_1, \cdots, \theta_i, \cdots, \theta_I]^T$:

$$\delta_i = \frac{1}{2}\mathbb{E}_{\theta_i}\left[\int\left|f_{y_m} - f_{y_m|\theta_i}\right|dy\right] \quad (3)$$

where f_{y_m} and $f_{y_m|\theta_i}$ are the unconditional and θ_i −conditioned output distributions, respectively. For completeness, also a traditional variance-based sensitivity measure, i.e., Sobol indexes (Saltelli et al., 2004) has been adopted:

$$S_i = \frac{Var\theta_i\left(\mathbb{E}_{\theta_i}[y_m|\theta_i]\right)}{Var(y_m)} \quad (4)$$

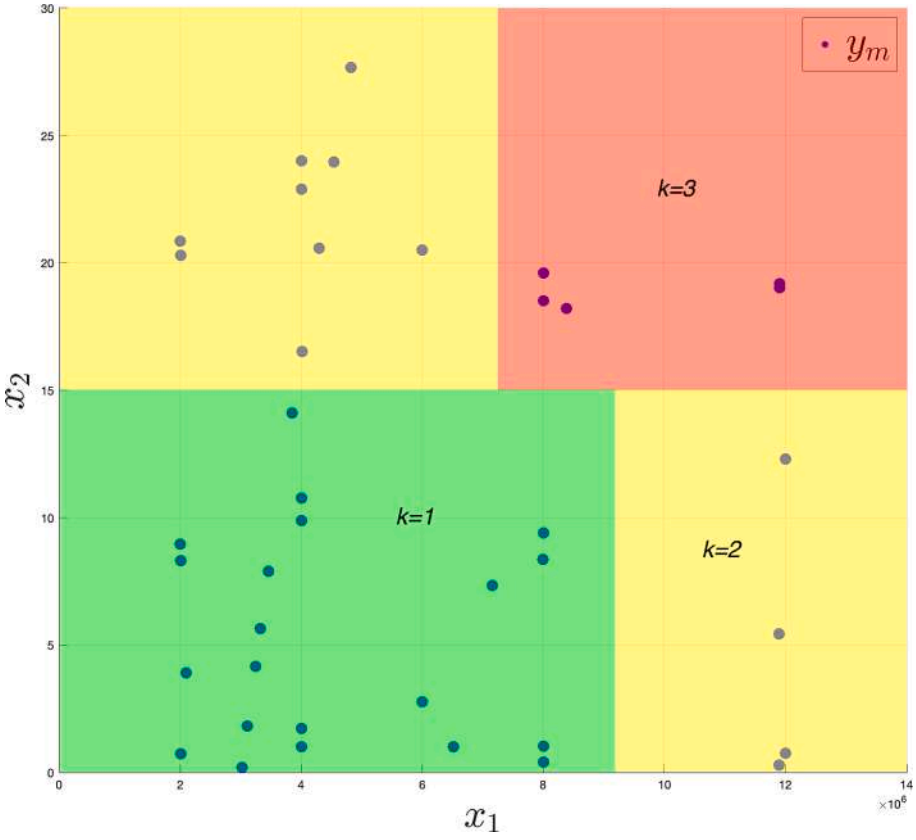


Fig. 3. Example of input space segmentation with $K = 3$ (i.e., a low, medium and high output values region).

$$S_i^T = \frac{\mathbb{E}_{\theta_{\sim i}}[Var_{\theta_i}(y_m|\theta_{\sim i})]}{Var(y_m)} \tag{5}$$

where Eq. (4) and Eq. (5) are also called the first order and total Sobol Index, respectively. It is worth noting that both Delta and Sobol Indexes can be generalized to the case of multivariate outputs ($P > 1$) (Xiao, Lu and Qin, 2016; Liu et al., 2019).

2.2. Inverse Uncertainty quantification (IUQ)

Design variables $\overline{x}$ carry the information regarding ICs, BCs and all the other observable inputs employed to describe the conditions or scenarios under which the considered phenomenon develops. Their uncertainty is usually easily estimated, whereas, the uncertainty of calibration parameters $\overline{\theta}$, that refer to physical quantities (e.g., material and fluid properties) or numerical coefficients (e.g., closure relations), requires IUQ (Wu et al., 2021). To estimate $\overline{\theta}$ we can build a relationship between code predictions and experimental data, i.e., the so called *model updating equation*:

$$\overline{y}_{obs}(\overline{x}) = \overline{y}_{real}(\overline{x}) + \varepsilon \tag{6}$$

where $\overline{y}_{real}(\overline{x})$ is the real (unknown) value of the measured QoI and $\varepsilon \sim \mathcal{N}(0, \Sigma_{obs})$ is an additive normally distributed experimental error, with covariance matrix:

$$\Sigma_{obs} = I\overline{\sigma}_{obs}^2 = \begin{bmatrix} \sigma_{obs_1}^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_{obs_P}^2 \end{bmatrix} \tag{7}$$

on the other hand, the code predictions $\overline{y}_m(\overline{x}, \overline{\theta})$ are related with $\overline{y}_{real}(\overline{x})$ as in Eq. (8):

$$\overline{y}_{real}(\overline{x}) = \overline{y}_m(\overline{x}, \overline{\theta}) + \delta(\overline{x}) \tag{8}$$

where $\delta(\overline{x})$ is called discrepancy, also known as model uncertainty, inadequacy or bias, and describes the mismatch between the simulation and reality (e.g., numerical errors, missing physics). The main purpose of $\delta(\overline{x})$ is to limit the probability of overfitting $\overline{\theta}$ during the IUQ. Nevertheless, its formal definition is still an open issue and various formulations have been proposed in literature (Ling et al., 2014), starting from gaussian distributed random variables to sophisticated kriging metamodels (Wu et al., 2018). By combining Eq. (6) and Eq. (8) it is possible to write Eq. (9):

$$\overline{y}_{obs}(\overline{x}) = \overline{y}_m(\overline{x}, \overline{\theta}) + \varepsilon + \delta(\overline{x}) \tag{9}$$

hereafter, for the sake of simplicity, we neglect $\delta(\overline{x})$: this is equivalent to consider the code $\mathcal{G}$ capable of perfectly reproducing the experimental observations for some given values of $\overline{\theta}$, that may lead to an underestimation of their uncertainty (overfitting). Nevertheless, in many applications, such as the ones investigated in this work, even though the experimental campaign is exhaustive, enough data to build a statistical model for $\delta(\overline{x})$ are not available. Since during the experimental campaigns multiple experiments are performed, it is convenient to extend Eq. (9) to the case of an array of measurements $Y_{obs}(\overline{x}) = [\overline{y}_{obs}(\overline{x}^1), \cdots, \overline{y}_{obs}(\overline{x}^n), \cdots, \overline{y}_{obs}(\overline{x}^N)]_{(N\times P)}$, where $\overline{y}_{obs}(\overline{x}^n)$ is the experimentally measured QoI under $n = 1, 2, \cdots, N$ varying ICs and BCs. Then, by collecting an array of model responses $Y_m(\overline{x}, \overline{\theta}) = [\mathcal{G}(\overline{x}^1, \overline{\theta}), \cdots, \mathcal{G}(\overline{x}^n, \overline{\theta}), \cdots, \mathcal{G}(\overline{x}^N, \overline{\theta})]_{(N\times P)}$ we obtain an extended formulation for the model updating equation as:

$$\begin{bmatrix} \overline{y}_{obs}(\overline{x}^1) \vdots \overline{y}_{obs}(\overline{x}^n) \vdots \overline{y}_{obs}(\overline{x}^N) \end{bmatrix} = \begin{bmatrix} \mathcal{G}(\overline{x}^1, \overline{\theta}) \vdots \mathcal{G}(\overline{x}^n, \overline{\theta}) \vdots \mathcal{G}(\overline{x}^N, \overline{\theta}) \end{bmatrix} + \begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \\ \vdots \\ \varepsilon^N \end{bmatrix} \tag{10}$$

2.2.1. Bayesian formulation of the IUQ problem

The purpose of IUQ is the estimation of the PDFs of $\overline{\theta}$ conditioned on the experimental data $Y_{obs}(\overline{x})$, hereafter Y_{obs}, which is the joint posterior PDF $\mathcal{J}(\overline{\theta}|Y_{obs})$. Indeed, by applying the Bayes rule of Eq. (11).

$$\mathcal{J}(\overline{\theta}|Y_{obs}) = \frac{\mathcal{L}(Y_{obs}|\overline{\theta})\mathcal{P}(\overline{\theta})}{\int \mathcal{L}(Y_{obs}|\overline{\theta})\mathcal{P}(\overline{\theta})}d\overline{\theta} \tag{11}$$

it is possible to obtain $\mathcal{J}(\overline{\theta}|Y_{obs})$ as the product of the joint prior PDF $\mathcal{P}(\overline{\theta})$ and the likelihood function $\mathcal{L}(Y_{obs}|\overline{\theta})$. The posterior and prior distributions represent the PDFs of $\overline{\theta}$ after and before the experimental data are observed, respectively. Whereas, the likelihood function is the probability of observing Y_{obs} given a particular value of $\overline{\theta}$. Then, from Eq. (10) and by assuming zero mean Gaussian distributed measurement error (i.e., $\varepsilon \sim \mathcal{N}(0, \Sigma_{obs})$) the likelihood function assumes the form:

$$\mathcal{L}(Y_{obs}|\overline{\theta}) = \prod_{n=1}^{N} \frac{1}{\sqrt{(2\pi)^P det(\Sigma_{obs})}} exp\Big[-\frac{1}{2}[\overline{y}_{obs}(\overline{x}^n) - \mathcal{G}(\overline{x}^n, \overline{\theta})]^T (\Sigma_{obs})^{-1} [\overline{y}_{obs}(\overline{x}^n) - \mathcal{G}(\overline{x}^n, \overline{\theta})] \Big] \tag{12}$$

where N is the number of independent experimental measurements and Σ_{obs} is the $(P \times P)$ covariance matrix for the measurement error. The denominator of Eq. (11) is the normalization constant of the posterior distribution, usually very difficult to calculate due to the complexity of the integral. To address this issue, a MCMC algorithm can be adopted for sampling from distributions known only up to a normalization constant (i.e., $\int \mathcal{L}(Y_{obs}|\overline{\theta})\mathcal{P}(\overline{\theta})\, d\overline{\theta}$) (Brooks et al., 2011). Then, MCMC samples can be used to reconstruct an estimated posterior distribution $\mathcal{J}(\widehat{\overline{\theta}}|Y_{obs})$. The main drawback of MCMC is that it usually takes thousands of iterations (i.e., forward model evaluations) to reach convergence, thus, for complex T-H models the problem becomes computationally intractable. To reduce the computational burden metamodelling and dimensionality reduction techniques can be adopted (Wu and Kozlowski, 2017; Roma et al., 2021, 2022). In this work, the developed models are simple enough to not require metamodelling. Finally, exploiting MCMC, we can estimate the posterior distribution as:

$$\mathcal{J}(\widehat{\overline{\theta}}|Y_{obs}) \propto \mathcal{P}(\overline{\theta}) \prod_{n=1}^{N} \frac{1}{\sqrt{(2\pi)^P det(\Sigma_{obs})}} exp\Big[-\frac{1}{2}[\overline{y}_{obs}(\overline{x}^n) - \mathcal{G}(\overline{x}^n, \overline{\theta})]^T (\Sigma_{obs})^{-1} [\overline{y}_{obs}(\overline{x}^n) - \mathcal{G}(\overline{x}^n, \overline{\theta})] \Big] \tag{13}$$

2.2.2. The local IUQ methodology

Through SA (see Section 2.1) the ICs & BCs domain of the SETs databases can be objectively segmented: then, for each k-th region of the $\overline{x}$ domain the traditional IUQ procedure presented in Section 2.2.1 can be performed independently just by exploiting the subset of related experimental data $Y_{obs}^k(\overline{x}^k)$, with $k = 1, 2, ..., K$. The outcome of this so called "local" IUQ analysis are K different posterior PDFs of $\widehat{\overline{\theta}}^k$, which are conditioned on the segmented domain $\overline{x}^k$, i.e., $\widehat{\overline{\theta}}^k|\overline{x}^k$. The conjecture beyond the adoption of this approach is that the estimated $\mathcal{J}\left(\widehat{\overline{\theta}}^k|Y_{obs}^k\right)$ should, in principle, be capable of better describe the localized behaviour of the model in the k-th region. Indeed, following this methodology,

only the data strictly related to the k-th response region of the model are used to tune the $\overline{\theta}$ parameters. Clearly, this procedure seems promising only in the case of highly sensitive models to the $\overline{x}$ domain, because otherwise, there is not a justifiable reason to use just a subset of $Y_{obs}(\overline{x})$ for the IUQ instead of the whole database.

2.3. IUQ reference verification and validation

Once the joint posterior PDF $\mathcal{J}(\widehat{\overline{\theta}}|Y_{obs})$ has been estimated, the validity of the obtained results must be checked accordingly to the Verification and Validation (V&V) framework (Trucano et al., 2006; Oberkampf and Trucano, 2007; Oberkampf and Roy, 2010). In the following, for Verification we refer to the evaluation of the capability of $\widehat{\overline{\theta}}$ to correctly represent the experimental observations $Y_{obs}(\overline{x})$ (i.e., the reference observations) when their uncertainty is propagated through the model $\mathcal{G}(\overline{x},\widehat{\overline{\theta}})$. Whereas, for Validation we refer to the uncertainty propagated for new (unknown) data collected for different $\overline{x}$. To perform this latter objective, a portion of the available experimental data is typically retained from the source database to build an independent (validation) database, i.e., $Y_{val}(\overline{x}) = \left[\overline{y}_{val}(\overline{x}^1),\cdots,\overline{y}_{val}(\overline{x}^v),\cdots,\overline{y}_{val}(\overline{x}^V)\right]_{(V\times P)}$, where $\overline{y}_{val}(\overline{x}^v)$ is the vector of experimentally measured QoI at $v = 1,2,\ldots,V$ varying ICs and BCs. Then, the model is evaluated by sampling from the posterior $\mathcal{J}(\widehat{\overline{\theta}}|Y_{obs})$ and an array of model responses is collected, i.e., $Y_m(\overline{x},\widehat{\overline{\theta}}) = \left[\mathcal{G}(\overline{x}^1,\widehat{\overline{\theta}}),\cdots,\mathcal{G}(\overline{x}^v,\widehat{\overline{\theta}}),\cdots,\mathcal{G}(\overline{x}^V,\widehat{\overline{\theta}})\right]_{(V\times P)}$. In the following, for notational simplicity, the case of a scalar output $(P = 1)$ will be assumed, i.e., $Y_m(\overline{x},\widehat{\overline{\theta}})$ is a $(V\times 1)$ array collecting the model responses $y_m(\overline{x}^v,\widehat{\overline{\theta}}) = \mathcal{G}(\overline{x}^v,\widehat{\overline{\theta}})$ at every v-ICs & BCs. The number of samples $s = 1,\ldots,N_s$ extracted from $\mathcal{J}(\widehat{\overline{\theta}}|Y_{obs})$ can vary according to the model complexity and the corresponding computational requirements. Typically, the estimation of the output PDF $f(y_m|\overline{x}^v,\widehat{\overline{\theta}}_s)$ is not feasible due to the large number of samples required, i.e., $N_s \longrightarrow \infty$. Therefore, for computationally intensive simulations, N_s could be decreased, by adopting order statistics (Wilks, 1941, 1942) to estimate a desired Tolerance Interval (TI) of $f(y_m|\overline{x}^v,\widehat{\overline{\theta}}_s)$. Indeed, the probability β that at least r (positive integer) observations within a random sample of size N_s are greater than the γ percentile of the unknown underlying distribution generating the samples can be written as (Zio, Di Maio and Tong, 2010):

$$\beta = \sum_{b=0}^{N_s-r}\binom{N_s}{b}\gamma^b(1-\gamma)^{N_s-b} \tag{14}$$

Then, by choosing $r = 2$, Eq. (14) reduces to the expression for the well known two sided tolerance interval (Nutt and Wallis, 2004):

$$\beta = 1-\gamma^{N_s}-N_s(1-\gamma)\gamma^{N_s-1} \tag{15}$$

Solving Eq. (15) for the desired confidence level β and coverage value γ it is possible to define the needed N_s. To objectively quantify the difference between expected and predicted results we could adopt different validation metrics. Typically, validation metrics are calculated using only point values for each response, e.g., the output corresponding to the mean or the Maximum A Posteriori (MAP) of $\mathcal{J}(\widehat{\overline{\theta}}|Y_{obs})$. In this work, we calculate the Mean Relative Absolute Error (MRAE) for the v-th observation of Y_{val} with respect to the model prediction when $\widehat{\overline{\theta}}_s$ is sampled from $\mathcal{J}(\widehat{\overline{\theta}}|Y_{obs})$, with $s = 1,\ldots,N_s$. The same metric can be extended to the entire validation domain obtaining the global validation performance indicator of Eq. (16):

$$\|\eta\| = \frac{1}{VN_s}\sum_{v=1}^{V}\sum_{s=1}^{N_s}\left|\frac{y_{val}(\overline{x}^v)-y_m(\overline{x}^v,\widehat{\overline{\theta}}_s)}{y_{val}(\overline{x}^v)}\right| \tag{16}$$

3. Case study

3.1. The ATRIUM project

The ATRIUM project (Ghione, 2023) has been promoted by OECD/NEA/CSNI/WGAMA for advancing the methodologies of IUQ in the framework of BEPU modelling for NPPs safety assessment. The final goal of the project is performing the IUQ of the uncertain parameters affecting the modelling of the critical mass flux during an IBLOCA that might occur in the Integral-Test Effects facility (IET) LSFT IB-HL-01 (NEA, 2017) (see Table 1) by exploiting SETs facilities BE models. The necessity of using SETs instead of the IET is related to the impossibility of directly performing the IUQ on the IET model, due to both the complexity of the model and the lack of experimental data. The available SETs experimental datasets are listed in Table 2, in which the name of the set, the label utilized throughout the paper, the number of experimental points N and the geometric (pipe diameter D, length over diameter ratio L/D) and T-H (pressure P_0, temperature T_0, quality X_0, all at stagnation point, and the QoI, critical mass flux $\dot{\Gamma}$ in steady state conditions) properties are also given.

As said in the Introduction, through the data adequacy methodology developed in (Di Maio, Coscia and Zio, 2024) it has been possible to rank the SETs of Table 2 that are more suitable to model the phenomena of interest (i.e., IBLOCA in LSTF), namely S-S and SMD. An overview of the T-H BE TRACE models adopted is presented in Section 3.2 and Section

Table 1
Experimental data of IET for exercise 1 of ATRIUM (J. Baccou et al., 2024).

Experiment Name	Label	N_P	L/D(−)	D (mm)	P_0 (bar)	T_0 (°C)	X_0 (−)	G (kg/m^{2}s)
LSTF IB-HL-01	LSTF	–	15	–	10–155	177–327	–	1500–46000

Table 2
Available experimental datasets of SETs for exercise 1 of ATRIUM (J. Baccou et al., 2024).

Experiment Name	Label	N_P	L/D (−)	D (mm)	P_0 (bar)	T_0 (°C)	X_0 (−)	G (kg/m^{2}s)
Sozzi Sutherland	S-S N2	358	0–140	12.7	56.0–71.3	232–286	−0.0044–0.0065	17528–75824
	S-S N3	58	0	12.7	42.7–69.0	212–285	−0.0059–0.0060	33161–61226
	S-S N4	23	0	19	56.0–66.3	271–282	−0.0003–0.0099	29295–51266
Super Moby Dick	SMD Div	27	18	20	20–120.1	192.3–324.4	<0	15300–62200
	SMD Exp	12	20	20	20–120.1	191.5–323.6	<0	16100–61800
Marvi ken	Marv 13	1	3	200				<89200
	Marv 17	1	3,7	300	~ 50 transient	ΔT_{sub} ~ 31 (°C)	<0	<61700
	Marv 24	1	0,33	500				<59750

Table 3
Summary of the prior $\mathscr{P}(\overline{\theta})$ adopted,

Symbol	Name	Trace	Nominal Value	Prior Support	Prior Distribution
θ_1	Subcooled choked-flow multiplier	CHM12	1.0	[0.5–2.0]	Uniform
θ_2	Two-phase choked-flow multiplier	CHM22	1.0	[0.5–2.0]	Uniform
θ_3	Choking relaxation constant 1 for set 2	C1RC2	2.0	[0.01–5.0]	Uniform
θ_4	Choking relaxation constant 2 for set 2	C2RC2	1.0	[0.01–2.5]	Uniform

Source (Domitr and Włostowski, 2021).

3.3, respectively.

The framework presented in Section 2 has been implemented to perform the IUQ on the models of both SETs and the results, shown in Section 4, are compared with the ones obtained with the traditional IUQ methodology described in Section 2.2. The following assumptions hold:

- The analysis has been performed independently for SMD and S-S.
- The QoI y_m is a scalar quantity, i.e., the critical mass flux $\dot{\Gamma}$ at the break section in steady state conditions, for both SMD and S-S.
- For SMD, the data related to the divergent nozzle (SMD Div) have been selected for the quantification database (Y_{obs}) while the ones from the expansion nozzle (SMD Exp) have been used for validation (Y_{val}).
- For S-S, from the dataset related to nozzle N2, a subset of $N = 91$ points (the ones with L/D closer to LSTF, i.e., between 2.5 and 12.5) have been used for Y_{obs} while $V = 30$ points have been retained for Y_{val} (choosing those with L/D closer to LSTF, i.e., 1.5 and 20, among the remaining data).
- For SMD, $J = 2$ design variables $\overline{x} = [P_0, \Delta T_{sub}]^T$ (namely pressure (P_0) and the degree of subcooling (ΔT_{sub}) at the stagnation point) have been selected. Instead, for S-S $J = 5$ design variables $\overline{x} = [P_0, T_0, X_0, h_0, L/D]^T$ (pressure (P_0), temperature (T_0), quality (X_0), enthalpy (h_0) all at the stagnation point plus the length over diameter ratio (L/D)) have been chosen.
- The same set of $I = 4$ calibration parameters $\overline{\theta} = [\theta_1, \theta_2, \theta_3, \theta_4]^T$ has been chosen both for SMD and S-S. They are numerical coefficients of the TRACE critical flow model (see Table 3) and their prior ranges have been determined through expert judgment in the work of (Domitr and Włostowski, 2021).
- The SA related to the inputs $\overline{x}$ (see Section 2.1.1) has been performed on a small ensemble of simulations (less than 10^2) to demonstrate its applicability in the frequent case of highly demanding computational models.
- A $K = 3$ region input segmentation has been performed for both SMD and S-S.
- The SA on the $\overline{\theta}$ (see Section 2.1.2) has been performed through UQLab MATLAB Sensitivity Analysis package (Lamas et al., 2022) using 10^4 simulations for each model to obtain robust results.
- The Affine Invariant Ensemble Sampler (AIES) (Goodman and Weare, 2010) is used to generate twenty parallel chains (seeds or walkers) with 10^5 iterations and a burn in period of 50 %. We have adopted the UQLab MATLAB Bayesian Inversion package (Wagner et al., 2019) to implement the MCMC algorithm and to perform the overall IUQ analysis.
- For the validation analysis, the 95 % two sided TI of the $y_m(\overline{x}', \widehat{\overline{\theta}})$ has been estimated through $N_s = 93$ forward model evaluations sampling $\widehat{\overline{\theta}}$ form $\mathscr{J}(\widehat{\overline{\theta}}|Y_{obs})$. The value of N_s is obtained from Eq. (15) by fixing $\beta = \gamma = 0.95$ (Nutt and Wallis, 2004; Zio et al., 2010).

3.2. Super Moby Dick (SMD)

The Super Moby-Dick (SMD) experiments (Rousseau, 1987) were performed at the Commissariat à l'énergie atomique et aux énergies alternatives (CEA) between 1982 and 1983 in order to investigate the two-phase critical flowrate in the medium and high pressure range. The test facility is designed to produce steam-water two-phase critical flow in a test section under stable steady state conditions. The experimental test procedure consisted in obtaining the desired conditions (pressure and temperature) at the inlet of the test section. Then, maintaining these inlet conditions constant, the critical mass flux was reached by decreasing the pressure in the condenser. The critical flow was detected when no effect on the mass flux and on the pressure upstream of the throat could be observed decreasing further the condenser pressure. The critical flow was measured by a turbine flowmeter and its uncertainty, for values between 10,000 and 60,000 kg/m^2s, was estimated as ± 2 % for this experimental setup (Rousseau, 1987). The nodalization of the SMD experiment is shown in Fig. 4: the nozzle where the critical flow occurs is modelled with a pipe component initially convergent (four cells), then straight (ten cells) and finally a conical expansion (eleven cells); an inlet break component sets the inlet conditions of each test (i.e. pressure and temperature), while in the outlet break component the pressure is gradually reduced to achieve the critical flow; the choked flow model is activated only at the edge where critical flow is expected to occur.

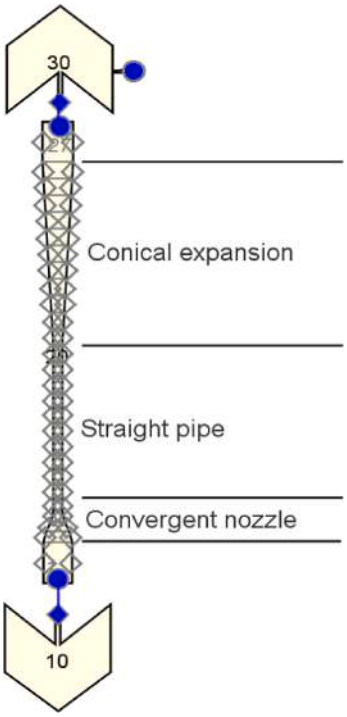


Fig. 4. TRACE nodalization of SMD nozzle developed through SNAP.

3.3. Sozzi-Sutherland (S-S)

The Sozzi-Sutherland experiments (Sozzi and Sutherland, 1975) were performed by General Electric in 1975. The goal was to obtain critical flow experimental data at typical conditions of LOCA. Subcooled and saturated water at relative high pressure (approximately 70 bar) was discharged from a vessel to the atmosphere through various flow

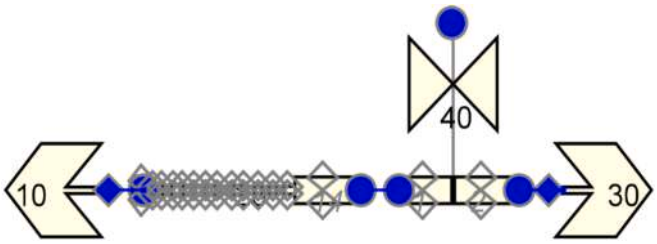


Fig. 5. TRACE nodalization of S-S N2 nozzle developed through SNAP.

geometries and critical flowrate measurements were obtained from the estimation of the rate of change of the fluid mass inventory in the vessel during the test. For the sake of simplicity and in absence of further information we have assumed an experimental error equal to the SMD measurements. The nodalization of the S-S nozzle N2 experiment is shown in Fig. 5. Also in this case only the nozzle where critical flow occurs was simulated with a horizontal pipe component. A break component fixes the inlet boundary conditions of each test (i.e., pressure and temperature for the subcooled cases or pressure and void fraction for the saturated cases). The break component at the outlet provides fixed atmospheric conditions and the critical flow is achieved by the opening of a valve component. Similarly to SMD, the choked flow model is activated only at the edge where critical flow is expected to occur. In this case, three cells were used to model the convergent inlet section whereas the following straight pipe nodalization was adjusted to simulate the various configurations, having a different length, maintaining as much as possible a similar L/D for each cell.

4. Results and discussion

In this Section the results of the application of the framework proposed in Section 2 to the case study illustrated in Section 3 are presented.

4.1. SMD

In Fig. 6, it is shown the discretized response surface of the SMD model with respect to the inputs $\overline{x} = [P_0, \Delta T_{sub}]^T$ along with the quantification $Y_{obs}(\overline{x})$ and the validation $Y_{val}(\overline{x})$ database. The procedure of Section 2.1.1 is applied on the $K = 3$ regions (namely Low ($k = 1$), Medium ($k = 2$), High ($k = 3$) values y_m) which will be identified by the same colour scheme (i.e., green, yellow, red) throughout the paper.

4.1.1. SMD: SA

In Fig. 7 the results of the distribution based SA are presented; in blue are plotted the unconditioned input distributions f_{x_1} and f_{x_2} (**(a,b)** upper portion) while the conditioned ones $f_{x_1|y_m^k}$ and $f_{x_2|y_m^k}$ (**(a,b)** lower portion) follow the three regions colour scheme. As shown by Fig. 7, there is a dependence of the model response by $\overline{x}$, because there are some visible differences between the f_{x_j} and $f_{x_j|y_m^k}$, especially for the High region (i.e., the red one). Nevertheless, the sensitivity measures reported in Fig. 7 **(c)** for each region, i.e., Eq. (11) and Eq. (12), do not show any clearly dominant localized behaviour for the model, since there is no region with the highest sensitivity measure for all the $\overline{x}$ (i.e., there is no subset of the input domain that univocally drives the model response into one specific k-th region).

Despite the SA for SMD reveals a weak dependency on ICs & BCs, the input space segmentation (see Section 2.1.1) is performed to obtain the $\overline{x}^k$ sub-regions, and the sensitivity of $\overline{\theta}$ is performed (see Section 2.1.2) accounting for such segmentation. The outcomes of the analysis are shown in Fig. 8, where the same colours of the K models are used ($k = 1$ green, $k = 2$ yellow, $k = 3$ red): these highlight large δ and Sobol Indexes for θ_1 and θ_2, and that δ and Sobol Indexes results are providing qualitatively the same results. The similarity and agreement between the results obtained from a variance-based (i.e., Sobol Indexes) and a distribution-based (i.e., Borgonovo Delta) sensitivity measure confirm the correctness of the assumption of unimodality of the output PDFs f_{y_m}

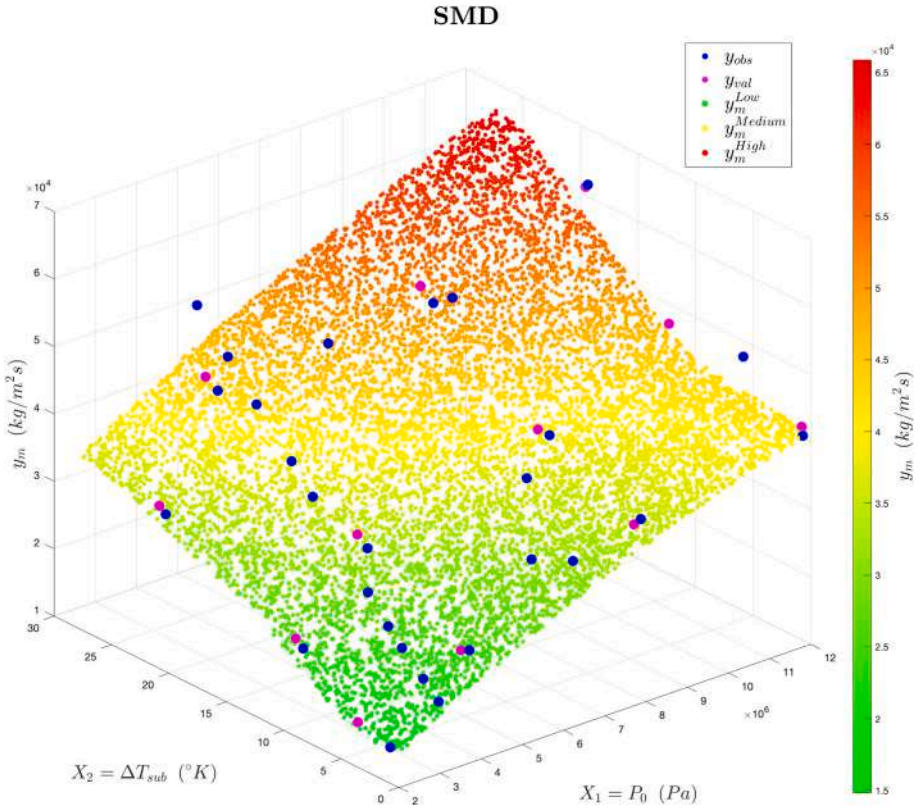
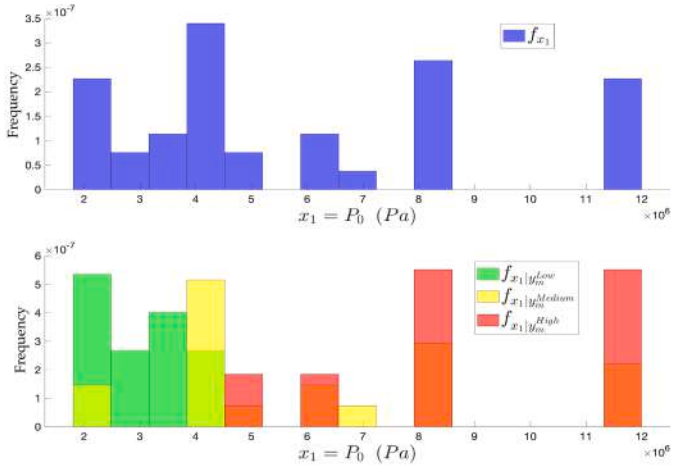
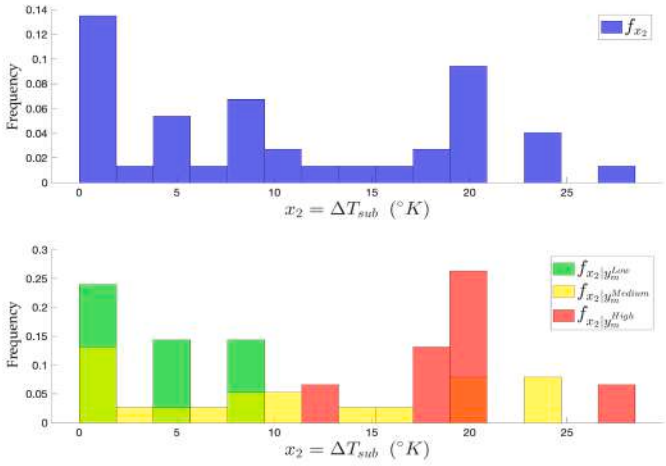


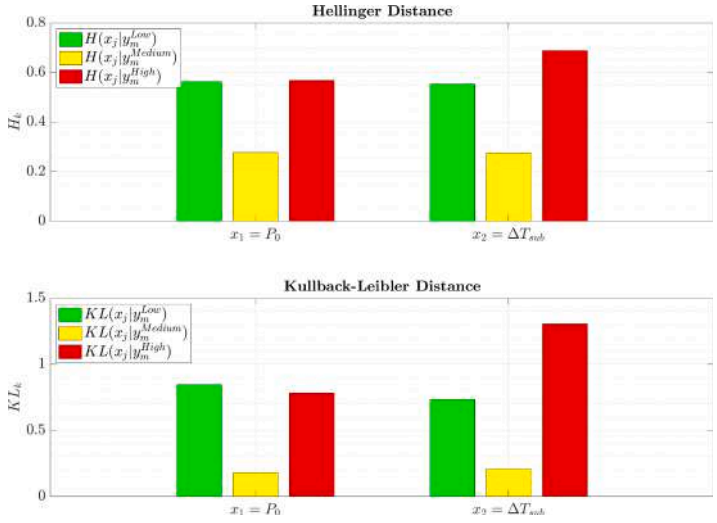
Fig. 6. Plot of the response surface of SMD as function of the ICs.



(a) f_{x_1} and $f_{x_1|y_m^k}$

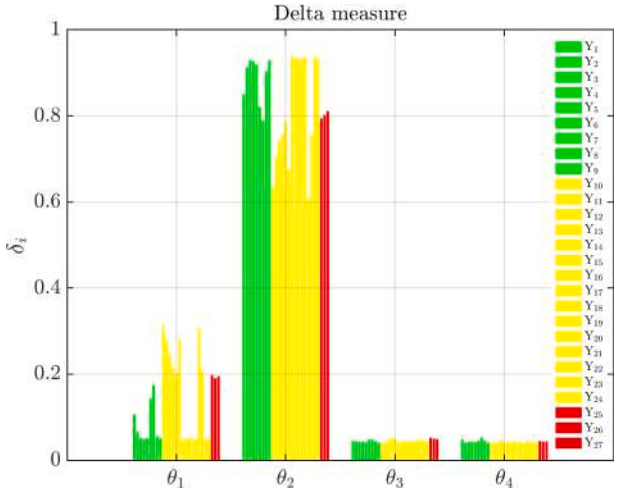


(b) f_{x_2} and $f_{x_2|y_m^k}$

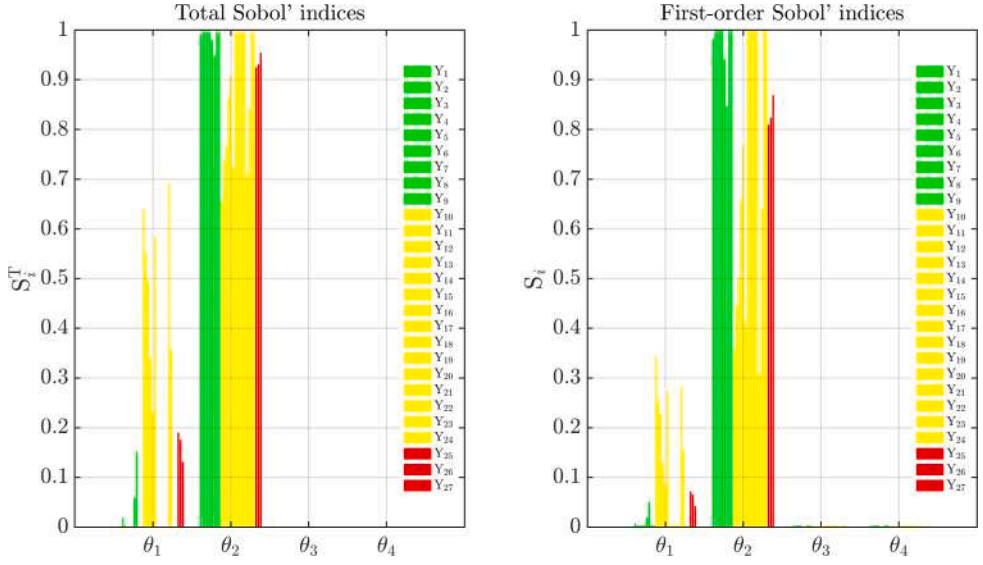


(c) KL and H sensitivity measures

Fig. 7. SA results for SMD.

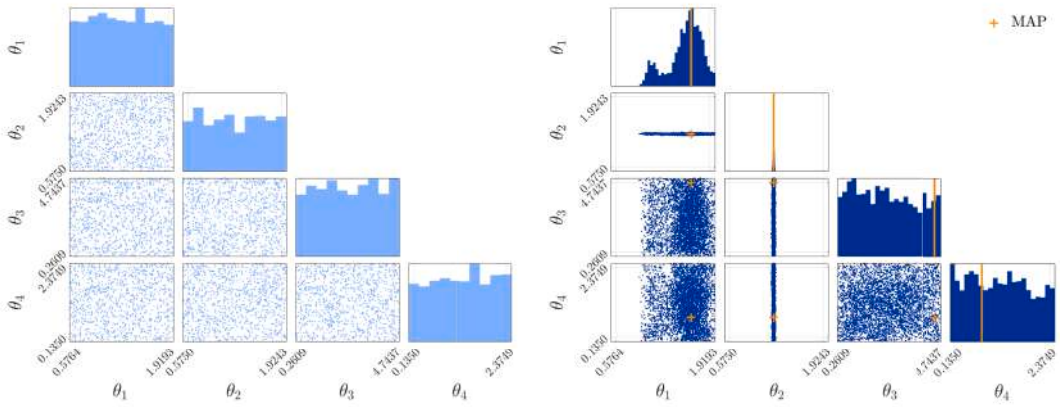


(a) Borgonovo δ

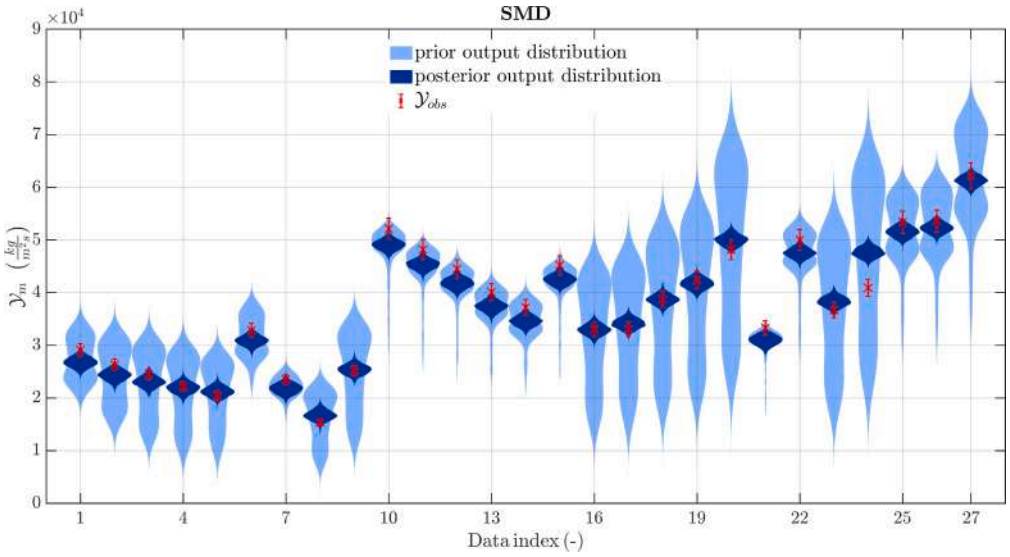


(b) Sobol indexes

Fig. 8. Sensitivity of $\overline{\theta}$ for SMD.



(a) Prior $\mathcal{P}(\overline{\theta})$ **(b)** Posterior $\mathcal{J}(\hat{\overline{\theta}})$



(c) IUQ Reference Verification

Fig. 9. Results of traditional IUQ analysis for SMD.

Table 4
Posterior PDFs statistics estimated through the traditional IUQ for SMD.

Parameter	Mean	Variance	MAP	[*Q(0.05)-Q(0.95)*]
$\widehat{\theta}_1$	1.5650	0.0582	1.6511	[1.0305–1.9355]
$\widehat{\theta}_2$	1.2135	$1.0211\ 10^{-04}$	1.2132	[1.1935–1.2332]
$\widehat{\theta}_3$	2.3838	2.0666	4.6986	[0.1432–4.8692]
$\widehat{\theta}_4$	1.1925	0.5119	0.7670	[0.0588–2.4312]

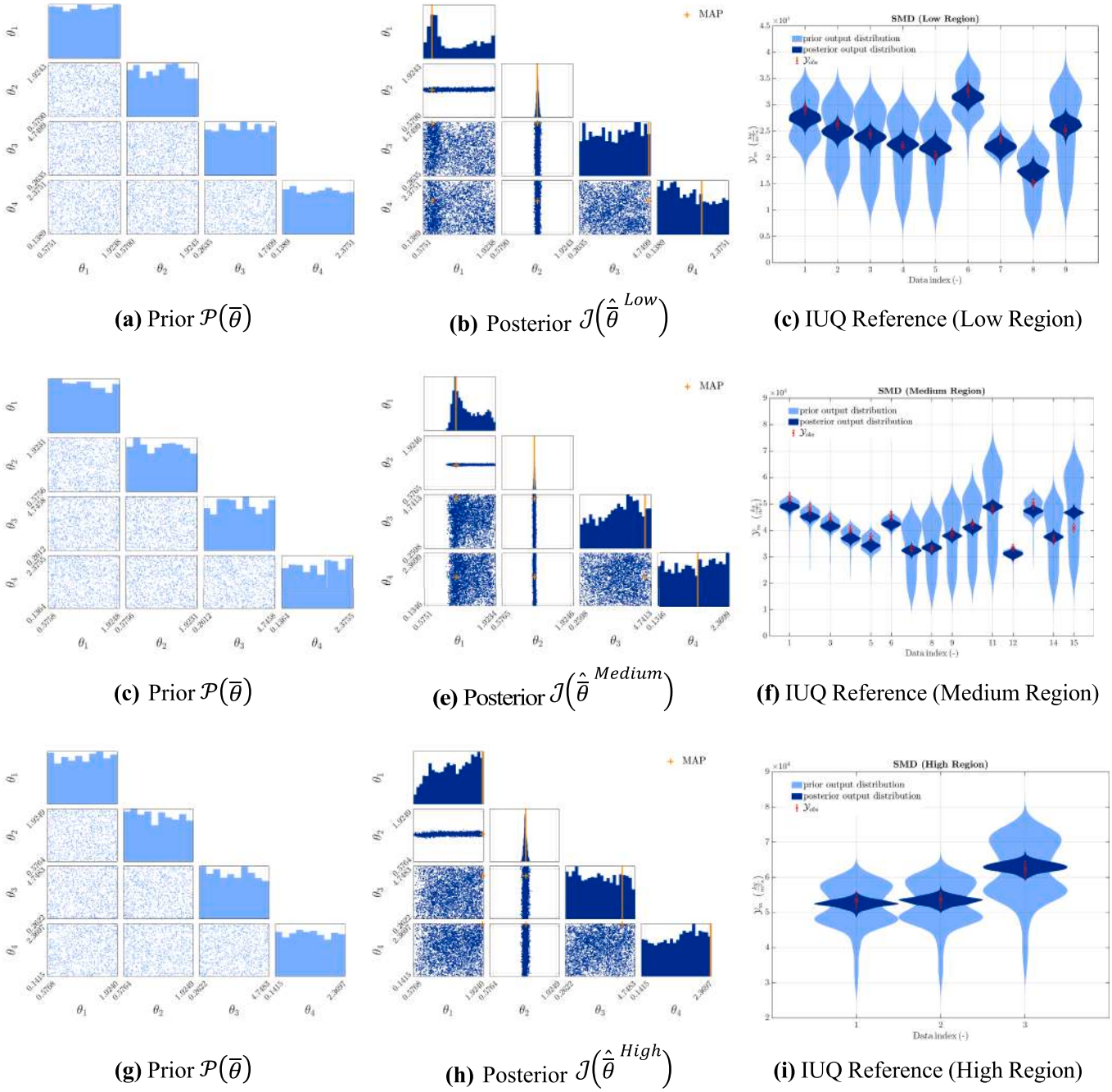
and $f_{y_m|\theta_i}$, when evaluated with fixed input $\bar{x}$.

4.1.2. SMD: IUQ

In Fig. 9, the quantitative results of the traditional IUQ procedure are presented (i.e., a not segmented input domain is adopted): Fig. 9 **(a)**,**(b)** show the prior $\mathcal{P}(\overline{\theta})$ and the posterior $\mathcal{J}(\widehat{\overline{\theta}})$ PDFs of $\overline{\theta}$ (on the main diagonal of the plot are shown the marginal PDFs (summarized in Table 4) while on the inner boxes the pairwise projections of the joint posterior PDF are plotted, respectively); Fig. 9 **(c)** shows the outcomes of the IUQ reference verification by forward propagation of the $\mathcal{P}(\overline{\theta})$ and $\mathcal{J}(\widehat{\overline{\theta}})$ with respect to the Y_{obs} (i.e., the red crosses). A violin plot (Hintze and Nelson, 1998) is obtained by fitting a set of 10^3 model realizations through kernel density estimation with a reflection boundary correction (Silverman, 1986). In Table 4, the following statistics of the marginalized joint posterior PDF $\mathcal{J}(\widehat{\overline{\theta}})$ are presented: Mean, Variance, MAP and the $[Q(0.05) - Q(0.95)]$ quantiles of the PDF. All the statistics are calculated based on 10^3 samples of $\mathcal{J}(\widehat{\overline{\theta}})$.

Similarly, in Fig. 10 and Table 5 the same results are summarized for the local IUQ analysis with $K = 3$: the prior, posterior and the output distributions after the IUQ are shown for each k-th region, respectively, in **(a)**, **(b)**, **(c)** for the low region, **(d)**, **(e)**, **(f)** for the medium region, **(g)**, **(h)**, **(i)** for the high region.

It can be seen that the results shown in Fig. 9 and Fig. 10 are similar and both satisfactory (i.e., the estimated $\mathcal{J}(\widehat{\overline{\theta}})$ reproduce the experimental observations used for the IUQ quite well adopting both the traditional (see Fig. 9 **(c)**) or the local IUQ approach (see Fig. 10 **(c)**, **(f)**, **(i)**, respectively), hence for SMD, the input space segmentation presents no substantial benefit. Moreover the highly sensitive parameters θ_1 and



(a) Prior $\mathcal{P}(\overline{\theta})$ **(b)** Posterior $\mathcal{J}\left(\hat{\overline{\theta}}^{Low}\right)$ **(c)** IUQ Reference (Low Region)

(c) Prior $\mathcal{P}(\overline{\theta})$ **(e)** Posterior $\mathcal{J}\left(\hat{\overline{\theta}}^{Medium}\right)$ **(f)** IUQ Reference (Medium Region)

(g) Prior $\mathcal{P}(\overline{\theta})$ **(h)** Posterior $\mathcal{J}\left(\hat{\overline{\theta}}^{High}\right)$ **(i)** IUQ Reference (High Region)

Fig. 10. Results of local IUQ analysis for SMD.

Table 5
Posterior PDFs statistics estimated through the local IUQ for SMD.

Parameter	Mean	Variance	MAP	[*Q(0.05)-Q(0.95)*]
$\widehat{\theta}_1^{Low}$	1.1530	0.2231	0.6830	[0.5519–1.9758]
$\widehat{\theta}_2^{Low}$	1.2627	6.2150 10^{-04}	1.2562	[1.2140–1.3150]
$\widehat{\theta}_3^{Low}$	2.5776	2.1252	4.8953	[0.1401–4.8664]
$\widehat{\theta}_4^{Low}$	1.1441	0.5163	1.5656	[0.0708–2.4301]
$\widehat{\theta}_1^{Medium}$	1.1422	0.0759	1.1753	[1.0601–1.9441]
$\widehat{\theta}_2^{Medium}$	1.2627	1.7478 10^{-04}	1.1806	[1.1582–1.2095]
$\widehat{\theta}_3^{Medium}$	2.5776	1.8805	4.5731	[0.1912–4.8960]
$\widehat{\theta}_4^{Medium}$	1.1441	0.5378	1.4065	[0.0671–2.4634]
$\widehat{\theta}_1^{High}$	1.5115	0.1676	1.9809	[0.6077–1.9674]
$\widehat{\theta}_2^{High}$	1.1577	0.0010	1.2815	[1.2092–1.3376]
$\widehat{\theta}_3^{High}$	2.3707	2.0782	4.0843	[0.1316–4.8219]
$\widehat{\theta}_4^{High}$	1.2962	0.5101	2.4730	[0.0804–2.4368]

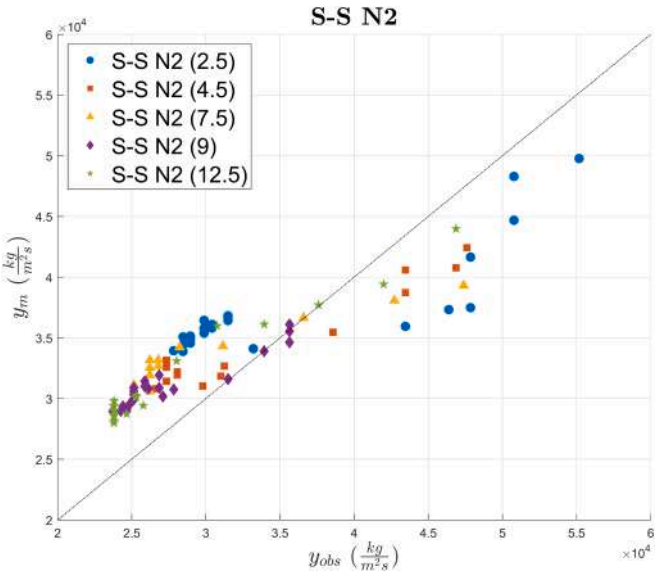
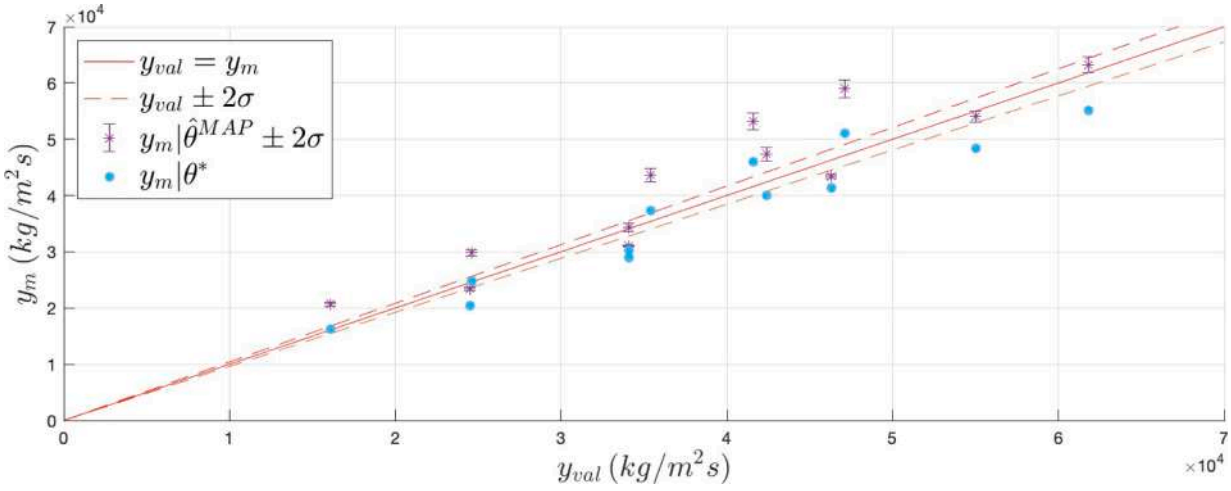
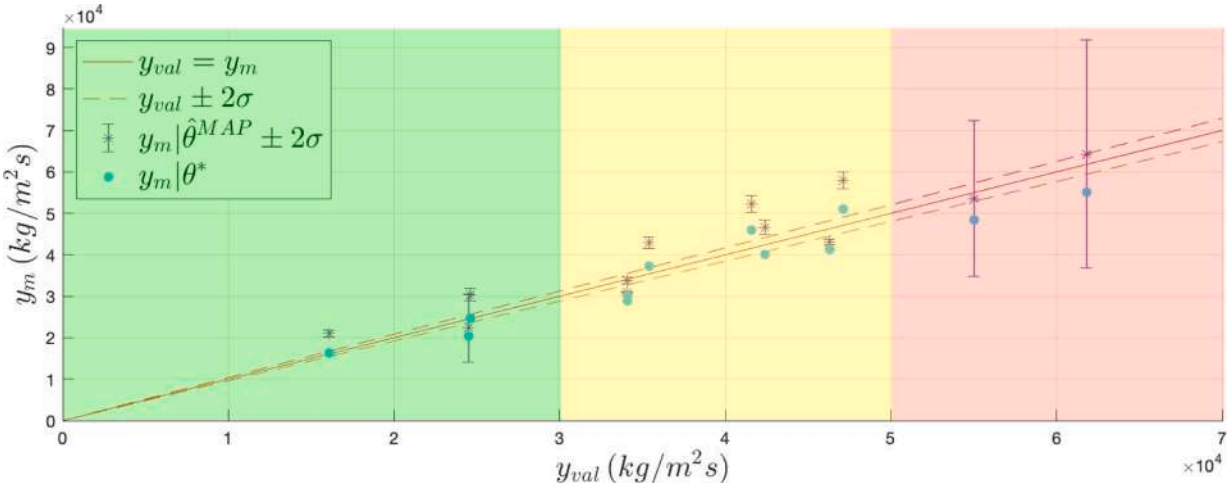


Fig. 12. Reference calculation results of S-S model before the IUQ analysis.

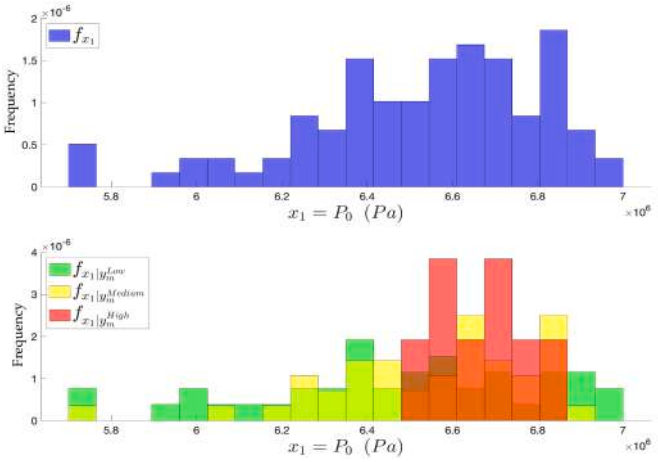


(a) IUQ

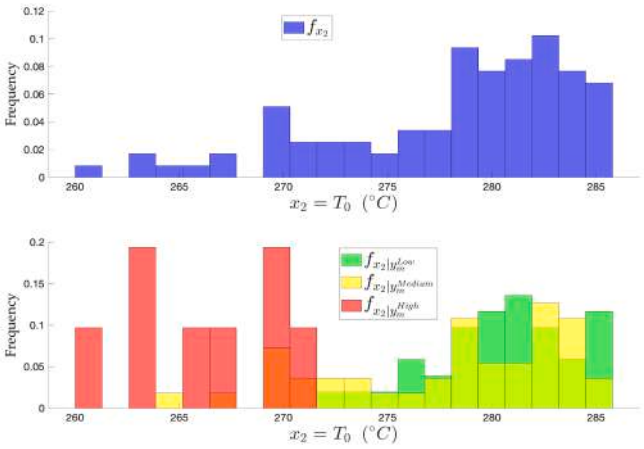


(b) Local IUQ

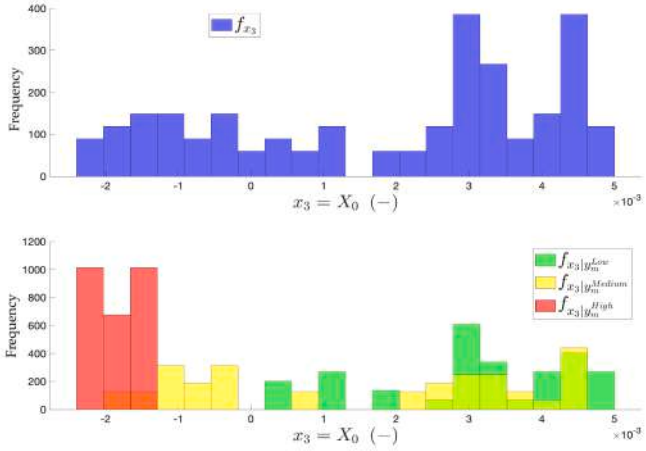
Fig. 11. Comparison of validation results of the traditional and the local IUQ for SMD.



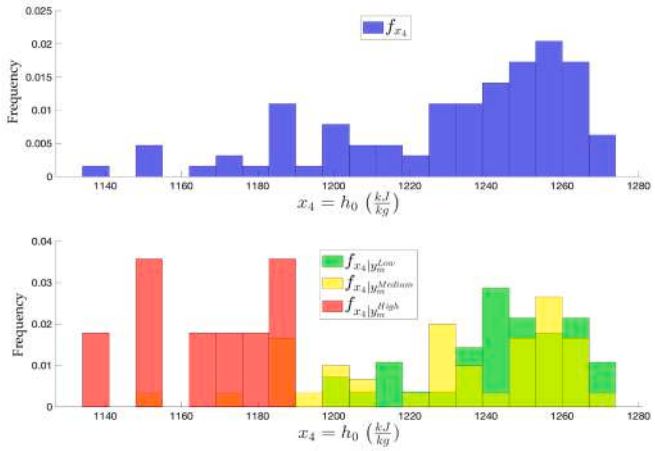
(a) f_{x_1} and $f_{x_1|y_m^k}$



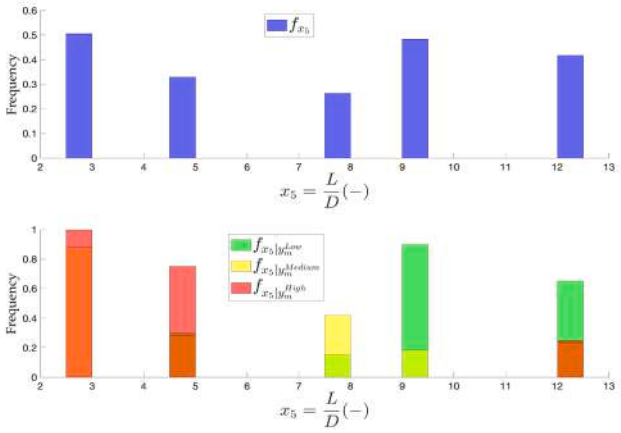
(b) f_{x_2} and $f_{x_2|y_m^k}$



(a) f_{x_3} and $f_{x_3|y_m^k}$



(d) f_{x_4} and $f_{x_4|y_m^k}$



(c) f_{x_5} and $f_{x_5|y_m^k}$

Fig. 13. SA results for S-S.

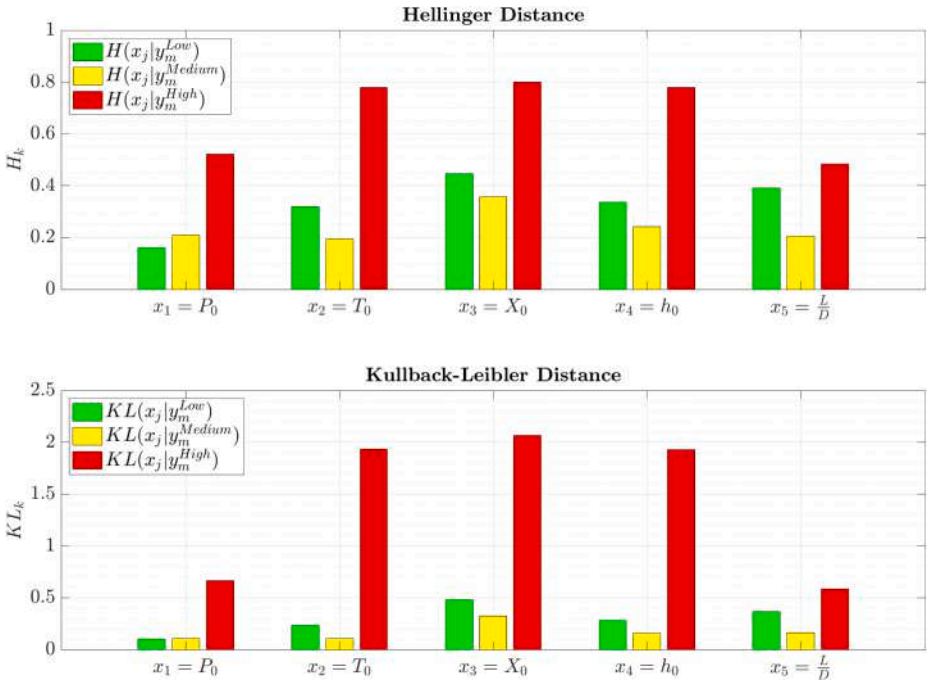


Fig. 14. SA quantitative results for S-S.

θ_2 show substantially different posterior PDFs with respect to their prior. Whereas, as expected by (Wu et al., 2019), θ_3 and θ_4 are basically left unchanged due to their low sensitivity (i.e., they are not identifiable).

4.1.3. SMD: Validation

In Fig. 11, the results of the validation procedure for the traditional IUQ **(a)** and the local IUQ **(b)** are presented. The bisector of the first quadrant is plotted in red with a $\sigma_{obs} = \pm 2\%$ (i.e., the dotted line) and represents the condition $y_m = y_{val}$ (i.e., perfect calibration). Then, for each validation point, the model responses with $\overline{\theta}$ fixed at the default prior values $\overline{\theta}^*$ (i.e., the blue dots) are compared with the two sided 95% TI of the model responses when sampling from the estimated $\mathcal{J}(\widehat{\overline{\theta}})$ and centered on the model response at the $\widehat{\overline{\theta}}^{MAP}$ (i.e., the purple asterisk).

From a visual comparison, the results of the two approaches are quite similar, because the overall trend of the responses obtained with the different posterior PDFs is the same. In both cases, few points are enveloped within the TI: it can be justified as the side effect of the fact that the validation is performed on data referring to different nozzle shapes (i.e., SMD Exp) from those used for the IUQ analysis (i.e., SMD Div). To measure the differences between the two approaches we have employed the metric of Eq. (16) to quantify the MRAE between the point estimates of the two approaches (i.e., the purple asterisks). Since the error is just around $2-5\%$ in every region we can claim that there are not substantial differences between the two methods. The only major dissimilarity is the width of the estimated TI, indeed, the input space segmentation shrinks the extension of experimental datasets (e.g., Fig. 11 **(b)** red region). Thus, the resulting posterior distributions estimated with the local IUQ are affected by a greater uncertainty due to the lower available evidence. This behaviour is clearly linked to the epistemic component of the $\overline{\theta}$ uncertainty, which is reduced by increasing the number of observations (Higdon et al., 2008).

4.2. S-S

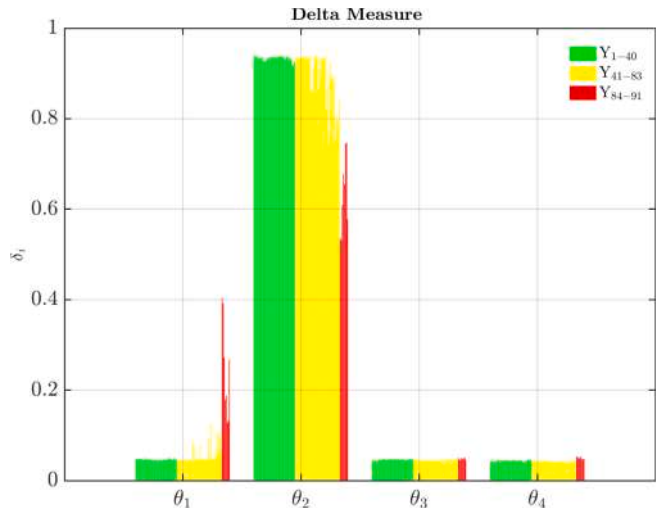
In Fig. 12, it is plotted the response of the S-S model (before IUQ) with respect to the experimental data $Y_{obs}(\overline{x})$. In this case, due to the dimensionality $\overline{x} = [P_0, T_0, X_0, h_0, L/D]^T$ it is not possible to give an immediate graphical representation of the model response surface as a function of the whole $\overline{x}$ domain. Nevertheless, the difference in terms of L/D for the nozzle N2 is highlighted using different markers. It is worth mentioning that, in order to evaluate the model at physically consistent ICs & BCs, for subcooled ICs the couple $[P_0, \Delta T_{sub}]^T$ has been imposed, while for two-phase ICs $[P_0, X_0]^T$ have been used. Nevertheless, T_0 and h_0 are still employed as additional inputs just for obtaining additional information (also to be consistent with the experimental observations ICs & BCs domain). Whereas, regarding L/D, the different configurations have been obtained by varying the corresponding nodalization from case to case.

4.2.1. S-S: SA

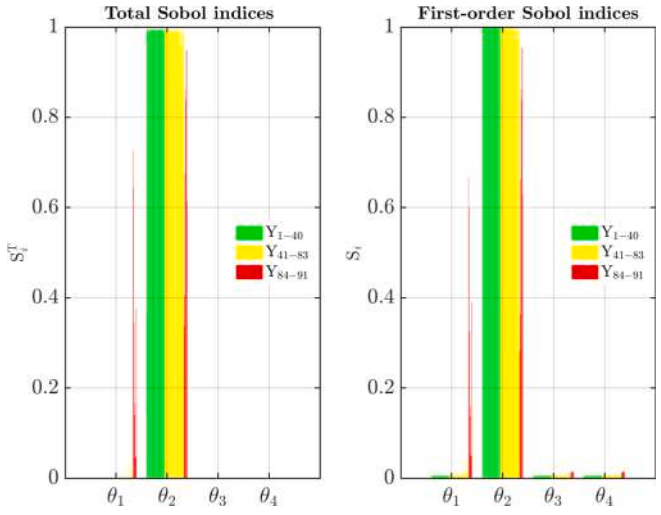
The results of the clustering analysis for S-S are presented in Fig. 13 and the related quantitative sensitivity measures can be found in Fig. 14. As shown by the aforementioned figures, in this case and unlike SMD, there is a clear and strong dependence of the model response by $\overline{x}$. Indeed, the sensitivity measures (see Fig. 14) highlight a clearly dominant region in terms of sensitivity, i.e., the high $(k = 3)$ response one. We could explain the presence of this peaked ICs & BCs dependent behaviour by resorting to physical considerations. In fact S-S is characterized by the presence of both multiphase and subcooled inlet conditions which are related to the inlet quality $x_3 = X_0$, defined in terms of saturated liquid and gas density as:

$$X_0 = \frac{(1/\rho) - (1/\rho_{l,sat})}{(1/\rho_{g,sat}) - (1/\rho_{l,sat})} \tag{17}$$

where $\rho_{l,sat}$, $\rho_{g,sat}$, ρ are saturated liquid, saturated vapour and the



(a) Borgonovo δ



(b) Sobol indexes

Fig. 15. Sensitivity of $\overline{\theta}$ for S-S.

two-phase mixture densities, respectively. Then, by observing the conditioned distributions of Fig. 13 **(c)** it is evident that the high response region of the output is driven by the subset of the input domain which corresponds to the subcooled initial conditions, since the support of $f_{x_3|y_m^{High}}$ is negative. Hence, we can claim that for S-S a strong dependency on the input $\overline{x}$ domain is observed.

In Fig. 15, the outcomes of the sensitivity analysis for $\overline{\theta}$ are illustrated. As already pointed out by the SMD results, θ_2 remains the clearly dominant parameter using both sensitivity measures. Nevertheless, in this case, there is also a localized peaked sensitivity for θ_1 just in the high response region. This behaviour could be easily linked to the previous physical consideration because θ_1 is a numerical coefficient directly related to the subcooled flow conditions (see Table 3). From this analysis we could expect a strong identifiability for θ_2 in every region and some relevant differences for θ_1 from region to region. Whereas θ_3 and θ_4 should remain of difficult identifiability due to their low sensitivity.

4.2.2. S-S: IUQ

In Fig. 16, the results of the traditional IUQ analysis are shown. In this case despite the overall outcomes of the procedure could be still considered satisfactory, nevertheless, there are some points in which the local behaviour of the calibrated model do not match the experimental data, specifically for some points in the medium and high response region (Fig. 16 **(c)** data index $>$ 55). In Table 6, a summary of the estimated $\mathscr{J}(\widehat{\overline{\theta}})$ can be found.

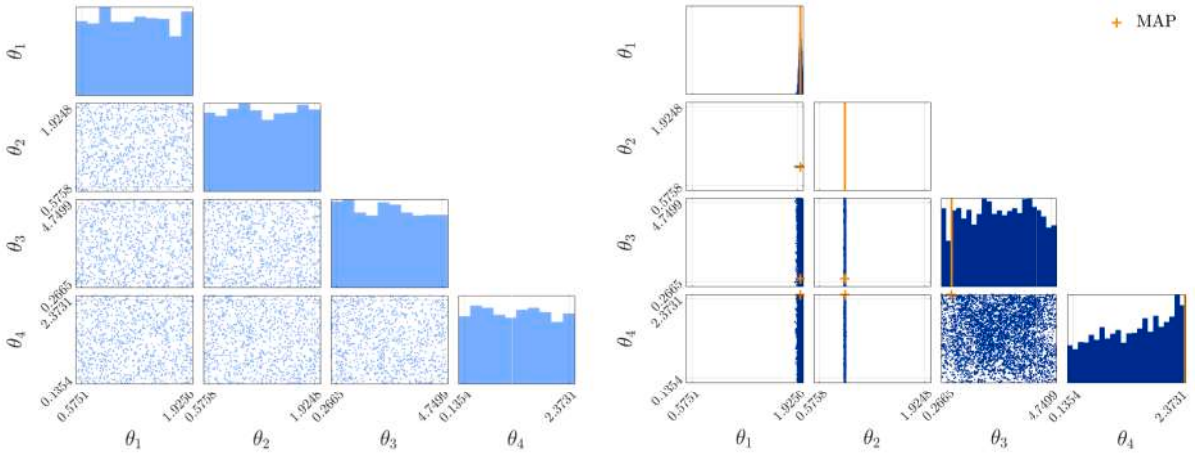
Then, in Fig. 17 and Table 7 the same information is provided for the results of the local IUQ analysis. It is worth noting that posteriors of θ_1 and θ_2 vary greatly across the three different regions. Despite this, the overall results of the reference verification procedure for both the low and the medium region (Fig. 17 **(c)**, **(f)**) are practically indiscernible from the ones obtained through the traditional IUQ (Fig. 16 **(c)**). Nevertheless, for the high region (Fig. 17 **(i)**), the accordance between the experimental data and the posterior output distribution is clearly improved with respect to the traditional approach (Fig. 16 **(c)**). We could link this fact to the peaked sensitivity previously identified through the SA (see Fig. 14).

4.2.3. S-S: Validation

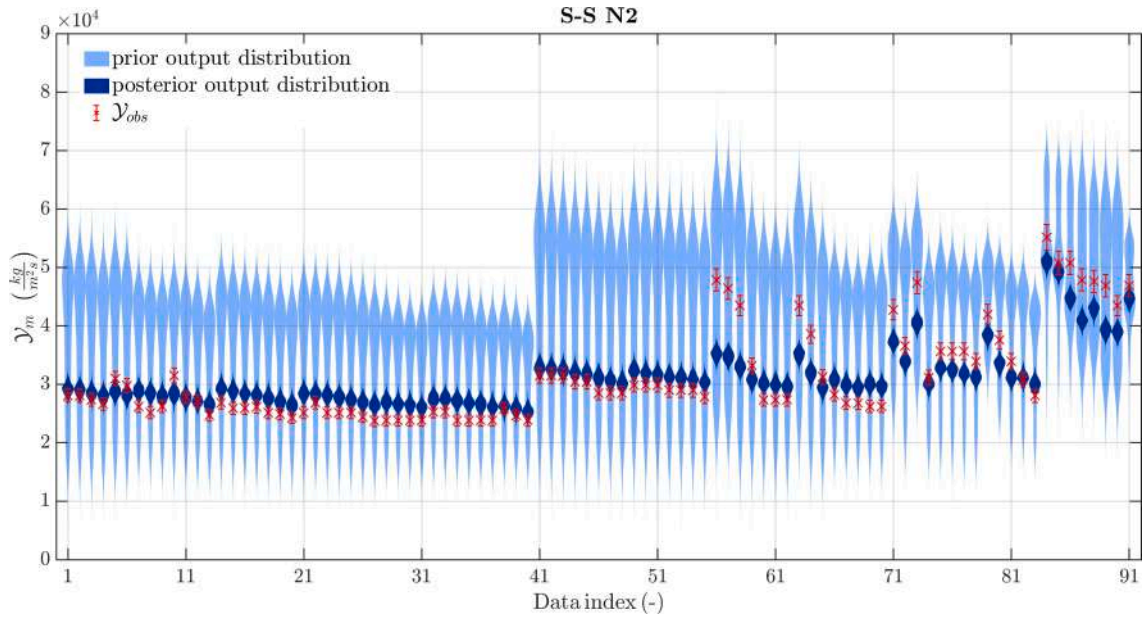
In Fig. 18, the results of the validation procedure are presented. In this case the overall trend of the model predictions is shifted downwards with respect to the experimental observations for both approaches. This behaviour can be considered as the side effect of some distortion between the configurations of nozzle N2 used for the IUQ (i.e., Y_{obs}) and those employed for the validation (i.e., Y_{val}). To investigate the causes of this effect, in Fig. 19, we have plotted in parallel coordinates the input values $\overline{x}$ of the quantification dataset Y_{obs} (blue continuous lines) and those of the validation database Y_{val} (red dashed lines). The input values of Y_{obs} cover homogeneously those of Y_{val}, except for L/D, where the largest and the lowest values are only represented in the Y_{val} dataset, whereas they are not in Y_{obs} (as shown by the red shadowed circles in Fig. 19) leading to a systematic bias in the estimation of y_m. Nevertheless, these results are quite illustrative of the improvement obtained through the application of the local IUQ procedure. Indeed, by comparing Fig. 18 (**(a)** and **(b)**) it is visible how minor differences are present for the low and medium regions (i.e., the low sensitivity ones) with respect to the high one, which is also the most sensitive one. In fact, for the high region the predictive accuracy of the model is greatly enhanced, so this kind of analysis shows that through input space segmentation it is possible to potentially increase the performances of Bayesian IUQ methodologies when a high dependency of the model from ICs & BCs is detected. In this case the evaluation of the mean squared error between the predictions of the two approaches underlines a difference of one order of magnitude for the high region ($\sim$ 24 %) with respect to the other two ($\sim$ 2 %).

4.3. Comparison between SMD and S-S

In Fig. 20, a comparison between SMD **(a)** and S-S **(b)** is presented in terms of validation results. The metric of Eq. (16) is applied to the subset of the validation domain corresponding to each response region. The results show that the local IUQ procedure has been capable to improve the model predictive performances for the High response region of S-S, which was also the most sensitive one with respect to input $\overline{x}$, whereas no appreciable differences can be observed for SMD. It is worth noting



(a) Prior $\mathcal{P}(\overline{\theta})$ **(b)** Posterior $\mathcal{J}(\hat{\theta})$

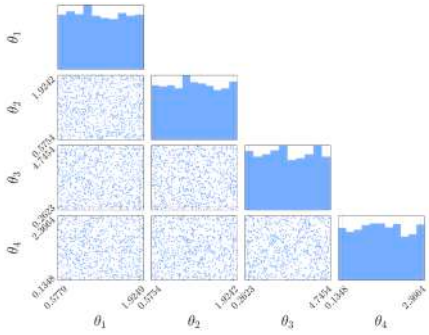


(a) IUQ Reference Verification

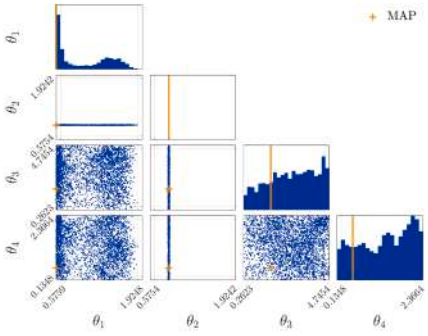
Fig. 16. Results of traditional IUQ analysis for S-S.

Table 6
Posterior PDFs statistics estimated through the traditional IUQ for S-S.

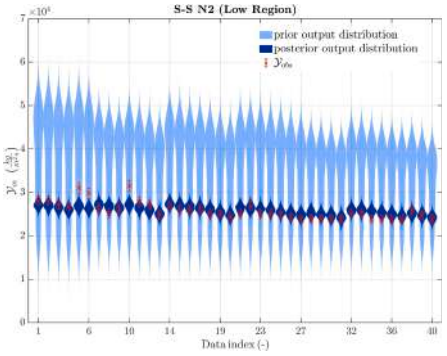
Parameter	Mean	Variance	MAP	[*Q(0.05)-Q(0.95)*]
$\widehat{\theta}_1$	1.9568	$3.2078\ 10^{-04}$	1.9651	[1.9275–1.9959]
$\widehat{\theta}_2$	0.8996	$1.4135\ 10^{-05}$	0.9001	[0.8924–0.9071]
$\widehat{\theta}_3$	2.5178	1.9580	0.4644	[0.1241–4.8282]
$\widehat{\theta}_4$	1.4071	0.5118	2.4932	[0.0877–2.4586]



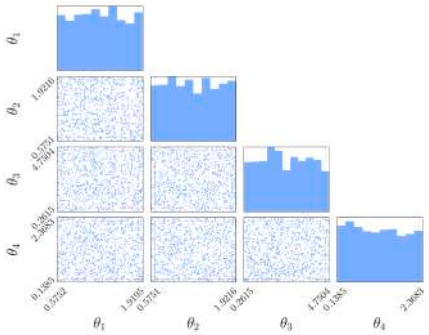
(a) Prior $\mathcal{P}(\overline{\theta})$



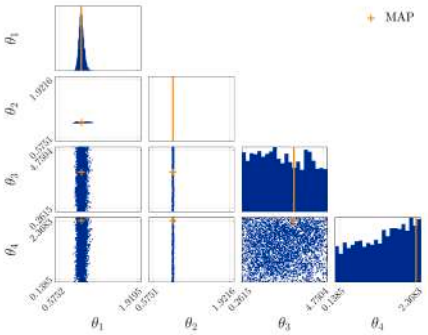
(b) Posterior $\mathcal{J}\left(\hat{\overline{\theta}}^{Low}\right)$



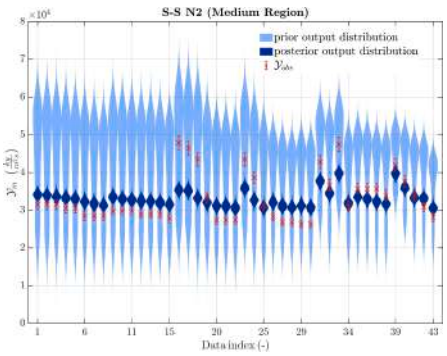
(c) IUQ Reference (Low Region)



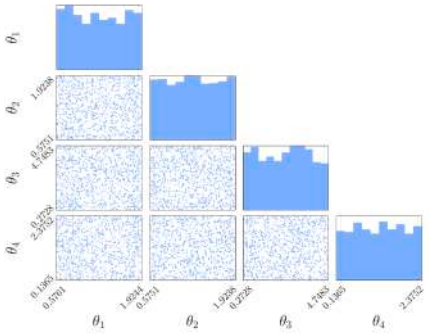
(a) Prior $\mathcal{P}(\overline{\theta})$



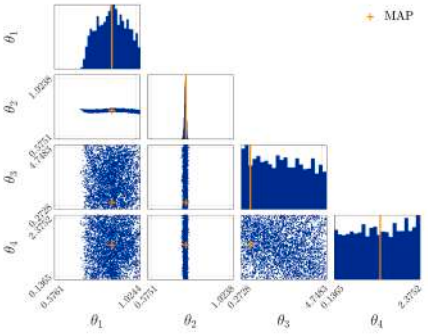
(e) Posterior $\mathcal{J}\left(\hat{\overline{\theta}}^{Medium}\right)$



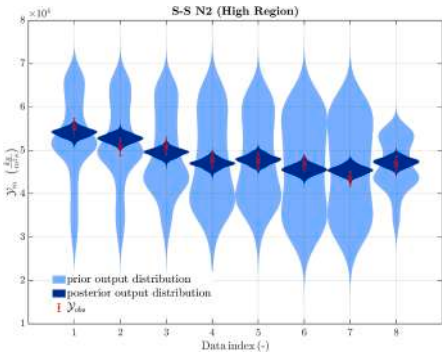
(f) IUQ Reference (Medium Region)



(g) Prior $\mathcal{P}(\overline{\theta})$



(h) Posterior $\mathcal{J}\left(\hat{\overline{\theta}}^{High}\right)$

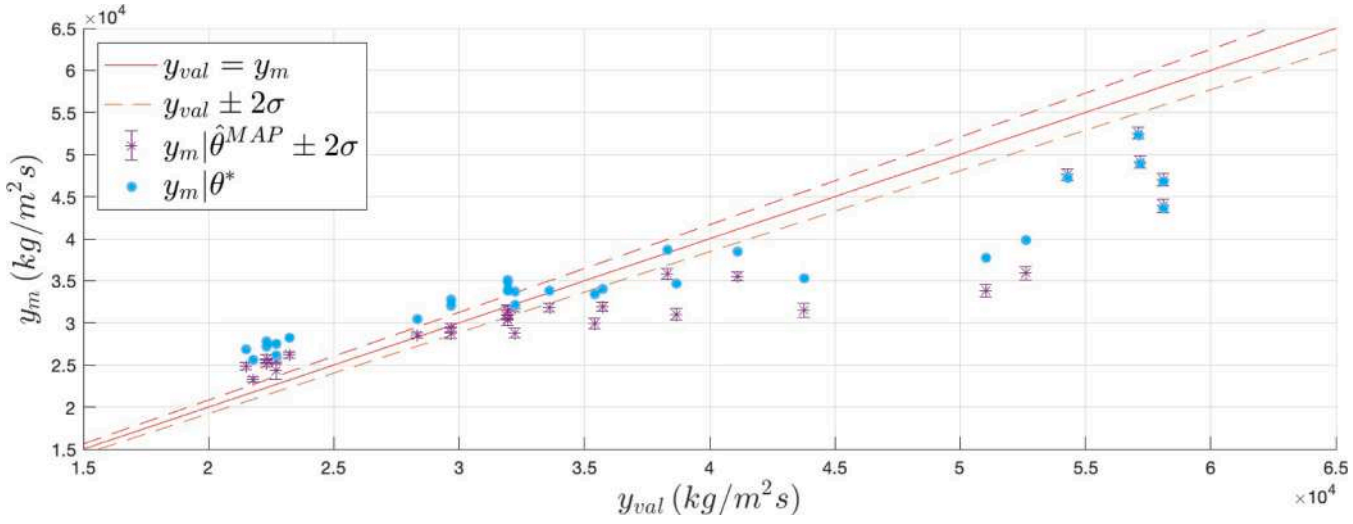


(i) IUQ Reference (High Region)

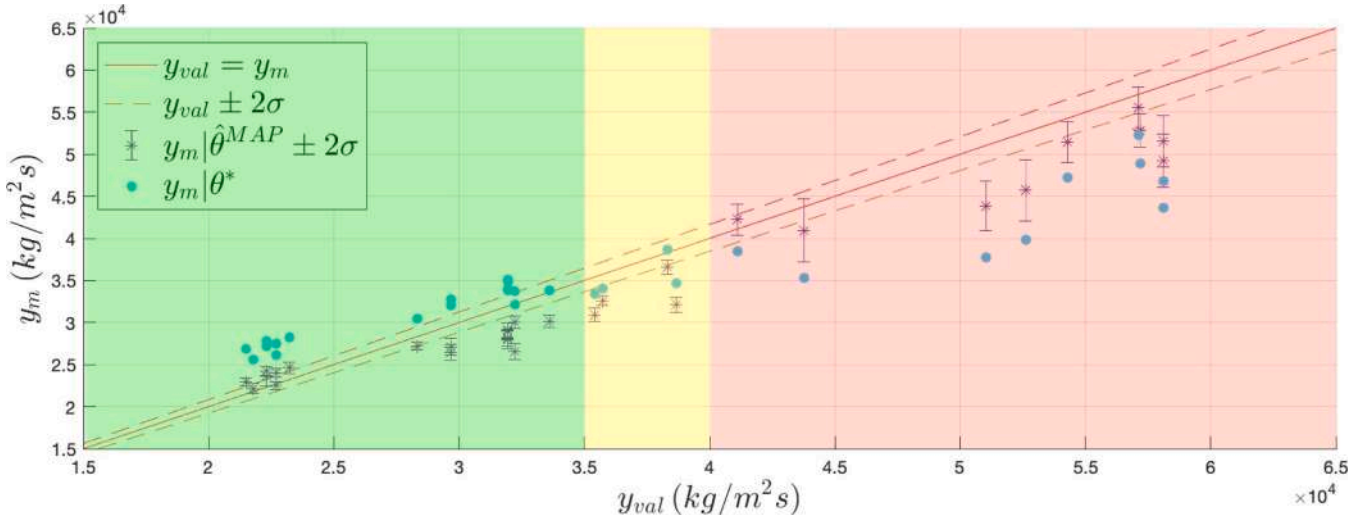
Fig. 17. Results of local IUQ analysis for S-S.

Table 7
Posterior PDFs statistics estimated through the local IUQ for S-S.

Parameter	Mean	Variance	MAP	[Q(0.05)-Q(0.95)]
$\widehat{\theta}_1^{Low}$	0.9926	0.1989	0.5009	[0.5036–1.7418]
$\widehat{\theta}_2^{Low}$	0.8318	$2.8422\ 10^{-05}$	0.8322	[0.8211–0.8422]
$\widehat{\theta}_3^{Low}$	2.8304	1.8819	1.6082	[0.2588–4.9125]
$\widehat{\theta}_4^{Low}$	1.3907	0.5411	0.4705	[0.0899–2.4427]
$\widehat{\theta}_1^{Medium}$	1.1530	0.2231	0.9554	[0.8783–1.0553]
$\widehat{\theta}_2^{Medium}$	1.2627	$6.2150\ 10^{-04}$	0.9276	[0.9180–0.9372]
$\widehat{\theta}_3^{Medium}$	2.5776	2.1252	3.0548	[0.1171–4.8520]
$\widehat{\theta}_4^{Medium}$	1.1441	0.5163	2.3577	[0.0871–2.4552]
$\widehat{\theta}_1^{High}$	1.5115	0.0633	1.5051	[1.0654–1.9510]
$\widehat{\theta}_2^{High}$	1.1577	$3.4527\ 10^{-04}$	1.1698	[1.1186–1.1919]
$\widehat{\theta}_3^{High}$	2.3707	2.0868	0.5602	[0.1307–4.8588]
$\widehat{\theta}_4^{High}$	1.2962	0.5293	1.3443	[0.0762–2.4538]



(a) IUQ



(b) Local IUQ

Fig. 18. Comparison of validation results of the traditional and the local IUQ procedure for S-S.

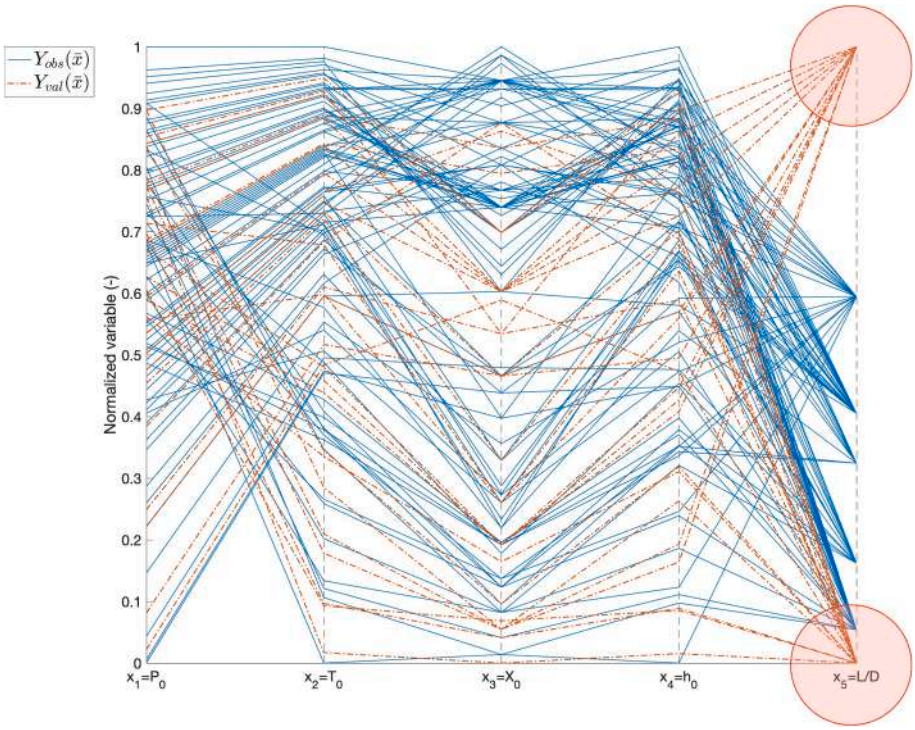
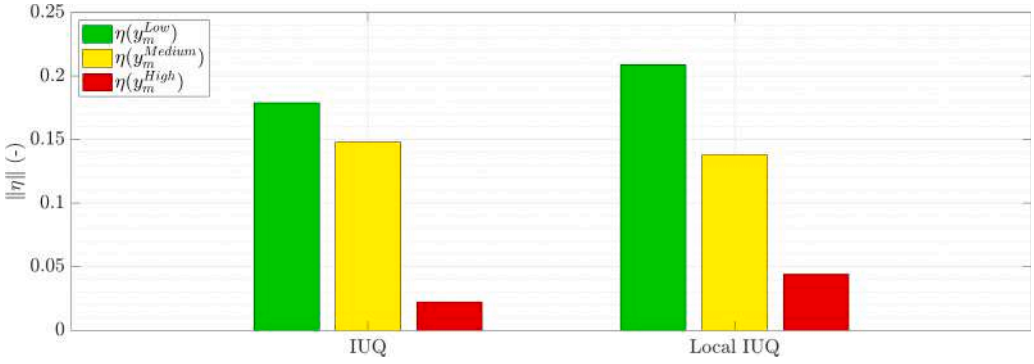
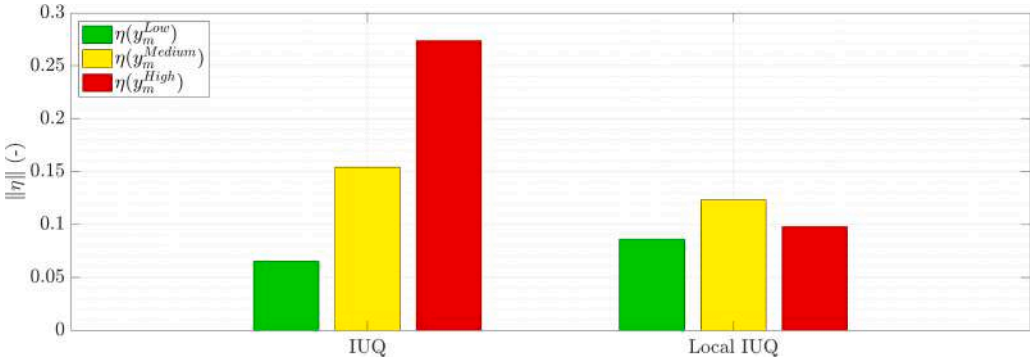


Fig. 19. Parallel coordinates representation of the input values $\bar{x}$ for the quantification database Y_{obs} (blue continuous lines) and the validation database Y_{val} (red dashed lines) for S-S.



(a) SMD



(b) S-S

Fig. 20. Validation results comparison between SMD and S-S.

that the order of magnitude of the calculated discrepancies between model predictions and experimental data is in the range (10 −25%), that is typically deemed acceptable for two-phase T-H models validation according to (Jaeger et al., 2013).

5. Conclusions

In this paper, the development of an IUQ procedure based on input space segmentation through SA is presented and tested with respect to two SETs BE models, namely SMD and S-S. The rationale behind this approach is highlighting and extending the role of SA during IUQ for characterizing the model local behaviour and guiding a tailored IUQ analysis on the data correlated with the subset of the input domain which mostly affects the QoI response in the region of interest. The approach is based on an input space segmentation performed through distribution based global SA measures like the KL divergence and the H distance and, then, on a localized IUQ. This novel methodology is integrated with the traditional IUQ procedure in a comprehensive framework of analysis inspired by the SAPIUM project guidelines (Baccou et al., 2020). The application of the developed approach to a case study from the ATRIUM project shows that this kind of analysis could improve the performance of the IUQ methodologies when the response of the model is greatly affected by the variation of the input domain (e.g., ICs & BCs). Further work is still needed to refine the methodology and to test it on more complex applications like IETs to verify the robustness of the approach. In principle, this framework should express all its potential in the case of multiphysics applications, which are typically characterized by a highly non linear behaviour and an extensive dimensionality. Therefore, in these cases, there is not a trivial procedure to partition the input space exact relying on an expert-based assessment. Future developments will be focused on the elaboration of a tailored "local" discrepancy term capable to properly characterize the impact of model uncertainty on the analysis results (Arendt et al., 2012; Wu et al., 2021): due to the lack of a consistent mathematical formulation and of a consensus on the functional form for the discrepancy term (which is supposedly strongly dependent on the particular case study under analysis), the problem could be robustly addressed by the generation of multiple (diverse, possibly empirical) models and the selection (and/or merging) of the best ones through validation and evaluation of the corresponding predictive capability (Ling et al., 2014). Also, the scalability and extrapolation issue of the estimated posterior distributions $\mathscr{J}(\widehat{\theta})$ of SETs for the applicability of these results on the IET have to be addressed: in this view, the possibility to merge the different posterior density functions by means of Bayesian Model Averaging (Fragoso et al., 2018) or Frequentist Model Averaging (Fu and Pan, 2018) could be explored, where the models predictive capability and the datasets adequacy (Di Maio et al., 2024) may be used as model weighting factors.

CRediT authorship contribution statement

Francesco Di Maio: Writing – review & editing, Supervision, Methodology, Conceptualization. **Thomas Matteo Coscia:** Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Nicola Pedroni:** Writing – review & editing, Conceptualization. **Andrea Bersano:** Writing – review & editing, Software, Data curation. **Fulvio Mascari:** Writing – review & editing, Software, Data curation. **Enrico Zio:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

The authors would like to acknowledge the contribution of all those individuals who had a key role and leadership in the conduct of the activity of the Application Tests for Realization of Inverse Uncertainty quantification and validation Methodologies in thermal-hydraulics (ATRIUM) of the NEA Working Group on the Analysis and Management of Accidents (WGAMA), especially the task leaders L. Sargentini and A. Ghione (CEA, France).

References

Agnello, G., et al., 2022. Cold Leg LBLOCA uncertainty analysis using TRACE/DAKOTA coupling. J. Phys.: Conf. Ser.. IOP Publishing Ltd. https://doi.org/10.1088/1742-6596/2177/1/012023.

Alibrandi, U., Andersen, L.V., Zio, E., 2022. Informational probabilistic sensitivity analysis and active learning surrogate modelling. Probab. Eng. Mech. 70, 103359 https://doi.org/10.1016/j.probengmech.2022.103359.

Applied Programming Technology Inc. (2021) 'Symbolic Nuclear Analysis Package (SNAP) Users Manual'. Bloomsburg, PA.

Arendt, P.D., Apley, D.W., Chen, W., 2012. Quantification of model uncertainty: Calibration, model discrepancy, and identifiability. J. Mech. Des. 134 (10) https://doi.org/10.1115/1.4007390.

Baccou, J., et al., 2019. Development of good practice guidance for quantification of thermal-hydraulic code model input uncertainty. Nucl. Eng. Des. 384 https://doi.org/10.1016/j.nucengdes.2019.110173.

Baccou, J., et al., 2024. A systematic approach for the adequacy analysis of a set of experimental databases: Application in the framework of the ATRIUM activity. Nucl. Eng. Des. 421 https://doi.org/10.1016/j.nucengdes.2024.113035.

Baccou, J. et al. (2020) SAPIUM: A Generic Framework for a Practical and Transparent Quantification of Thermal-Hydraulic Code Model Input Uncertainty. Nucl. Sci. Eng., 194(8–9), pp. 721–736. https://doi.org/10.1080/00295639.2020.1759310.

Bersano, A. et al. (2020). Ingress of Coolant Event simulation with TRACE code with accuracy evaluation and coupled DAKOTA Uncertainty Analysis. Fusion Eng. Des., 159, p. 111944. Available at: https://doi.org/10.1016/J.FUSENGDES.2020.111944.

Borgonovo, E., 2007. A new uncertainty importance measure. Reliab. Eng. Syst. Saf. 92 (6), 771–784. https://doi.org/10.1016/j.ress.2006.04.015.

Borgonovo, E., Plischke, E., 2016. Sensitivity analysis: A review of recent advances. Eur. J. Oper. Res. 248 (3), 869–887. https://doi.org/10.1016/j.ejor.2015.06.032.

Brooks, S., et al., 2011. Handbook of Markov Chain Monte Carlo. Handbooks of Modern Statistical Methods, Chapman & Hall/CRC.

D'Auria, F., Camargo, C., Mazzantini, O., 2012. The Best Estimate Plus Uncertainty (BEPU) approach in licensing of current nuclear reactors. Nucl. Eng. Des. 248, 317–328. https://doi.org/10.1016/j.nucengdes.2012.04.002.

D'Auria, F., Galassi, G., Mazzantini, O., 2022. The best estimate plus uncertainty approach in licensing of Atucha II. Pressurized Heavy Water Reactors 8, 51–204.

De Crécy, A., 2001. Determination of the uncertainties of the constitutive relationships of the CATHARE 2 code. M & C conference, Salt Lake City, Utah, USA.

Di Maio, F., et al., 2015. Finite mixture models for sensitivity analysis of thermal hydraulic codes for passive safety systems analysis. Nucl. Eng. Des. 289, 144–154. https://doi.org/10.1016/j.nucengdes.2015.04.035.

Di Maio, F., Coscia, T.M., Zio, E., 2024. Data adequacy by an extended analytic hierarchy process for inverse uncertainty quantification in nuclear safety analysis. Nucl. Eng. Des. 419 https://doi.org/10.1016/j.nucengdes.2024.112971.

Domitr, P., Włostowski, M., 2021. The use of machine learning for inverse uncertainty quantification in TRACE code based on Marviken experiment. Nucl. Eng. Des. 384, 111498 https://doi.org/10.1016/j.nucengdes.2021.111498.

Fragoso, T.M., Bertoli, W., Louzada, F. (2018). Bayesian model averaging: a systematic review and conceptual classification. Int. Statistical Rev., 86(1), pp. 1–28. Available at: https://doi.org/10.1111/insr.12243.

Fu, P., Pan, J. (2018). A review on high-dimensional frequentist model averaging. Open J. Statistics, 08(03), pp. 513–518. Available at: https://doi.org/10.4236/ojs.2018.83033.

Ghione, A., 2023. Application of SAPIUM guidelines to Input Uncertainty Quantification: the ATRIUM project. In 20th International Topical Meeting on Nuclear Reactor Thermal Hydraulics (NURETH20).

Gibbs, A.L., Su, F.E. (2002). On choosing and bounding probability metrics. Int. Statist. Rev., 70(3), pp. 419–435. Available at: https://doi.org/10.1111/j.1751-5823.2002.tb00178.x.

Goodman, J., Weare, J., 2010. Ensemble samplers with affine invariance. Commun. Appl. Mathematics Computat. Sci. 5 (1).

Higdon, D., et al., 2008. A Bayesian calibration approach to the thermal problem. Comput. Methods Appl. Mech. Eng. 197 (29), 2431–2441. https://doi.org/10.1016/j.cma.2007.05.031.

Hintze, J.L., Nelson, R.D. (1998). Violin Plots: A Box Plot-Density Trace Synergism. Am. Statistician, 52(2), pp. 181–184. Available at: https://doi.org/10.1080/00031305.1998.10480559.

Hou, J., Avramova, M., Ivanov, K. (2020) Best-Estimate Plus Uncertainty Framework for Multiscale, Multiphysics light water reactor core analysis. Sci. Technol. Nucl. Installat., 2020, pp. 1–18. Available at: https://doi.org/10.1155/2020/7526864.
IAEA, 2008. Best Estimate Safety Analysis for Nuclear Power Plants. Uncertainty Evaluation.
IAEA (2014) Progress in Methodologies for the Assessment of Passive Safety System Reliability in Advanced Reactors.
Jaeger, W. et al. (2013). Validation and comparison of two-phase flow modeling capabilities of CFD, sub channel and system codes by means of post-test calculations of BFBT transient tests. Nucl. Eng. Des., 263, pp. 313–326. Available at: https://doi.org/10.1016/J.NUCENGDES.2013.06.002.
Kennedy, M.C., O'Hagan, A. (2001) Bayesian calibration of computer models. J. Royal Statistical Soc.. Ser. B: Statist. Methodol., 63(3), pp. 425–464. Available at: https://doi.org/10.1111/1467-9868.00294.
Kovtonyuk, A., et al., 2017. Development and assessment of a method for evaluating uncertainty of input parameters. Nucl. Eng. Des. 321, 219–229. https://doi.org/10.1016/j.nucengdes.2016.08.021.
Lamas, C. et al. (2022). UQLab user manual: Sensitivity analysis. Report UQLab-V2.0-106.
Ling, Y., Mullins, J., Mahadevan, S., 2014. Selection of model discrepancy priors in Bayesian calibration. J. Comput. Phys. 276, 665–680. https://doi.org/10.1016/j.jcp.2014.08.005.
Liu, F., et al., 2019. Global sensitivity analysis for multivariate outputs based on multiple response Gaussian process model. Reliab. Eng. Syst. Saf. 189, 287–298. https://doi.org/10.1016/j.ress.2019.04.039.
Marquès, M., et al., 2005. Methodology for the reliability evaluation of a passive system and its integration into a Probabilistic Safety Assessment. Nucl. Eng. Des. 235 (24), 2612–2631. https://doi.org/10.1016/j.nucengdes.2005.06.008.
NEA (2011) BEMUSE Phase VI Report: Status report on the area, classification of the methods, conclusions and recommendations, Nuclear Safety, NEA/CSNI/R(2011).
NEA (2016) PREMIUM, a benchmark on the quantification of the uncertainty of the physical models in the system thermal-hydraulic codes: methodologies and data review, Nuclear Safety, NEA/CSNI/R(2016)9.
NEA (2017) Post-BEMUSE Reflood Model Input Uncertainty Methods (PREMIUM) Benchmark, Final Report, Nuclear Safety, NEA/CSNI/R(2016)18.
Nutt, W.T., Wallis, G.B., 2004. Evaluation of nuclear safety from the outputs of computer codes in the presence of uncertainties. Reliab. Eng. Syst. Saf. 83 (1), 57–77. https://doi.org/10.1016/j.ress.2003.08.008.
Oberkampf, W.L., Trucano, T.G. (2007) Verification and Validation Benchmarks.
Oberkampf, W., Roy, C., 2010. Verification and Validation in Scientific Computing. Cambridge University Press.
Park, C.K., Ahn, K.-I., 1994. A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment. Reliab. Eng. Syst. Saf. 46 (3), 253–261. https://doi.org/10.1016/0951-8320(94)90119-8.
Perret, G., et al., 2019. Global Sensitivity and Registration Strategy for Temperature Profile of Reflood Experiment Simulations. Nucl. Technol. 205, 1638–1651. https://api.semanticscholar.org/CorpusID:164742293.
Perret, G., et al., 2022. Global sensitivity analysis and bayesian calibration on a series of reflood experiments with varying boundary conditions. Nucl. Technol. 208 (4), 711–722. https://doi.org/10.1080/00295450.2021.1936879.
Petruzzi, A., 2019. The CASUALIDAD Method for uncertainty evaluation of best-estimate system thermal-hydraulic calculations. Nucl. Technol. 205 (12), 1554–1566. https://doi.org/10.1080/00295450.2019.1632092.
Porter, N.W., Mousseau, V.A. (2019). Bayesian calibration of empirical models common in melcor and other nuclear safety codes. In: 18th International Topical Meeting on Nuclear Reactor Thermal Hydraulics (NURETH-19).
Porter, N.W., Mousseau, V.A., Avramova, M.N., 2019. Quantified validation with uncertainty analysis for turbulent single-phase friction models. Nucl. Technol. 205 (12), 1607–1617. https://doi.org/10.1080/00295450.2018.1548221.
Puppo, L., et al., 2021. A Framework based on Finite Mixture Models and Adaptive Kriging for Characterizing Non-Smooth and Multimodal Failure Regions in a Nuclear Passive Safety System. Reliab. Eng. Syst. Saf. 216, 107963 https://doi.org/10.1016/j.ress.2021.107963.
Roma, G., et al., 2021. A Bayesian framework of inverse uncertainty quantification with principal component analysis and Kriging for the reliability analysis of passive safety systems. Nucl. Eng. Des. 379, 111230 https://doi.org/10.1016/j.nucengdes.2021.111230.
Roma, G., et al., 2022. Passive safety systems analysis: A novel approach for inverse uncertainty quantification based on Stacked Sparse Autoencoders and Kriging metamodeling. Prog. Nucl. Energy 148, 104209. https://doi.org/10.1016/j.pnucene.2022.104209.
Rousseau, J.C., 1987. Flashing flow. In: Multiphase Science and Technology. Hemisphere Publishing Coorporation, USA, pp. 378–389.
Saltelli, A., et al., 2004. Sensitivity Analysis in Practice. Wiley.
Sánchez, V.H., Thieme, M., Tietsch, W., 2012. Validation and application of the thermal hydraulic system Code TRACE for analysis of BWR transients. Sci. Technol. Nucl. Installat. https://doi.org/10.1155/2012/247482.
Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability 26. Chapman & Hall/CRC.
Sozzi, G.L., Sutherland, W.A. (1975) Critical flow of saturated and subcooled water at high pressure, Report NEDO-13418. San Jose, USA.
Stuart, A.M., 2010. Inverse problems: A Bayesian perspective. Acta Numer. 19, 451–559. https://api.semanticscholar.org/CorpusID:122815877.
Trucano, T.G., et al., 2006. Calibration, validation, and sensitivity analysis: What's what. Reliab. Eng. Syst. Saf. 91 (10), 1331–1357. https://doi.org/10.1016/j.ress.2005.11.031.
U.S. Nuclear Regulatory Commission (2020) 'TRACE v5.0 patch 6 Theory Manual'.
Unal, C., et al., 2011. Improved best estimate plus uncertainty methodology, including advanced validation concepts, to license evolving nuclear reactors. Nucl. Eng. Des. 241 (5), 1813–1833. https://doi.org/10.1016/j.nucengdes.2011.01.048.
Vinai, P., Macian-Juan, R., Chawla, R., 2007. A statistical methodology for quantification of uncertainty in best estimate code physical models. Ann. Nucl. Energy 34 (8), 628–640. https://doi.org/10.1016/j.anucene.2007.03.003.
Wagner, P.R. et al. (2019). UQLab user manual: Bayesian inference for model calibration and validation.
Wilks, S.S., 1941. Determination of Sample Sizes for Setting Tolerance Limits. Ann. Math. Stat. 12 (1), 91–96.
Wilks, S.S., 1942. Statistical Prediction with special reference to the problem of tolerance limits. Ann. Math. Stat. 13 (4), 400–409.
Wu, X., et al., 2021. A comprehensive survey of inverse uncertainty quantification of physical model parameters in nuclear system thermal-hydraulics codes. Nucl. Eng. Des. 384 https://doi.org/10.1016/j.nucengdes.2021.111460.
Wu, X., Kozlowski, T., 2017. Inverse uncertainty quantification of reactor simulations under the Bayesian framework using surrogate models constructed by polynomial chaos expansion. Nucl. Eng. Des. 313, 29–52. https://doi.org/10.1016/j.nucengdes.2016.11.032.
Wu, X., Kozlowski, T., Meidani, H., 2018. 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., Shirvan, K., Kozlowski, T., 2019. Demonstration of the relationship between sensitivity and identifiability for inverse uncertainty quantification. J. Comput. Phys. 396, 12–30. https://doi.org/10.1016/j.jcp.2019.06.032.
Xiao, S., Lu, Z., Qin, F., 2016. Estimation of the Generalized Sobol's sensitivity index for multivariate output model using unscented transformation. J. Struct. Eng. [Preprint]. https://doi.org/10.1061/(ASCE)ST.1943.
Xiong, Q., Gou, J., Wen, Y., et al., 2020a. Global sensitivity analysis of LOFT large break loss of coolant accident with optimized moment-independent method. Ann. Nucl. Energy 139, 107289. https://doi.org/10.1016/j.anucene.2019.107289.
Xiong, Q., Gou, J., Mao, H., et al., 2020b. Optimization of sensitivity analysis in best estimate plus uncertainty and the application to large break LOCA of a three-loop pressurized water reactor. Prog. Nucl. Energy 126, 103396. https://doi.org/10.1016/j.pnucene.2020.103396.
Zio, E., Di Maio, F., Tong, J., 2010. Safety margins confidence estimation for a passive residual heat removal system. Reliab. Eng. Syst. Saf. 95 (8), 828–836. https://doi.org/10.1016/j.ress.2010.03.006.