

UNCERTAINTY QUANTIFICATION OF AN INDUSTRIAL SCALE CFD APPLICATION USING STOCHASTIC SPECTRAL METHODS

Philipp J. Wenig¹, Stephan Kelm², and Markus Klein¹

¹University of the Bundeswehr Munich
Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany
e-mail: {philipp.wenig,markus.klein}@unibw.de

² Institute of Energy and Climate Research, Forschungszentrum Jülich GmbH
52425 Jülich, Germany
e-mail: s.kelm@fz-juelich.de

Abstract. *Accurate prediction of flow phenomena is crucial for a wide range of industrial and scientific applications. However, due to the complexity of many fluid systems, even state-of-the-art Computational Fluid Dynamics (CFD) simulations involve various sources of uncertainty, which need to be quantified. Therefore, this work focuses on the Uncertainty Quantification (UQ) of a CFD simulation of a technical scale experiment, which investigates buoyancy-driven mixing processes between two miscible fluids. The CFD model is subject to uncertainties in initial and boundary conditions, as well as in thermo-physical properties. Since computational runs of the CFD model are computationally expensive, the uncertainty analysis requires the use of efficient methods. For this reason, stochastic spectral methods, such as Polynomial Chaos Expansion (PCE) and Karhunen-Loève Expansion (KLE), are applied along with novel approaches like Stochastic Model Composition (SMC) for the uncertainty quantification of responses. Fluctuations due to turbulence are identified as considerable source of uncertainty and are modeled with tailored stochastic models. Results include a comprehensive probabilistic representation of response quantities with probability density functions (PDFs), statistical moments, variance-based decomposition and supplementary error estimates of the underlying stochastic models. This enables well-founded result interpretation and informed decision making. The UQ results show that the uncertain input parameters considerably affect the duration of the mixing process. Through the investigations, the applied UQ methods were tested on an engineering application and demonstrate great potential as a viable technique for uncertainty quantification in technical-scale computations.*

Keywords: Uncertainty Quantification, Computational Fluid Dynamics, Polynomial Chaos Expansion, Karhunen-Loève Expansion, Error estimation, Variance-based decomposition.

1 INTRODUCTION

For the prediction of physical phenomena or the reliable design of components in engineering applications by means of numerical simulations, the consideration of uncertainties is essential. They result from numerical approximations, uncertain initial and boundary conditions, thermodynamic properties etc. However, UQ of industrial scale applications in the field of CFD is computationally expensive. Therefore, the maximum number of computational runs needs to be limited and efficient methods for the reliable prediction of stochastic results have to be applied. Within the scope of this work, UQ of an application-oriented CFD validation case is conducted, which investigates buoyancy-driven mixing processes between two miscible fluids within a vessel of volume $V \approx 60 \text{ m}^3$. Various uncertainties in initial and boundary conditions, as well as in the thermo-physical properties affect the CFD model results. The impact of these uncertainties on responses needs to be quantified. Therefore, by using a generic test case, different methods were initially developed and qualified as suitable for the application to engineering applications [1, 2, 3, 4]. Stochastic spectral methods, such as PCE and KLE, were proven a promising approach and form the basis for the approximation of stochastic results in the present work. The novel SMC approach is applied, which proposes stochastic model construction through function composition of individual stochastic models. Moreover, response transients, which exhibit fluctuations due to turbulence, are considered in particular with combined stochastic models consisting of representations for the trend and fluctuating component of the response. On the basis of stochastic models, variance-based measures are derived and enable the assignment of response uncertainties to input uncertainties. Supplementary model error estimates ensure the reliability of presented results. Based on the evaluations, the applied UQ methods demonstrate great potential as a viable technique for uncertainty quantification in technical-scale computations, and are appropriate for the risk and safety assessment of engineering applications.

The paper is structured as follows. In Section 2, the employed techniques for the prediction and stochastic representation of the underlying flow phenomena are detailed. The obtained results are presented and discussed in Section 3. A summary and conclusion of the work is provided in Section 4.

2 METHODS

Various methods are used within the framework for the Uncertainty Quantification of the application case. To ensure a thorough understanding of the underlying methodology, a flow chart detailing the sequential steps of the approach is provided in Figure 1. The initial step involves generating a computational model that adequately represents the application case using numerical techniques. If the model is subject to uncertainties, these uncertainties are identified and described probabilistically. Subsequently, suitable UQ methods for stochastic modeling of response variables are selected. According to the chosen UQ method and the input uncertainties, corresponding input parameter data is generated for individual simulation runs of the computational model and simulation setups are created. Afterwards, the simulation runs are conducted and data of response quantities is collected. The response data obtained from the simulations is utilized to construct stochastic models. The evaluation of model response data and simulation response data enables the estimation of the model error. Based on the model error, it can be assessed whether the model performs well when the error estimate is small or improvements of the model are required when the error estimate is large. In the latter case, the process has to be repeated by increasing the number of computational runs or by using more appropriate methods. On the other hand, if the model accuracy is high, model evaluations can be carried

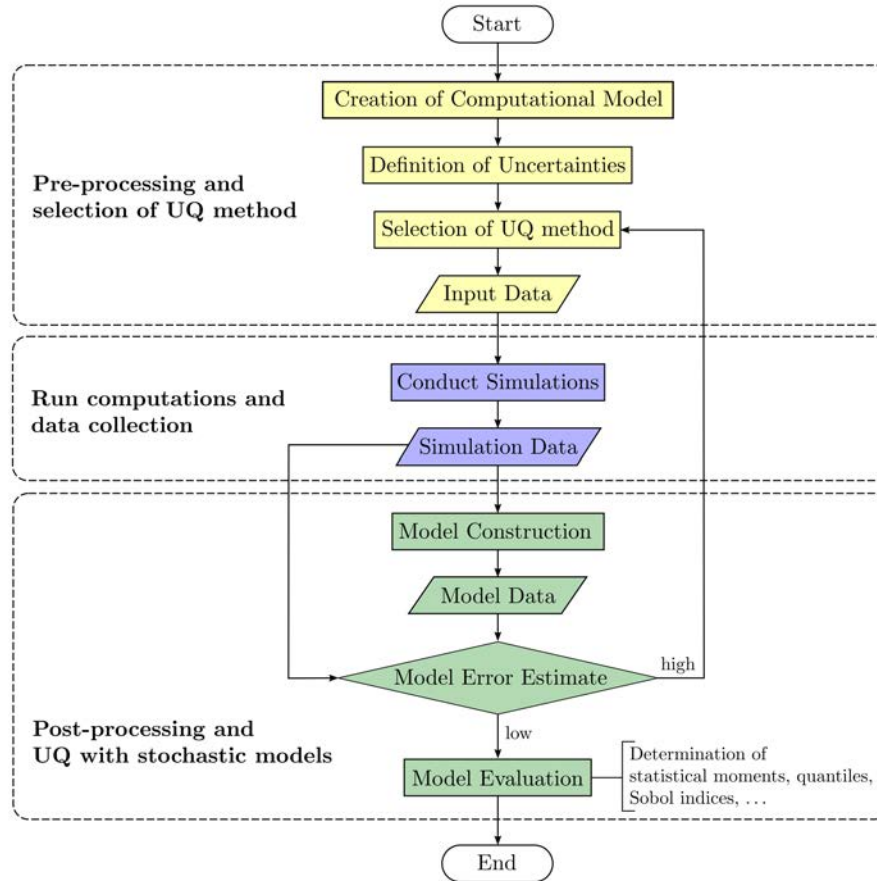


Figure 1: Flow chart illustrating the step-by-step process for conducting the underlying uncertainty analysis.

out to determine numerous statistics for UQ.

Overall, the process can thus be divided into three basic steps, which are highlighted in three different colors in Figure 1. The first step is to establish the foundation for the uncertainty study by using an appropriate computational setup and suitable methods. Subsequently, computational runs are conducted and associated response data is collected. In the last step, post-processing of the collected data eventually yields stochastic models using modeling techniques. These stochastic models enable the derivation of detailed statistical information for the comprehensive probabilistic description of uncertain responses.

The individual methods used in the aforementioned steps are explained in more detail in the next sections. Section 2.1 presents applied mathematical models and numerical techniques for the creation of the CFD model along with the definition of uncertainties. Section 2.2 covers the applied UQ methods for model construction and model evaluation.

2.1 CFD methods

Section 2.1.1 explains the computational setup of the application case and the definition of uncertainties in the CFD model. The numerical framework and discretization techniques are detailed in Section 2.1.2.

2.1.1 Application Setup and Uncertainties

Investigation of flow and transport phenomena within a unique, technical-scale experimental facility is conducted utilizing CFD methodologies. The experimental facility refers to the THAI (Thermal-hydraulics, Hydrogen, Aerosols and Iodine) test containment [5]. Experimental data obtained from the THAI test series is used to validate and improve computational models used for the prediction of thermal-hydraulic flows and phenomena. Therefore, CFD methods are applied to model the THAI TH32 experiment [6], which investigates buoyancy-driven mixing processes between two miscible fluids within a vessel of volume $V \approx 60 \text{ m}^3$, while an underlying natural circulation flow is created by heating the lower and cooling the upper wall sections. The experiment comprises multiple consecutive phases, which are shown over time in Figure 2 and mimics the release and dispersion of hydrogen in a nuclear reactor containment. The first phase is the initial phase in which natural circulation of air is established. Heated air rises along the heating jackets, which are indicated in red. At the top part of the vessel the air is cooled at the cooling jackets, which are indicated in blue. This causes the air to sink and pass through the inner cylinder, which is located in the center of the containment and serves to guide the flow. This is followed by the injection phase, in which helium is injected into the test containment and the circulation is suppressed. After the injection process has been completed and a stable stratification has formed, the erosion phase of the helium layer begins. Once the circulation forms again, the stratification is eroded rapidly.

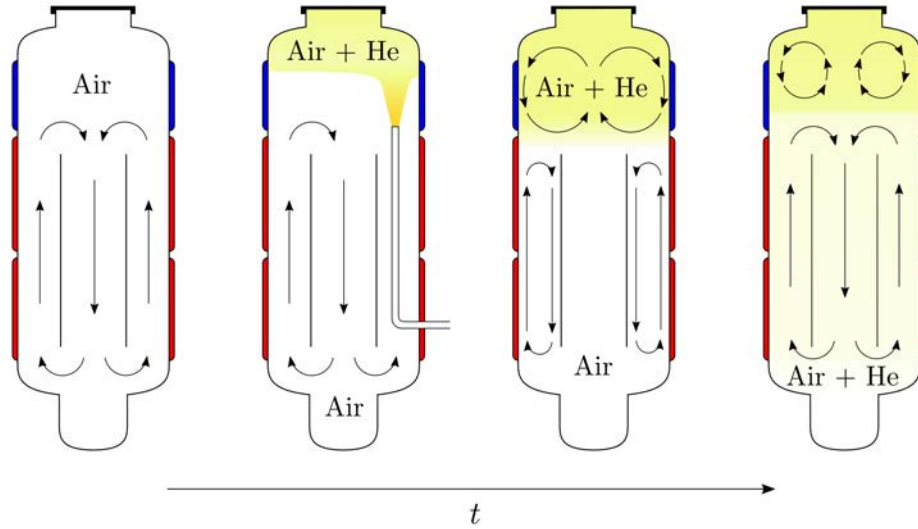


Figure 2: Consecutive phases of the THAI TH32 experiment.

In this application case, we investigate a low-Mach number flow of two Newtonian viscous fluids within the THAI containment using unsteady Reynolds-Averaged Navier-Stokes (URANS) simulations. To account for effects due to turbulence, the k - ω Shear Stress Transport (SST) turbulence model [7] is applied with included buoyancy terms. The gradient flux approach is utilized to consider turbulent heat and mass transport. The molecular diffusion coefficient was derived according to the Fuller's method [8]. The conjugated heat transfer between the fluid and the vessel walls or structures is taken into account, and thus the heat conduction equation in the solid region is additionally solved. A coupled boundary condition is used to ensure that the heat transfer between the solid and fluid regions is taken into account

at their shared interface. This boundary condition enforces the continuity of temperature and heat flux across the interface, thus thermally coupling the two regions. The exterior of the vessel walls is subject to a mixed boundary condition, which includes the definition of ambient temperatures and heat transfer coefficients. Radiation heat transfer within the containment is modelled with the finite volume Discrete Ordinates Method (fvDOM) [9]. Surface-to-surface radiation is considered. The fluid is assumed to be a non-participative medium in radiation. During the initial phase, experimentally measured temperatures are mapped to the solid domain and remain constant to establish an initial flow field and thermal stratification in the vessel. As the injection phase commences, the temperatures are no longer kept constant. Therefore, the temperature distribution results according to the given local thermal conditions. Inflow data during injection is defined according to given experimental data.

Table 1: Uncertain input parameters for the CFD simulation of the THAI TH32 experiment.

Physical Quantity	Region	Uncertainty
Initial structural temperatures T_0	Cooling Jacket Middle Heating Jacket Lower Heating Jacket	$\mathcal{U}(T_0^* - 2.5K, T_0^* + 2.5K)$
Heat transfer coefficient h	Cooling Jacket	$\mathcal{U}(0.8 \times h^*, 1.2 \times h^*)$
Helium mass flow rate during injection \dot{m}_{He}	Injection Pipe	$\mathcal{U}(0.98 \times \dot{m}_{\text{He}}^*, 1.02 \times \dot{m}_{\text{He}}^*)$

The CFD simulation of the THAI TH32 experiment requires the definition of a large number of input parameters. However, the definition of initial and boundary conditions is subject to uncertainties in the present setup, since measurement errors in the experiment occur or a lack of required input variables exists [10, 11]. For this reason, the propagation of predefined input uncertainties to response variables is investigated. From an initial global sensitivity analysis, the five most important parameters were selected and are considered in the following. The uncertain parameters, which are listed in Table 1, include initial structural temperatures, heat transfer coefficients at the outer walls of the containment and the helium mass flow rate during injection. The uncertainty of all parameters is defined with a uniform distribution with the lower and upper bounds a and b , respectively. This is denoted by the operator $\mathcal{U}(a, b)$. Quantities marked with an asterisk correspond to nominal values. The computational model of the containment and regions for the definition of initial and boundary conditions or uncertain parameters are depicted in Figure 3. The initial temperature of the lower heating jacket, the middle heating jacket and the cooling jacket is uniformly distributed within the interval $[T_0^* - 2.5K, T_0^* + 2.5K]$ in terms of their respective nominal temperatures. The heat transfer coefficient at the cooling jacket is defined with the lower bound $0.8 \times h^*$ and upper bound $1.2 \times h^*$. The injected helium mass is also considered uniformly distributed between $0.98 \times \dot{m}_{\text{He}}^*$ and $1.02 \times \dot{m}_{\text{He}}^*$.

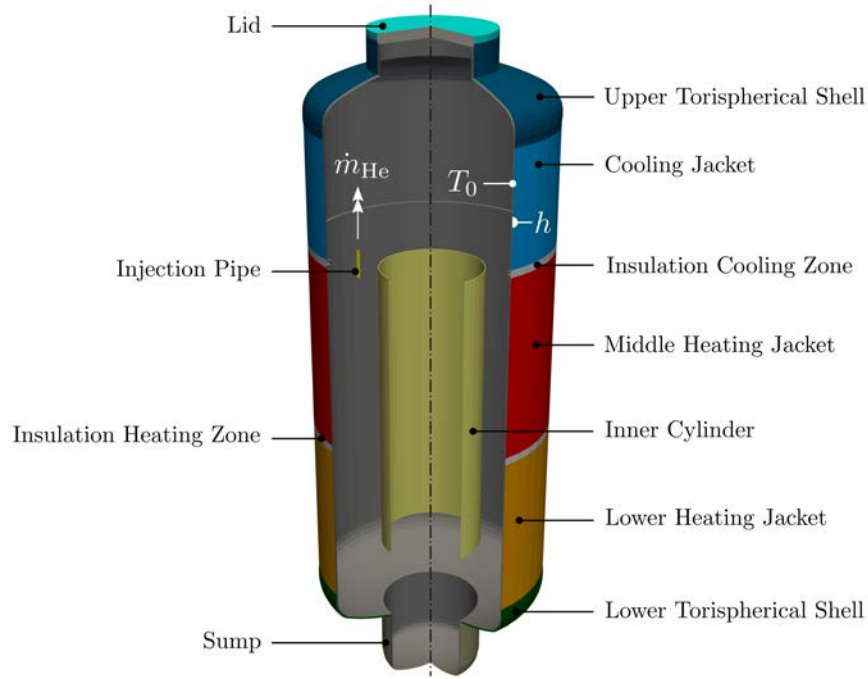


Figure 3: Computational model of the containment and regions for the definition of initial and boundary conditions.

2.1.2 Framework and Discretization

The set of equations, which govern the buoyancy-driven mixing process within the containment, is solved using the *containmentFOAM* solver [12], a tailored solver and model library based on *OpenFOAM*[®]. The pressure–velocity coupling was addressed by using the PIMPLE algorithm. It is ensured that the normalized residuals of the pressure-velocity coupling fall below the value 10^{-4} , while for the matrix solvers below 10^{-6} . The numerical schemes used in the present work are second-order accurate in time and space. Convection terms were discretized by second-order linear upwind. Gradients were evaluated with linear interpolation. Diffusive fluxes or Laplacian terms were evaluated with limited linear interpolation. The convective flux of the helium mass fraction was discretized by the limited linear scheme that is bounded between 0 and 1. Temporal advancement was achieved by blending of 10% Euler and 90% Crank–Nicolson scheme, as a compromise between accuracy and robustness. During the erosion phase, the CFL number is ensured to be below the value of 0.9. The underlying mesh comprises 1.3 Mio. Cells. To ensure appropriate resolution of the wall-boundary layer, average y^+ values are set below a value of 1 throughout the transient process.

2.2 UQ methods

Various mathematical techniques employed for stochastic modeling and model evaluation are explained in the following. Specifically, the model construction using PCEs is elaborated in Section 2.2.1, followed by the stochastic representation of random fields using Karhunen–Loève Expansions in Section 2.2.2. Applied techniques for the reliable prediction of noisy responses with stochastic models are introduced in Section 2.2.3. Additionally, Section 2.2.5 outlines one model evaluation aspect that involves variance-based decomposition, and Section 2.2.6 describes the methods used to estimate errors in the stochastic models.

2.2.1 Polynomial Chaos Expansion

The utilization of non-intrusive PCE [13, 14] was motivated by its ability to produce stochastic results with high convergence rates as the number of simulation runs increases. As a result, accurate results can be obtained with an adequate number of calculations. Let \mathbf{Q} be random input variables $\mathbf{Q} : \Omega \rightarrow \Upsilon \subset \mathbb{R}^n$ that map events $\omega \in \Omega$ from sample space Ω to realizations $\mathbf{q} \in \Upsilon$. In the employment of PCE, random response functions $\mathcal{R}(\omega) = R(\mathbf{Q})$ are approximated through a series of suitable multidimensional orthogonal polynomials $\Psi_{\beta}(\mathbf{Q})$ with corresponding expansion coefficients α_{β} . The expression for the truncated PCE is given in the following:

$$\mathcal{R}(\omega) \approx R^{\mathcal{B}}(\mathbf{Q}) = \sum_{\beta \in \mathcal{B}^{n,d}} \alpha_{\beta} \Psi_{\beta}(\mathbf{Q}), \quad (1)$$

where $\beta = (\beta_1, \dots, \beta_n)$ with $\beta_i \geq 0$ is an n -dimensional multi-index from the finite index set $\mathcal{B}^{n,d} = \{\beta \in \mathbb{N}^n : |\beta| \leq d\}$, which specifies the univariate polynomial degree tuples of each polynomial basis element. Therefore, Ψ_{β_i} denote degree- β_i orthogonal polynomials, which yield $\Psi_{\beta} = \prod_{i=1}^n \Psi_{\beta_i}(\mathbf{Q}_i)$. For computational purposes, the PCE is truncated by retaining a total polynomial degree of $|\beta| = \sum_{i=1}^n \beta_i \leq d$, which was chosen to be $d = 2$. This truncation leads to a total number of expansion terms P . Thus, PCE can also be formulated with a term-based index $p \in [1, P]$, which indicates the p -th multi-index $\beta^{(p)}$ in the polynomial basis $\mathcal{B}^{n,d}$:

$$\mathcal{R}(\omega) \approx R^P(\mathbf{Q}) = \sum_{p=1}^P \alpha_{\beta^{(p)}} \Psi_{\beta^{(p)}}(\mathbf{Q}). \quad (2)$$

The orthogonal polynomials were chosen to be Legendre polynomials as it provides the optimal basis for the uniform distributions of the random input variables. This is due to its orthogonality with respect to the corresponding probability density functions. The expansion coefficients $\alpha_{\beta^{(p)}}$ were estimated by using linear regression [15, 16]. In the present work, the ordinary least squares method (OLS) was applied and the coefficient estimations $\hat{\alpha}$ are determined according to

$$\hat{\alpha} = (\Psi^T \Psi)^{-1} \Psi^T \mathbf{R}, \quad (3)$$

where $\Psi \in \mathbb{R}^{S \times P}$ is the regressor matrix whose entries are defined as $\Psi_{sp} = \Psi_{\beta^{(p)}}(\mathbf{q}^{(s)})$ from a finite number of simulation runs S , which result from a Latin hypercube sample set with the input parameter vectors $\mathbf{q}^{(s)}$. In addition, \mathbf{R} is the response vector for different input samples $\mathbf{R} = \{R(\mathbf{q}^{(1)}), \dots, R(\mathbf{q}^{(S)})\}^T$.

A collocation ratio of 2 was applied, which denotes the ratio of simulation runs per PCE term. Therefore, 42 simulations runs were conducted for the determination of 21 PCE coefficients. The open-source software Dakota 6.10 [17] was used as framework for the determination of polynomial chaos expansions.

2.2.2 Karhunen-Loève Expansion

Stochastic processes or random fields were approximated using Karhunen-Loève Expansion [18, 19], also known as Principal Component Analysis (PCA). The method is also closely related to Proper Orthogonal Decomposition (POD), which is commonly used in fluid dynamics.

A stochastic model is constructed through a finite linear combination of orthogonal deterministic basis functions, or principal components, multiplied by uncorrelated random variables. With this approach, the variance in the random response data is captured through its most important modes of variability. Let $\mathcal{R}(\mathbf{x}, \omega)$ be a random field (RF), which is dependent on the deterministic field variable $\mathbf{x} \in \mathcal{X}$ and random event $\omega \in \Omega$, where \mathcal{X} denotes the field domain and Ω is the sample space. The RF can be divided into its mean $\mu_{\mathcal{R}}(\mathbf{x})$ and centered field $\mathcal{R}_0(\mathbf{x}, \omega)$ in the following way:

$$\mathcal{R}(\mathbf{x}, \omega) = \mu_{\mathcal{R}}(\mathbf{x}) + \mathcal{R}_0(\mathbf{x}, \omega) \quad (4)$$

Subsequently, a truncated KLE is applied to approximate the centered field $\mathcal{R}_0(\mathbf{x}, \omega)$

$$\mathcal{R}_0(\mathbf{x}, \omega) \approx R_0^K(\mathbf{x}, \boldsymbol{\zeta}) = \sum_{k=1}^K \sqrt{\lambda_k} \varphi_k(\mathbf{x}) \zeta_k, \quad (5)$$

where $\boldsymbol{\zeta} = \zeta_k(\omega)$ are mutually uncorrelated random variables with zero mean and unit variance. The scalars λ_k and field-dependent deterministic functions $\varphi_k(\mathbf{x})$ are, respectively, the eigenvalues and orthogonal eigenfunctions to the homogeneous Fredholm equation of second kind. For solving the integral within the Fredholm equation, the discrete Karhunen-Loève method [20] was employed along with a uniform discretization of the field domain with B points. The expectation of the random field $\mathcal{R}(\mathbf{x}, \omega)$ was determined by utilizing OLS to first approximate the random field using PCE and taking the respective coefficients of the first PCE term $\hat{\boldsymbol{\alpha}}_{\beta(1)} = \{\alpha_{\beta(1)}^1, \dots, \alpha_{\beta(1)}^B\}$, which correspond to the estimated expectation of the response at different field locations. The covariance matrix was constructed in an analogous manner through the approximation of $[R_0(x_a, \mathbf{Q}) R_0(x_b, \mathbf{Q})]$ via PCE and the determination of its expectation by picking the first PCE term. The truncation of the KLE was accomplished through a practical rule based on the eigenvalue decay rate, i.e. terms with eigenvalues λ_k , which decayed to some fraction (e.g. 10%) of the largest eigenvalue λ_1 , are omitted. In the present work, the decay rate $\lambda_{K+1}/\lambda_1 \leq 0.02$ was applied. The mutually uncorrelated random variables ζ_k were approximated with PCE [21, 22], to establish the relationship between \mathbf{Q}_i and ζ_k :

$$\zeta_k^P(\omega) = \sum_{p=1}^P \gamma_{k, \beta(p)} \Psi_{\beta(p)}(\mathbf{Q}). \quad (6)$$

The substitution of $\zeta_k^P(\omega)$ into the KLE formulation yields a PCE based on KLE with field-dependent expansion coefficients:

$$\mathcal{R}(\mathbf{x}, \omega) \approx \mu_{\mathcal{R}}(\mathbf{x}) + \sum_{p=1}^P \left[\alpha_{\beta(p)}^K(\mathbf{x}) \right] \Psi_{\beta(p)}(\mathbf{Q}) \quad (7)$$

KLE provides an optimal representation of the underlying random field in the least-square sense, meaning it minimizes the mean squared error (MSE) between the approximation and the original field with a given number of terms. This property might facilitate an efficient construction of a random field approximation while preserving essential characteristics of the random field.

2.2.3 Noisy response prediction

Turbulent fluctuations are an important and ubiquitous phenomenon in many engineering applications. These fluctuations can generate noise in the data obtained from such systems, making it challenging to extract accurate information and insights. In addition, the noise caused by turbulent fluctuations can be a major source of uncertainty. Therefore, it is crucial to consider this in particular and apply appropriate methods to derive accurate stochastic models, which provide reliable prediction of uncertainties.

In the present work, the computed time series $\mathbf{R}_L = \{r_1, \dots, r_L\}$ with L time instances were decomposed into their trend components $\bar{\mathbf{R}}_L$ and their fluctuating components \mathbf{R}'_L . This decomposition was accomplished through the data-driven filtering approach called Singular Spectrum Analysis (SSA) [23, 24], such that $\mathbf{R}_L \approx \bar{\mathbf{R}}_L + \mathbf{R}'_L$. This enabled the independent analyses of the constituent subseries and the construction of individual stochastic models. In the context of UQ in the field of CFD, these time series are discrete-time realizations of the corresponding random response functions:

$$\mathcal{R}(t, \omega) = \bar{\mathcal{R}}(t, \omega) + \mathcal{R}'(t, \omega) \quad (8)$$

where $\mathcal{R}(t, \omega)$, $\bar{\mathcal{R}}(t, \omega)$ and $\mathcal{R}'(t, \omega)$ correspond to \mathbf{R}_L , $\bar{\mathbf{R}}_L$ and \mathbf{R}'_L , respectively. The random response $\mathcal{R}(t, \omega)$ consists of its trend $\bar{\mathcal{R}}(t, \omega)$ and fluctuating component $\mathcal{R}'(t, \omega)$. The construction of a stochastic model for $\bar{\mathcal{R}}(t, \omega)$ is conducted using the methodology outlined in Sections 2.2.1 and 2.2.2. The obtained model is based on KLE in conjunction with PCE and is derived as a function of the input variables:

$$\bar{\mathcal{R}}(t, \omega) \approx \bar{R}^{P,K}(t, \mathbf{Q}) = \mu_{\bar{\mathcal{R}}}(t) + \sum_{p=1}^P \left[\bar{\alpha}_{\beta^{(p)}}^K(t) \right] \Psi_{\beta^{(p)}}(\mathbf{Q}). \quad (9)$$

The fluctuating component, on the other hand, is not assumed to be a function of the random input parameters, because the turbulent fluctuations are an inherent property of the underlying flow and are not significantly altered by small changes in the input parameters. For this reason, the fluctuations are considered as Gaussian processes [25] and are approximated via the KLE method:

$$\mathcal{R}'(t, \omega) \approx [R']^{K'}(t, \boldsymbol{\zeta}') = \mu_{\mathcal{R}'}(t) + \sum_{k=1}^{K'} \sqrt{\lambda_k'} \varphi_k'(t) \zeta_k'. \quad (10)$$

It is worth noting that in this way the complexity of these fluctuations can be adequately captured without the need for a substantial number of computational runs. Finally, the approximation of the original stochastic process is obtained by adding the two submodels:

$$\begin{aligned} \mathcal{R}(t, \omega) \approx R^{P,K,K'}(t, \mathbf{Q}, \boldsymbol{\zeta}') &= \mu_{\bar{\mathcal{R}}}(t) + \sum_{p=1}^P \left[\bar{\alpha}_{\beta^{(p)}}^K(t) \right] \Psi_{\beta^{(p)}}(\mathbf{Q}) + \\ &\mu_{\mathcal{R}'}(t) + \sum_{k=1}^{K'} \sqrt{\lambda_k'} \varphi_k'(t) \zeta_k'. \end{aligned} \quad (11)$$

This approach enables accurate model construction according to the characteristics of the computational data. In addition, uncertainties could be attributed to both random input parameters and fluctuations and thus allows for better result interpretation and informed decision making.

2.2.4 Stochastic Model Composition

The field of CFD often involves the prediction of transient processes that are characterized by highly dynamic behavior. As a result, the stochastic approximation of random responses in the presence of uncertainties might be impeded. To address this issue, Stochastic Model Composition (SMC), or Random Field Composition (RFC), was introduced in [4], as a novel approach for modeling time-dependent stochastic processes with function composition. In the present work, the SMC methodology is employed to perform UQ of the buoyancy-induced mixing process. The progress of desired responses during the mixing process can be described by the mixing state, which starts from an inhomogeneous state and ends with a fully mixed homogeneous state. Therefore, the trend and fluctuating components of the response can be represented as functions of the mixing state. This allows for a reduction in the complexity of the response function, provided that the considered mixing processes exhibit similar phenomenological behavior with respect to the mixing state. To establish the relationship between the mixing state with physical time, an additional stochastic model is constructed, which represents the mixing state as a function of time. Subsequently, the state-dependent stochastic models of responses are combined with the time-dependent stochastic model of the mixing state through function composition. In this way, SMC enables the approximation of intricate stochastic processes $\mathcal{R}(t, \omega)$ by partitioning the stochasticity into two stochastic models.

2.2.5 Variance-based decomposition

Through the application of variance-based measures of sensitivity [26], it becomes possible to identify the sources of response uncertainties and attribute them to uncertainties in input parameters or to inherent random fluctuations, e.g., due to turbulence. The stochastic process was decomposed into its trend and fluctuating component: $\mathcal{R}(t, \omega) = \bar{\mathcal{R}}(t, \omega) + \mathcal{R}'(t, \omega)$. Hence, the total variance $\mathbb{V}\text{ar}(\mathcal{R})$ is the sum of the individual variances and the twofold covariance of both stochastic processes:

$$\mathbb{V}\text{ar}(\mathcal{R}) = \mathbb{V}\text{ar}(\bar{\mathcal{R}} + \mathcal{R}') = \mathbb{V}\text{ar}(\bar{\mathcal{R}}) + \mathbb{V}\text{ar}(\mathcal{R}') + 2 \mathbb{C}\text{ov}(\bar{\mathcal{R}}, \mathcal{R}') , \quad (12)$$

where $\mathbb{V}\text{ar}(\bar{\mathcal{R}})$ represents the variance due to the trend evolution of the stochastic process. $\mathbb{V}\text{ar}(\mathcal{R}')$ denotes the variance caused by the fluctuations. $\mathbb{C}\text{ov}(\bar{\mathcal{R}}, \mathcal{R}')$, on the other hand, accounts for a correlation between $\bar{\mathcal{R}}(t, \omega)$ and $\mathcal{R}'(t, \omega)$ and thus for their joint variability. The trend response is a function of the random input variables, and therefore the corresponding variance fraction can be further split into the variances that arise from individual parameters or from their interaction behavior:

$$\mathbb{V}\text{ar}(\bar{\mathcal{R}}) = \sum_{i=1}^n V_{Q_i} + \sum_{i < j}^n V_{Q_i Q_j} + \cdots + V_{Q_i Q_j \dots Q_n} . \quad (13)$$

From this, variance-based measures of sensitivity can be inferred. In the present work total-order Sobol indices $S_{T_{Q_i}}$ are considered, which characterize the overall variance caused by an individual parameter through its univariate effect and interaction with other parameters:

$$S_{T_{Q_i}} = \frac{V_{T_{Q_i}}}{\mathbb{V}\text{ar}(\mathcal{R})} , \quad V_{T_{Q_i}} = V_{Q_i} + \sum_{\substack{j=1 \\ j \neq i}}^n V_{Q_i Q_j} + \cdots + V_{Q_i Q_j \dots Q_n} . \quad (14)$$

Sobol indices for the variance of the fluctuations and for the covariance between the two stochastic subprocesses were also considered. For this purpose, these variance components were examined without an additional partitioning. Thus, the total-order Sobol indices $S_{T_{\langle \cdot \rangle}}$ correspond to the first-order Sobol indices $S_{\langle \cdot \rangle}$ and are defined as follows:

$$S_{T_{\mathcal{R}'}} = S_{\mathcal{R}'} = \frac{\mathbb{V}\text{ar}(\mathcal{R}')}{\mathbb{V}\text{ar}(\mathcal{R})}, \quad S_{T_{\text{cov}}} = S_{\text{cov}} = \frac{2 \mathbb{C}\text{ov}(\bar{\mathcal{R}}, \mathcal{R}')}{\mathbb{V}\text{ar}(\mathcal{R})}, \quad (15)$$

where $S_{T_{\mathcal{R}'}}$ denotes the total-order index with respect to the fluctuations and $S_{T_{\text{cov}}}$ refers to the total-order index of the covariance.

The variance-based decomposition is a useful tool to analyze stochastic processes, as it enables the allocation of response uncertainties to input uncertainties and allows for the comprehensive assessment of the stochastic processes.

2.2.6 Model error estimation

The accuracy of the stochastic models is quantified by the estimation of the root-mean-square error (RMSE), which corresponds to the root of the generalization error [27, 28] in the context of statistical learning. For the determination of the error measure, the trend component of the response was taken into account. Hence, the RMSE is defined as:

$$\text{RMSE} = \left(\mathbb{E} \left[\left(\hat{R}(t, \mathcal{Q}) - \bar{R}(t, \omega) \right)^2 \right] \right)^{-1/2}, \quad (16)$$

where $\hat{R}(t, \mathcal{Q})$ denotes the stochastic model, which approximates the stochastic process of the trend response $\bar{R}(t, \omega)$. The RMSE is the square root of the expectation of the squared differences between the predicted and actual process. In this way, the accuracy of the stochastic model predictions can be assessed with respect to the underlying computational model.

The expectation of the squared error or the mean square error (MSE) was estimated by initially approximating it through PCE with the available computational runs. The extraction of the respective coefficients of the first PCE term $\hat{\alpha}_{\beta(1)}$ provide an estimate of its expectation, as already described analogously in Section 2.2.2: The square root of the estimated MSE yields an approximation of the root-mean-square error $\text{RMSE} = \sqrt{\text{MSE}}$. As global measure, the temporal mean RMSE, denoted by $\langle \text{RMSE} \rangle$, is determined through the square root of the integral mean of the MSE over the whole time span of the model $\Delta t = t_{\text{end}} - t_{\text{start}}$. It is determined as follows:

$$\langle \text{RMSE} \rangle = \left(\Delta t^{-1} \int_{t_{\text{start}}}^{t_{\text{end}}} \text{MSE} \, dt \right)^{-1/2}. \quad (17)$$

These error measures allow for the evaluation of the stochastic model accuracy and for the identification of potential model improvements.

3 RESULTS

In the following, results for the probabilistic characterization of selected result quantities are presented. These are referred to as Quantities of Interest (QoI) in the context of UQ. The natural convection flow within the THAI test vessel leads to the gradual diffusive and convective mixing of air and helium until the inhomogeneous mixture reaches the homogenous state. The temporal progress of the mixing process is quantified by the mixture uniformity

$\sigma_X = (V^{-1} \int_V (X - \bar{X})^2 dV)^{1/2}$, which is the volume-weighted standard deviation of the mole fraction X from the homogeneous equilibrium state mole fraction \bar{X} over the total fluid volume V . In addition, heat transfer is of great importance in natural convection flows. For this reason, the integral heat flow rate $\dot{Q} = \int_A \vec{q} \cdot d\vec{a}$ into the fluid is taken into account, where \vec{q} denotes the heat flux at the enclosing walls of the fluid, $d\vec{a}$ are surface vectors and A is the total area of the enclosing walls. The third QoI, which is considered in the present work, is the global kinetic energy $\langle E_k \rangle_M = M^{-1} \int_M u^2/2 \, dm$. m denotes the mass in a cell and M denotes the total mass of the gas mixture in the fluid domain. The global kinetic energy provides information about occurring convection mechanisms.

A comprehensive stochastic representation of the mixture uniformity σ_X is given in Figure 4. The left subplot shows the probability density functions (PDFs) $f_{\mathcal{R}}(R)$ at different points in time. In addition, discrete levels of the cumulative distribution functions (CDFs) $F_{\mathcal{R}}(R)$ are indicated with colors according to a colormap. The top right subplot shows further statistics like expectation μ , standard deviation σ , median $Q_{0.50}$ along with quantiles Q , which are located at the discrete color transitions according to the values of $F_{\mathcal{R}}(R)$. The center right subplot shows the error estimates of the underlying stochastic model by means of the RMSE. Furthermore, stacked total-order Sobol indices in the bottom right subplot enable the assignment of result uncertainties to uncertain input parameters or uncertainties due to fluctuations. From the left and top right plot one can see that the mixture uniformity σ_X increases in the time interval between $t = 0$ s and $t = 234$ s, which corresponds to the start and the end of the injection. After the injection phase, σ_X gradually decreases with a narrow uncertainty band. In this phase the mixing is mostly driven by diffusive mass transport. At a certain point in time, the erosion phase takes place and the helium cloud is eroded by the reestablishment of the circulation flow.

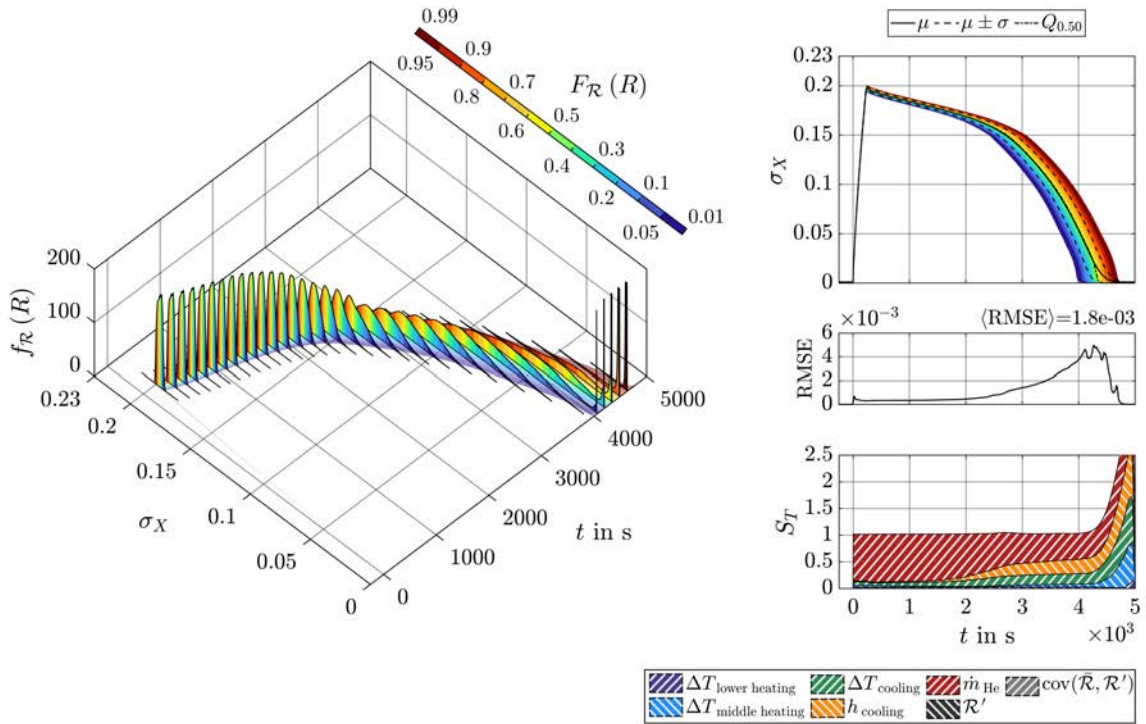


Figure 4: Probabilistic characterization of σ_X with PDFs $f_{\mathcal{R}}(R)$, discrete levels of the CDFs $F_{\mathcal{R}}(R)$, expectation μ , standard deviation σ , quantiles Q , RMSE estimation and stacked total-order Sobol indices S_T .

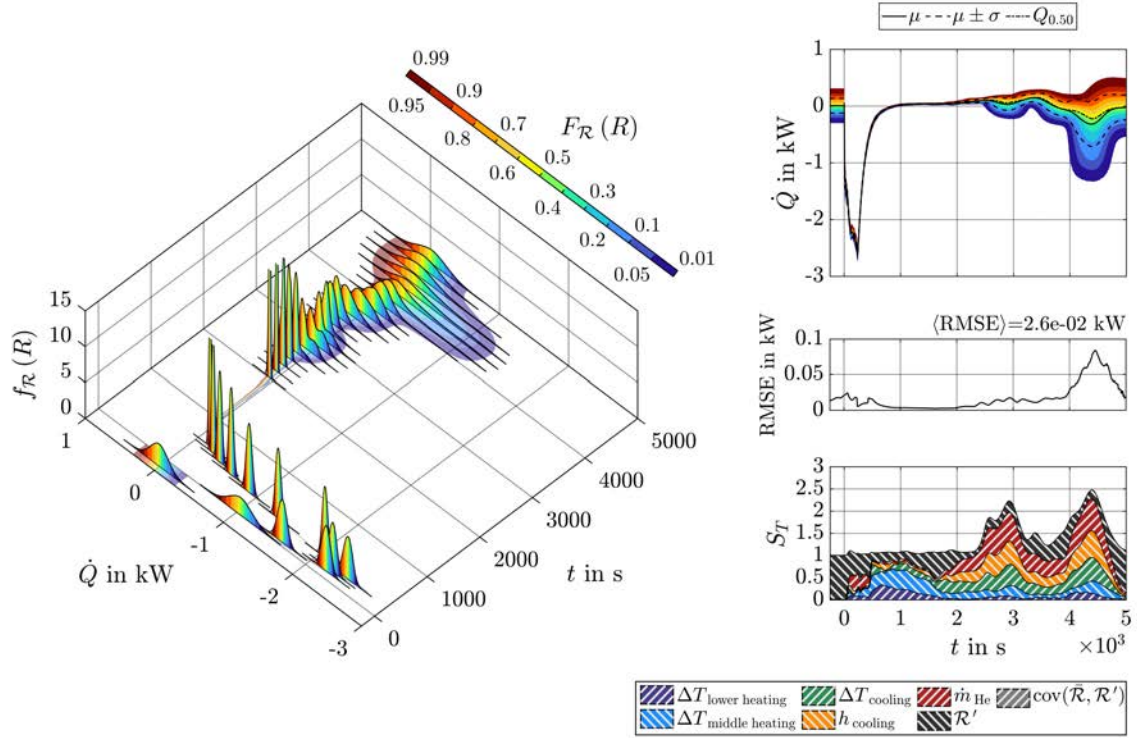


Figure 5: Probabilistic characterization of \dot{Q} with PDFs $f_{\mathcal{R}}(R)$, discrete levels of the CDFs $F_{\mathcal{R}}(R)$, expectation μ , standard deviation σ , quantiles Q , RMSE estimation and stacked total-order Sobol indices S_T .

Therefore, the mixture uniformity decreases with a steeper gradient. However, the start time of this erosion is subject to uncertainty and therefore takes place at different points in time. This leads to a progressively wider uncertainty band starting from $t = 2500$ s. At the end of the mixing process, the mixture uniformity decreases to zero and the probability accumulates around this value. The RMSE error estimates in the center right subplot show negligibly small values. The stochastic model thus is an accurate approximation of the result quantity and reliable statistical information can be derived from the model. In the stacked total-order Sobol indices subplot at the bottom right, the red area indicates that the helium mass flow rate during injection \dot{m}_{He} causes a major part of the uncertainty of σ_X . With the start of the erosion phase, the heat transfer coefficient at the cooling jacket h_{cooling} and the initial temperature of the cooling jacket $T_{0/\text{cooling}}$ also show significant impact.

The UQ results for the integral heat flow rate \dot{Q} are depicted in Figure 5. Based on the left and top right subplots, it can be inferred that \dot{Q} is also subject to uncertainty in the initial phase ($t < 0$ s). During the injection, the net heat flow rate takes on negative values, indicating that heat is transferred from the fluid to solid structures, since the fluid is heated up during compression by approximately 0.2 bar due to helium injection. Subsequently, the net heat flow rate rises to zero and stagnates with a narrow uncertainty band, i.e. only minor local convection takes place. Following the initiation of the erosion phase at approximately $t = 2500$ s, a transient process characterized by temporal oscillations occurs prior to attaining the new equilibrium state at the end. Before finally reaching the new state, an undershoot of \dot{Q} between $t = 4000$ s and $t = 4800$ s occurs, when the mixing is completed and the full vessel surface participates in heat transfer. It lasts until the vessel atmosphere is again in thermal equilibrium and leads to a considerable increase in the uncertainty band. The stacked total-order Sobol indices in the bottom

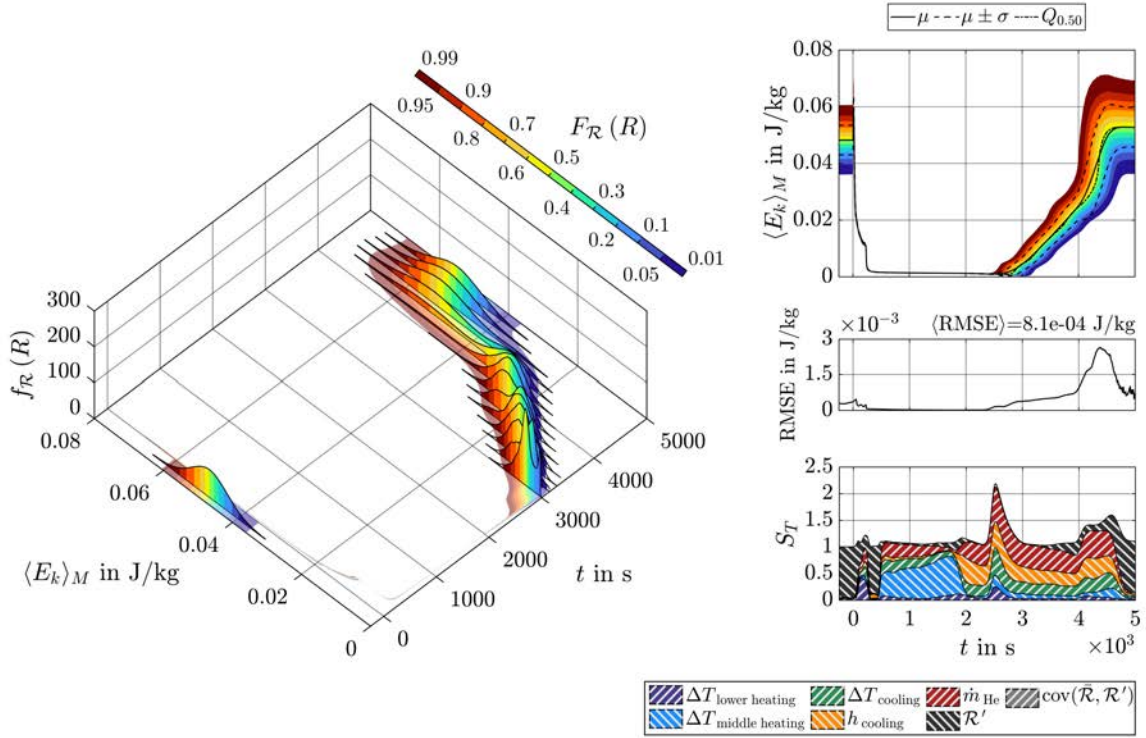


Figure 6: Probabilistic characterization of $\langle E_k \rangle_M$ with PDFs $f_{\mathcal{R}}(R)$, discrete levels of the CDFs $F_{\mathcal{R}}(R)$, expectation μ , standard deviation σ , quantiles Q , RMSE estimation and stacked total-order Sobol indices S_T .

right subplot reveal that fluctuations \mathcal{R}' contribute significantly to the uncertainty, as represented by the dark grey area indicating the variance components caused by \mathcal{R}' . During the transient transition stages, the input parameter uncertainties also show significant influence on the result uncertainty. The reliability of the model predictions is ensured, because the center right subplot merely shows moderate values for the error estimate at the time when the undershoot occurs.

Results for the global kinetic energy $\langle E_k \rangle_M$ are given in Figure 6. From the left and top right figure, it can be seen that the global kinetic energy in the initial phase ($t < 0$ s) is associated with uncertainty, due to fluctuations. During the injection, $\langle E_k \rangle_M$ decreases significantly, because the circulation flow within the containment is suppressed. As the amplitude of fluctuations also reduces, the width of the uncertainty band shrinks. The global kinetic energy stagnates up to around $t = 2500$ s. Afterwards, the circulation flow reestablishes and $\langle E_k \rangle_M$ gradually increases till it reaches its new equilibrium state. The reestablishment of the circulation flow takes place at different points in time. Therefore, the uncertainty band shows a larger width from $t > 2500$ s. From the total-order Sobol indices, it is evident, that the uncertainty during the initial phase and after reaching the new equilibrium state is mostly caused through fluctuations. The uncertainty during the mixing process, on the other hand, is mainly determined by input uncertainties. The RMSE estimates take on moderate values during the transition stage to the new equilibrium state. In the remaining temporal domain the estimated error is negligibly small.

4 SUMMARY AND CONCLUSION

In the present work, UQ of an industrial scale application in the field of CFD was successfully conducted. The application case refers to a technical scale experiment, which investigates the buoyancy-driven mixing processes between two miscible fluids within a vessel of $V \approx 60 \text{ m}^3$,

while an underlying natural circulation flow is created by heating the lower and cooling the upper wall sections. The CFD model for this experiment is subject to a number of uncertainties in initial and boundary conditions along with thermo-physical properties. The propagation of these uncertainties to response quantities was quantified. As UQ methodology, stochastic spectral methods, such as PCE and KLE, were used for the approximation of stochastic results and were proven a promising method for UQ of technical scale computations. For the accurate prediction of the underlying phenomena, the novel SMC method was applied, which allows for the computationally efficient representation of a respective stochastic process by partitioning its stochasticity into two stochastic models. In addition, response transients were decomposed into their trend and fluctuating components using SSA. This enabled the construction of tailored stochastic models for the trend process and its fluctuating counterpart. Based on the constructed models, a variance-based decomposition was conducted by means of total-order Sobol indices. Through this approach, proportions of the response uncertainties were attributed to uncertainties in input parameters or to fluctuations. Supplementary RMSE error estimates were determined for the assessment of the respective model accuracy.

It is evident from the UQ results for σ_X , that the uncertain input parameters lead to noticeable uncertainty in the duration of the mixing process. Due to the uncertain timing of consecutive mixing phenomena, uncertainties from input parameters additionally lead to considerable uncertainties in the temporal evolution of the integral heat flow rate \dot{Q} and the global kinetic energy $\langle E_k \rangle_M$. By means of the total-order Sobol indices S_T , the helium mass flow rate during injection \dot{m}_{He} , the heat transfer coefficient at the cooling jacket h_{cooling} and the initial temperature of the cooling jacket $T_{0/\text{cooling}}$ were identified as the most important parameters. Furthermore, the results for S_T indicate that fluctuations, which are caused through turbulent phenomena, can have a significant impact on the result uncertainty and therefore have to be taken into account for UQ in the field of CFD. Modest values for the RMSE error estimates demonstrate the good performance of the stochastic models. With a total computational cost of 2.1 Mio CPUh, a number of 42 parametric CFD simulations of the considered application case were conducted for the statistical assessment. Therefore, the applied UQ framework enabled the accurate stochastic representation of uncertain responses in a computationally efficient manner. In addition, it enabled the derivation of detailed statistical information for the comprehensive probabilistic description of uncertain responses and thus allows for better result interpretation and informed decision making.

REFERENCES

- [1] P. J. Wenig, R. Ji, S. Kelm, M. Klein, Towards Uncertainty Quantification of LES and URANS for the Buoyancy-Driven Mixing Process between two Miscible Fluids - Differentially Heated Cavity of Aspect Ratio 4, *Fluids* 6. doi:10.3390/FLUIDS6040161.
- [2] P. J. Wenig, R. Ji, S. Kelm, M. Klein, Uncertainty Quantification for the Buoyancy-Driven Mixing Process between two Miscible Fluids using Multifidelity Polynomial Chaos Expansions, in: ETMM-13, 2021.
- [3] P. J. Wenig, S. Kelm, M. Klein, Uncertainty Quantification of LES for Buoyancy-Driven Mixing Processes using PCE-HDMR, in: Proceedings of DLES-13, 2022.

- [4] P. J. Wenig, S. Kelm, M. Klein, CFD Uncertainty Quantification using Stochastic Spectral Methods - Exemplary Application to a Buoyancy-Driven Mixing Process, *Nuclear Engineering and Design* doi:10.2139/ssrn.4157189.
- [5] S. Gupta, E. Schmidt, B. Von Laufenberg, M. Freitag, G. Poss, F. Funke, G. Weber, THAI test facility for experimental research on hydrogen and fission product behaviour in light water reactor containments, *Nuclear Engineering and Design* 294 (2015) 183–201. doi:10.1016/j.nucengdes.2015.09.013.
- [6] M. Freitag, E. Schmidt, Simulation Benchmark Based on THAI - Experiment on Generation and Dissolution of a Light Gas Stratification By Natural Convection, in: *NURETH-19*, 2022, pp. 1–17.
- [7] F. R. Menter, Two-equation eddy-viscosity turbulence models for engineering applications, *AIAA Journal* 32 (8) (1994) 1598–1605. doi:10.2514/3.12149.
- [8] S. Kabelac, M. Kind, H. Martin, D. Mewes, K. Schaber, P. Stephan, *VDI-Wärmeatlas*, 11th Edition, 2013.
- [9] S. C. Mishra, H. K. Roy, Solving transient conduction and radiation heat transfer problems using the lattice Boltzmann method and the finite volume method, *Journal of Computational Physics* 223 (1) (2007) 89–107. doi:10.1016/j.jcp.2006.08.021.
- [10] D. Bestion, R. Camy, S. Gupta, C. Heib, O. Marfaing, F. Moretti, T. Skorek, Requirements for cfd-grade experiments for nuclear reactor thermalhydraulics, in: *CFD4NRS: Computational Fluid Dynamics for Nuclear Reactor Safety*, Shanghai, China, 2019, p. 20. URL <https://hal-cea.archives-ouvertes.fr/cea-02328974>
- [11] B. L. Smith, The difference between traditional experiments and CFD validation benchmark experiments, *Nuclear Engineering and Design* 312 (2017) 42–47. doi:10.1016/j.nucengdes.2016.10.007.
- [12] S. Kelm, M. Kampili, X. Liu, A. George, D. Schumacher, C. Druska, S. Struth, A. Kuhr, L. Ramacher, H. J. Allelein, K. A. Prakash, G. V. Kumar, L. M. Cammiade, R. Ji, The tailored CFD package 'containmentFOAM' for analysis of containment atmosphere mixing, H₂/CO mitigation and aerosol transport, *Fluids* 6. doi:10.3390/FLUIDS6030100.
- [13] N. Wiener, The Homogeneous Chaos, *American Journal of Mathematics* 60 (4) (1938) 897–936. doi:10.2307/2371268.
- [14] R. G. Ghanem, P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer New York, NY, 1991. doi:10.1007/978-1-4612-3094-6.
- [15] S. Hosder, R. W. Walters, M. Balch, Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables, in: *Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Vol. 3, Honolulu, HI, 2007, pp. 2946–2961. doi:10.2514/6.2007-1939.
- [16] R. W. Walters, Towards stochastic fluid mechanics via polynomial chaos, in: *Proceedings of the 41st AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, 2003.

- [17] B. Adams, W. Bohnhoff, K. Dalbey, M. Ebeida, J. Eddy, M. Eldred, R. Hooper, P. Hough, K. Hu, J. Jakeman, M. Khalil, K. Maupin, J. Monschke, E. Ridgway, A. Rushdi, D. Seidl, J. Stephens, L. Swiler, J. Winokur, Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.10 User's Manual, Sandia Technical Report SAND2014-4633, May 2019.
- [18] K. Karhunen, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.* 37 (1947) 1–79.
- [19] M. Loève, Probability theory. Vol. II, 4th ed. Graduate Texts in Mathematics., 4th Edition, Springer New York, 1978.
- [20] C. A. Schenk, G. I. Schuëller, Uncertainty Assessment of Large Finite Element Systems, 1st Edition, Springer Berlin Heidelberg, 2005. doi:10.1007/11673941.
- [21] X. Huan, C. Safta, Z. P. Vane, G. Lacaze, J. C. Oefelein, H. N. Najm, Uncertainty propagation using conditional random fields in large-eddy simulations of scramjet computations, *AIAA Scitech 2019 Forum* (2019) 1–17. doi:10.2514/6.2019-0724.
- [22] A. Jivani, X. Huan, C. Safta, B. Y. Zhou, N. R. Gauger, Uncertainty quantification for a turbulent round jet using multifidelity karhunen-loève expansions, *AIAA Scitech 2021 Forum* (2021) 1–16. doi:10.2514/6.2021-1367.
- [23] D. S. Broomhead, G. P. King, Extracting qualitative dynamics from experimental data, *Physica D: Nonlinear Phenomena* 20 (2-3) (1986) 217–236. doi:10.1016/0167-2789(86)90031-X.
- [24] H. Hassani, Singular Spectrum Analysis: Methodology and Comparison, *Journal of Data Science* 5 (2) (2021) 239–257. doi:10.6339/jds.2007.05(2).396.
- [25] C. Rasmussen, C. Williams, Gaussian processes for machine learning., in: Adaptive computation and machine learning, MIT Press, Cambridge, Massachusetts, 2006.
- [26] I. M. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Mathematics and Computers in Simulation* 55 (2001) 271–280. doi:10.1016/S0378-4754(00)00270-6.
- [27] V. N. Vapnik, The Nature of Statistical Learning Theory, 2nd Edition, Springer New York, NY, 2000. doi:10.1007/978-1-4757-3264-1.
- [28] G. Blatman, B. Sudret, Adaptive sparse polynomial chaos expansion based on least angle regression, *Journal of Computational Physics* 230 (6) (2011) 2345–2367. doi:10.1016/j.jcp.2010.12.021.