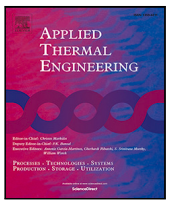
Applied Thermal Engineering 246 (2024) 122902

Contents lists available at ScienceDirect

Applied Thermal Engineering

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



Research Paper

Incremental sampling methods for multi-fidelity surrogate modeling of a MILD combustion furnace

A. Özden [a,b,c,*], A. Procacci [a,b], R. Malpica Galassi [d], F. Contino [c], A. Parente [a,b]

[a] Aero-Thermo-Mechanics Lab., Université Libre de Bruxelles, Avenue F. D. Roosevelt, Brussels, 1050, Belgium
[b] BRITE - Brussels Institute for Thermal-fluid Systems and Clean Energy, Brussels, Belgium
[c] Institute of Mechanics, Materials, and Civil Engineering, UC Louvain, Louvain-la-Neuve, Belgium
[d] Mechanical and Aerospace Engineering Department, Sapienza University of Rome, Via Eudossiana 18, Roma, Italy

ARTICLE INFO

Keywords:
Multi-fidelity ROM
Digital twin
Co-Kriging
Proper orthogonal decomposition
Manifold alignment

ABSTRACT

This study introduces a framework for developing a multi-fidelity reduced order model (MF-ROM) of a combustion furnace operating under Moderate and Intense Low-oxygen Dilution (MILD) conditions. It integrates Proper Orthogonal Decomposition for data compression, Procrustes manifold alignment for fidelity transfer, and CoKriging for interpolation. Design parameters such as air injector diameter, fuel composition, and equivalence ratio were used to generate two- and three-dimensional simulations and build the MF-ROM. Additionally, the key question concerning the optimal number of high-fidelity simulations to balance accuracy and training cost is addressed when building the MF-ROM. Through incremental sampling strategies, it is demonstrated that around half of the training cost can be conserved while maintaining comparable error values.

1. Introduction

The global energy crisis forces society to meet the increasing energy demands while addressing environmental concerns such as climate change and air pollution. Traditional combustion technologies, which have been the primary source of energy for decades, release significant amounts of greenhouse gases and pollutants that contribute to environmental deterioration [1]. The necessity to find solutions that can reduce greenhouse gas emissions, and improve air quality while enhancing energy efficiency and supporting long-term sustainability, leads scientists to develop environmentally friendly combustion technologies that can also accommodate fuel flexibility, in response to the availability of novel energy carriers. Moderate and Intense Low-oxygen Dilution (MILD) combustion [2] is a promising solution to reach emission targets with the capability of providing high combustion efficiency for a wide variety of fuels. MILD combustion carries the step-changing advantage of achieving strong reductions in NOx emissions, yielded by (i) its homogeneous nature and subsequent lower temperature peaks and thermal-NOx formation reduction, (ii) reduction of nitrogen availability due to dilution of air with combustion products, and (iii) alteration of NOx formation chemistry [3–5].

The physical understanding and the robust design of such a complicated multi-scale and multi-physics combustion system requires accurate yet inexpensive tools to explore and predict a potentially wide range of operational conditions and device geometries.

The improvements in computational fluid dynamics (CFD) tools are undeniable [6], but obtaining realistic predictions of complex flows in real-scale domains – which require the use of detailed chemical mechanisms and advanced turbulent-chemistry interactions – demands a large amount of computational resources. Hence, for real-scale problems, CFD simulations are not suitable for carrying out tasks such as design exploration, optimization, or real-time access to predictions, which require many model evaluations and, consequently, an excessive computational load. To this end, reduced-order models (ROMs) may be used in place of the expensive full-order models to approximate the underlying relationship between inputs (e.g., operating conditions) and outputs (e.g., the predictions of the system state), using available numerical and/or experimental observations to estimate the system response at new conditions. A ROM can be developed by extracting the key latent features to relate the quantities of interest within a predefined parameter space [7]. Once the relationship between inputs and outputs is discovered by the model, the system state can be predicted for new operating conditions. The modal analysis can be performed effectively by combining a method that extracts the low-dimensional subspace with a non-linear regression model to reconstruct and predict the complex reacting systems. A popular method for extracting this latent subspace is the Proper Orthogonal Decomposition (POD) [8]. Some ROM techniques use Galerkin methods [9] to project the governing equation of the physical model onto the extracted subspace.

* Corresponding author at: Aero-Thermo-Mechanics Lab., Université Libre de Bruxelles, Avenue F. D. Roosevelt, Brussels, 1050, Belgium.
E-mail address: aysu.ozden@ulb.be (A. Özden).

https://doi.org/10.1016/j.applthermaleng.2024.122902
Received 30 October 2023; Received in revised form 4 March 2024; Accepted 7 March 2024
Available online 11 March 2024

These are considered intrusive ROMs and they have the disadvantage of requiring access to the underlying simulation source code [10]. The alternatives are the ROMs that combine POD with a regression method, i.e., Kriging [11,12], in the latent space to predict the new solutions and treating the source code as a black box, hence their definition of non-intrusive ROMs. Non-intrusive ROMs are similar to traditional surrogate models but instead of estimating some integrated quantities (e.g. lift, drag, moment), they estimate the high-dimensional field solution, which provides a physically richer result [13]. Even though non-intrusive models are less accurate for some highly non-linear problems [14], they are more practical and flexible than intrusive ROMs.

ROMs are typically trained with a limited number of (expensive) model evaluations over a wide range of possible operating conditions and embed the critical parameters of detailed simulations into simplified relationships between the inputs and outputs, that can be used for predictions in real-time and/or other outer-loop applications.

Nonetheless, in the context of a combustion system that has highly non-linear responses and multiple design parameters, even the *limited* number of training data[1] becomes computationally expensive, therefore, the training of the ROM itself may become impractical, especially when dealing with large-scale systems. A possible solution can be achieved through the integration/fusion of a few high-fidelity observations/simulations and many cheaper low-fidelity ones, within a framework known as Multi-fidelity Reduced Order Modelling (MF-ROM). Different levels of fidelity may be obtained for CFD simulations by varying the physical model, the size of the computational grid, the computational time step, the convergence level of the simulations, and/or combining numerical data with experimental results [15]. Several methods can be used to fuse data and construct a multi-fidelity framework, including non-intrusive polynomial chaos [16], CoKriging [17], and stochastic radial basis function (RBF) [18]. Multi-fidelity methods have been successfully applied to different research areas in the past decade. Han et al. [19] proposed a CoKriging method for variable-fidelity surrogate modeling and showed that this method can potentially be applied to efficient aerodynamic data production and shape optimization on an RAE 2822 airfoil. Perron et al. [10] introduced a non-intrusive multi-fidelity ROM and assessed its performance on a transonic airfoil and proved that the computational training cost was reduced compared to a single-fidelity approach with comparable accuracy. Lamberti et al. [20] blended RANS and LES in a multi-fidelity framework to provide accurate wind load predictions for lower computational costs and showed the potential of using significant LES simulations for the design process while achieving higher accuracy than standard empirical models. A more comprehensive overview of multi-fidelity applications in different fields including uncertainty propagation, inference, and optimization can be found in [21].

In general, the fusion of low-fidelity and high-fidelity data in the construction of an MF-ROM responds to both an accuracy requirement and a computational budget availability. In other words, some high-fidelity (HiFi) simulations are replaced with low-fidelity (LoFi) ones to reduce the training cost, with the simultaneous goal of incurring in a limited accuracy loss of the resulting MF-ROM. Therefore, a fundamental question is: given a design of experiment (DoE), how many HiFi simulations can be replaced with LoFi ones to train an MF-ROM whose prediction error remains satisfactory?

In this paper, the effects of data fusion on the predictive performance of the MF-ROM of a semi-industrial combustion furnace that can operate in MILD combustion regime are addressed. The datasets are composed of 3-dimensional (HiFi) and 2-dimensional (LoFi) CFD simulations of the aforementioned furnace with a design space consisting of (i) the inlet fuel mixture proportions (CH_4-H_2), (ii) the fuel/air equivalence ratio, and (iii) the injector diameter. The goal is to analyze how the number and position in the DoE of the HiFi simulations play a role in the MF-ROM accuracy.

The paper is organized to first give a theoretical basis of the algorithmic and numerical approaches involved in developing the multi-fidelity framework. Then the details of the datasets used in this study are introduced and followed by the results obtained with an incremental sampling strategy for the HiFi data. Lastly, conclusions are drawn and prospects about future developments are shared.

2. Methodology

Data fusion is central to this workflow: a multi-fidelity surrogate model utilizes insights gathered from diverse HiFi and LoFi datasets, which could vary in dimensionality and topology. These datasets represent solutions to the same physical problem but with different levels of approximation and several inconsistencies (e.g., physical models, geometry, and dimensionality). Effective data fusion aims to uncover shared intrinsic representations, potentially of low dimensionality, across the different datasets. To achieve this, Proper Orthogonal Decomposition is employed to discover latent spaces within the data and identify correlations among the heterogeneous datasets. This newfound correlation is then leveraged through a manifold alignment technique. Once a unified latent representation is established, the latent variables from the training samples are mapped onto the design space, allowing for the creation of a continuous predictive model.

2.1. Proper orthogonal decomposition

Proper Orthogonal Decomposition (POD), also known as Karhunen–Loéve transformation [22] or Principal Component Analysis [23], is a dimensionality reduction technique applied to the data matrix $\mathbf{X} \in \mathbb{R}^{p\times n}$ which is decomposed into a matrix $\mathbf{\Phi} \in \mathbb{R}^{p\times n}$ which is the physics invariant information and a matrix $\mathbf{Z} \in \mathbb{R}^{n\times n}$ which is the parameter-dependent information, with the assumption that a large number of interrelated variables exists [24]:

$$\mathbf{X}(\boldsymbol{r},\boldsymbol{d}) = \mathbf{\Phi}(\boldsymbol{r})\mathbf{Z}^T(\boldsymbol{d}). \tag{1}$$

The training data matrix used in this study is constructed from the CFD simulation results, having $\boldsymbol{r}$ physical degrees of freedom, each one corresponding to a design condition $\boldsymbol{d} = \{d_i\}_{i=1,\dots,N_{\text{desVars}}}$ in the DoE, for a total of n designs, i.e. n simulations. Each simulation is rearranged as a column vector $\boldsymbol{x} \in \mathbb{R}^p$ whose length p is the number of computational cells times the number of features (i.e., temperature and chemical species mass fractions), $p = n_{cells} \times n_{features}$, and the data matrix $\mathbf{X}$ is formed by grouping n numerical simulations, i.e. $\mathbf{X} = \{\boldsymbol{x}_i\}_{i=1,\dots,n}$ where $p \gg n$ in typical combustion problems.

Dimensionality reduction is achieved by retaining only the first k eigenvectors, resulting in the truncated matrices $\mathbf{\Phi}_k$ and $\mathbf{Z}_k$ with dimensions $(p\times k)$ and $(n\times k)$, where $k < min(n,p)$. The last $n-k$ eigenvectors can be discarded while incurring in a minimal information loss because the POD algorithm ensures the minimization of the Frobenius norm of the reconstruction error $\mathbf{E} = \mathbf{X} - \mathbf{\Phi}_k \boldsymbol{Z}_k^T$.

The determination of the lower dimension k can be attained by calculating the singular value decomposition (SVD) of the matrix $\mathbf{X}$, which corresponds to solving the eigenproblem of the covariance matrix $\mathbf{C} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X}$. The lower dimension k is the number of retained eigenvectors and it was chosen according to the Relative Information content (RIC) value [25], which can be interpreted as the total variance captured by the POD. Commonly values such as 99.9% or above are chosen as the threshold. Therefore, it is possible to approximate the matrix $\mathbf{X}$ by using the truncated matrices $\mathbf{\Phi}_k$ and $\mathbf{Z}_k$ with dimensions $(p\times k)$ and $(n\times k)$ as:

$$\mathbf{X} \approx \mathbf{\Phi}_k \mathbf{Z}_k^T. \tag{2}$$

[1] One should associate one training data to one multi-dimensional reacting CFD simulation.

The columns of $\mathbf{\Phi}_k$ are called the POD modes, or eigenflames [26], and the matrix $\mathbf{Z}_k$ is called the POD coefficients matrix. Since the eigenvectors of $\mathbf{C}$ are orthogonal and the principal components are linear combinations of the original variables, the approximated/reconstructed system's response $\mathbf{x}_{rec}$ corresponding to a given set of POD coefficients $\mathbf{z}$ can be calculated as $\mathbf{x}_{rec} = \mathbf{\Phi}_k\, \boldsymbol{z}$. Therefore, the ROM is built by regressing the POD coefficients $\boldsymbol{z}$ over the design space, i.e., by building a model $\alpha^k = \alpha^k(\boldsymbol{d}; \boldsymbol{z})$ for each coefficient k, which allows one to estimate/predict at virtually no cost $\mathbf{x}_{pred}(\boldsymbol{d}^*) = \mathbf{\Phi}_k\ \boldsymbol{\alpha}(\boldsymbol{d}^*)$ for any unseen design condition $\boldsymbol{d}^*$. A typical regression technique is Kriging, as employed in [12].

2.2. Manifold alignment

In a multi-fidelity framework, even though the data are the results of equivalent problems, information transfer from one dataset to the other might be problematic due to topological or dimensionality differences.

Manifold alignment allows to identify a correlation structure between datasets and therefore transfers information from one manifold to the other by aligning their underlying low-dimensional representation [27]. The main assumption behind this technique is that these datasets do share a common low-dimensional representation which allows the comparison of the embedded data due to their intrinsic similarity [10]. Manifold alignment is applied as a useful technique in many different disciplines such as protein alignment [28], image matching and classification [29,30] or automatic machine translation [31]. In the literature, there are two types of manifold alignment approaches. The first method searches for a shared latent space with a semi-supervised approach by constructing a joint graph Laplacian and the resultant latent space should cover the individual and shared characteristics of both manifolds [10]. The other method uses dimensionality reduction techniques to discover the intrinsic latent space for each dataset and then apply Procrustes analysis to align both spaces.

The Procrustes manifold alignment technique is the logical approach to follow in this framework since dimensionality reduction is unavoidable considering the size of each dataset used. As mentioned before, the first step of this method is applying POD to identify the intrinsic latent space for each dataset. Then, the Procrustes analysis [32] is used to adjust the translational, rotational, and scaling components from one manifold to achieve an optimal alignment with the other.

HiFi and LoFi datasets are obtained as $\mathbf{X} \in \mathbb{R}^{p\times n}$ and $\mathbf{Y} \in \mathbb{R}^{q\times m}$, respectively. Here n and m indicate the number of observations with $n < m$, while p and q represent the value obtained by multiplication of grid points and features, $p = n_{cells,hf} \times n_{features}$ and $q = n_{cells,lf} \times n_{features}$. The LoFi dataset includes a linked subset $\mathbf{Y}_L \in \mathbb{R}^{q\times n}$ which shares the same input design parameters as the HiFi dataset, and the rest of the data belongs to the unlinked subset, $\mathbf{Y}_U \in \mathbb{R}^{q\times(m-n)}$, hence $\mathbf{Y} = [\mathbf{Y}_L, \mathbf{Y}_U]$. The first step of this method requires obtaining the reduced manifold for HiFi and LoFi datasets. Following the application of POD, for a latent subspace dimension of k, the eigenflames $\mathbf{\Phi} \in \mathbb{R}^{p\times k}$ and $\mathbf{\Psi} \in \mathbb{R}^{q\times k}$, with POD coefficients $\mathbf{Z} \in \mathbb{R}^{k\times n}$ and $\mathbf{W} \in \mathbb{R}^{k\times m}$, are obtained for HiFi and LoFi inputs, respectively. Consequently, the LoFi manifold $\mathbf{W}$ has also linked and unlinked coefficients subsets, $\mathbf{W} = [\mathbf{W}_L, \mathbf{W}_U]$. This orthogonal Procrustes problem is solved using the procedure of Wang and Mahadevan [28] which seeks the manifold $\mathbf{T}$ minimizing $\|\mathbf{Z} - \mathbf{T}\|$. The first step is to translate the linked manifolds $\mathbf{Z}$ and $\mathbf{W}_L$ to make sure that both were centered at the origin. To center the $\mathbf{W}_L$, the following shift is required:

$$\overline{\mathbf{W}}_L = \frac{1}{p}\sum_{i=1}^{p} \mathbf{W}_{L,i} = 0. \tag{3}$$

After translation, scaling and rotation are performed by calculating and applying a scaling factor (s) and a rotation matrix ($\mathbf{Q}$). These parameters are obtained through the computation of Singular Value Decomposition of $\mathbf{W}_L\mathbf{Z}^T$:

$$\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{W}_L\mathbf{Z}^T. \tag{4}$$

Here $\mathbf{U} \in \mathbb{R}^{k\times k}$, $\mathbf{V} \in \mathbb{R}^{k\times k}$, and $\mathbf{\Sigma} \in \mathbb{R}^{k\times k}$ are the left and right singular vectors, and the diagonal matrix containing the singular values, respectively. The scaling factor and the rotation matrix are calculated as follows:

$$\mathrm{s} = \frac{\mathrm{tr}(\mathbf{\Sigma})}{\mathrm{tr}(\mathbf{W}_\mathrm{L}\mathbf{W}_\mathrm{L}^\mathrm{T})}, \tag{5}$$

$$\mathbf{Q} = \mathbf{V}\mathbf{U}^T. \tag{6}$$

Finally, the transformed manifold is obtained as

$$\mathbf{T} = s\,\mathbf{Q}\mathbf{W}. \tag{7}$$

Manifold alignment introduces a transformed matrix which is a representation of the LoFi data, $\mathbf{Y}$ in the latent space of the HiFi data, $\mathbf{X}$. As a result of this alignment, both manifolds are embedded in the same space which allowed an easy comparison and ability to be used in the multi-fidelity framework.

2.3. CoKriging

Once the shared latent representation is obtained, with POD coefficients matrices $\mathbf{Z}$ and $\mathbf{T}$ for the HiFi and LoFi data respectively, a regression model is needed to link the POD coefficients with the design space, i.e., a model of the form $\alpha^k = \alpha^k(\boldsymbol{d}; \boldsymbol{z}, \boldsymbol{t})$ for each POD coefficient k.

The Gaussian Process Regression (GPR) [33], also known as Kriging when a spatial covariance function or kernel is used, is a very popular machine learning technique used for predicting data points based on probabilistic models that define distributions over functions and estimate uncertainties. Kriging is a two-step process: training and prediction. The model finds the underlying patterns between input–output from the available data during training. After model training, for a given new input, the model provides a predictive distribution over possible function values at that input which includes a mean value and a measure of uncertainty.

Kriging is used to predict a single variable based on the values of that same variable at observed locations. However, in the multi-fidelity approach, it is required to use a regression model to predict one variable combining the values of two correlated variables at observed locations. For this reason, CoKriging [13,34,35], also known as collocated Kriging, is preferred for these applications. CoKriging is an extension of Kriging and it improves the prediction accuracy by taking advantage of the spatial relations between correlated variables. This method is particularly useful in situations where one variable is measured more accurately than the others, making it a good choice for multi-fidelity frameworks.

The CoKriging regression model adopted in this study follows the auto-regressive model of Kennedy and O'Hagan [34] which is based on the following assumption:

$$\mathrm{cov}\{\alpha_{\mathrm{HiFi}}(\boldsymbol{d}), \alpha_{\mathrm{LoFi}}(\boldsymbol{d}')|\alpha_{\mathrm{LoFi}}(\boldsymbol{d})\} = 0 \text{ for all } \boldsymbol{d} \neq \boldsymbol{d}', \tag{8}$$

where α_{HiFi} is the HiFi model, while α_{LoFi} is the LoFi model to form the training datasets. This Markov property implies that the HiFi data is correct and nothing more can be learned from the LoFi data if a HiFi value is available [36]. It is possible to explain the working principle of CoKriging as building two Kriging models in sequence: the first Kriging model is based on the LoFi data only, and the second model is constructed on the residuals of HiFi and LoFi data:

$$\alpha_{\mathrm{HiFi}}(\boldsymbol{d}) = \rho\,\alpha_{\mathrm{LoFi}}(\boldsymbol{d}) + \delta(\boldsymbol{d}), \tag{9}$$

where ρ is a scaling factor, and δ is the modeling of the difference between the LoFi and HiFi data. Both models require the specification of a kernel function (or covariance), typically a radial basis function (RBF), along with the calibration of its hyper-parameters, such as

the signal variance and the length-scale,[2] in addition to the calibration of the scaling factor ρ. Since the hyper-parameters values have a considerable effect on the fitting properties of the model – their choice is indeed a model selection problem – a common approach is to optimize them. To this end, the Bayesian approach proposed by Le Gratiet [37] was embraced and then implemented in openMDAO [38], which provides a closed-form expression for the scaling factor and a numerically tractable marginal likelihood of the hyper-parameters, whose maximization returned the best estimate for the parameters given the training data.

Once fitted, the CoKriging model returns a mean prediction function that estimates $\alpha_{\text{HiFi}}(\boldsymbol{d})$ and its uncertainty $\sigma_{\alpha_{\text{HiFi}}}(\boldsymbol{d})$, with the assumption that the uncertainty associated with the LoFi data α_{LoFi} and the residual δ are uncorrelated and can be defined as:

$$\sigma_{\alpha_{\text{HiFi}}}(\boldsymbol{d}) = \sqrt{\sigma^2_{\alpha_{\text{LoFi}}}(\boldsymbol{d}) + \sigma^2_{\delta}(\boldsymbol{d})}. \tag{10}$$

An example of CoKriging regression is shown in Fig. 1 using the toy datasets provided in [36]. The autoregressive feature is noticeable from the fact that the mean prediction goes through the HiFi samples and the uncertainty in those locations is zero.

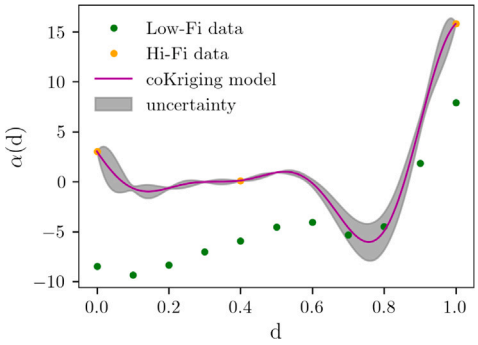


Fig. 1. Example of CoKriging regression using the toy datasets provided in [36].

2.4. MF-ROM prediction and error metrics

A general flow diagram of the multi-fidelity digital twin is provided in Fig. 2, where $\mathbf{X}$ is the HiFi data, $\mathbf{Y_L}$ is the linked and $\mathbf{Y_U}$ is the unlinked LoFi data. The POD modes $\mathbf{\Phi}$ and $\mathbf{\Psi}$ and the POD coefficients $\mathbf{Z}$ and $\mathbf{W}$ are obtained for HiFi and LoFi models,respectively. The dimensionality reduction is achieved by truncating the matrices according to a given retained total variance. Manifold alignment is used to transform LoFi POD coefficients $\mathbf{W}$ to $\mathbf{T}$ in order to achieve information transfer between two manifolds. The CoKriging regression model is trained with the POD coefficients and gave a mean prediction $\mu_{\alpha^k}(\boldsymbol{d})$ and uncertainty estimation $\sigma_{\alpha^k}(\boldsymbol{d})$ for the POD coefficients in an unseen design point $\boldsymbol{d}$. Finally, the predicted physical solution of this unseen design point $\boldsymbol{x}_{pred}(\boldsymbol{d})$ is obtained by multiplication of the obtained CoKriging coefficients with the eigenflames. The numerical implementation of the workflow was made in Python, using SciPy, Gpytorch, cvxpy, and openMDAO, and is available online [39].

In summary, the obtained MF-ROM is a functional relation of the kind:

$$\{\boldsymbol{x}_{pred}, \sigma(\boldsymbol{x}_{pred})\}(\boldsymbol{d}^*) = \mathcal{M}[\boldsymbol{X}, \boldsymbol{Y}](\boldsymbol{d}^*), \tag{11}$$

where the model $\mathcal{M}$ depends on the employed HiFi and LoFi training data $\mathbf{X}$ and $\mathbf{Y}$, as well as on the CoKriging hyper-parameters, and returns a predicted flowfield $\boldsymbol{x}_{pred}$ and its point-wise uncertainty $\sigma(\boldsymbol{x}_{pred})$ in an unseen design condition $\boldsymbol{d}^*$.

In this work, the ROM prediction accuracy is quantified using the Normalized Root Mean Square (NRMS) error, which is a common

[2] The training data are considered as noise-free data, thus the noise variance of the kernel function is set to zero.

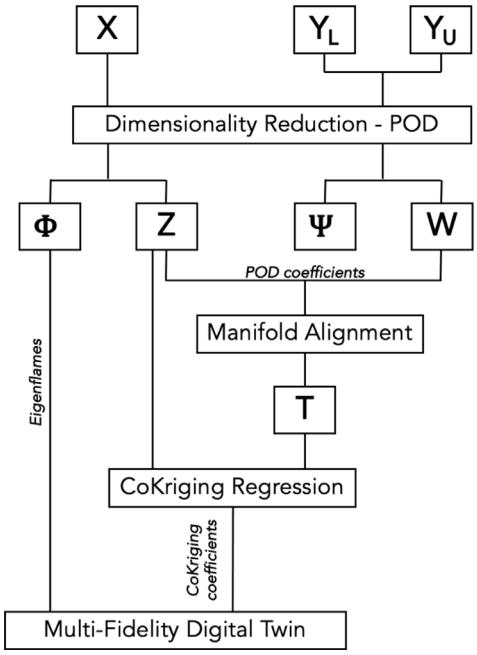


Fig. 2. Multi-fidelity reduced order model flow diagram.

metric to assess model performance. Specifically, NRMS error is used to evaluate the accuracy of the predictions compared to test points, namely CFD simulation results for selected test cases. The NRMS error is obtained by dividing the Root Mean Square (RMS) error by the data mean:

$$\mathcal{E}^{QoI}_{\text{NRMS}}(\boldsymbol{d}^{test}) = \frac{1}{\mu[\boldsymbol{x}(\boldsymbol{d}^{test})]} \sqrt{\frac{\sum_{i=1}^{n} \left(x_i(\boldsymbol{d}^{test}) - x_{i,pred}(\boldsymbol{d}^{test})\right)^2}{n}}, \tag{12}$$

where $x_i(\boldsymbol{d}^{test})$ and $x_{i,pred}(\boldsymbol{d}^{test})$ are the test simulation data and the ROM predicted data respectively for a given quantity of interest (QoI), e.g. temperature or a chemical species, and n is the number of cells of the flow field. More synthetic measures of error are also employed by averaging the NRMS error over a set of N_{test} test conditions:

$$\mathcal{E}^{QoIs}_{\text{NRMS}} = \frac{\sum_{t}^{N_{test}} \mathcal{E}^{QoIs}_{\text{NRMS}}(\boldsymbol{d}^t)}{N_{test}}, \tag{13}$$

and in turn over a set of N_{QoIs} QoIs:

$$E_{\text{NRMS}} = \frac{\sum_{q}^{N_{QoIs}} \mathcal{E}^{q}_{\text{NRMS}}}{N_{QoIs}}. \tag{14}$$

In addition to global accuracy measures, the prediction quality is assessed by employing topological and combustion-specific observables, such as the wall temperature, the OH radical peak position and value, and the flame length. In such cases, relative errors are used:

$$\epsilon_{\text{obs}} = \frac{\text{obs}(\boldsymbol{d}^{test}) - \text{obs}_{pred}(\boldsymbol{d}^{test})}{\text{obs}(\boldsymbol{d}^{test})}. \tag{15}$$

2.5. Incremental sampling

The amount of training data required to guarantee a given ROM's level of accuracy is typically problem-specific and determined by the non-linearity, the dimensionality, the noisy or smooth behavior of the underlying input/output relationship, and the training method [40]. It follows that a-priori sampling strategies which provide a distribution of training points without any knowledge of the underlying input/output relationship often end up being inefficient. Therefore, the ROM training strategies have moved from static to dynamically updating models, due

to the ability to improve their fitting capability by adaptive sampling or training [41]. The importance of adaptive sampling strategies lies in adding training points to the locations in the design space where they are the most impactful, keeping the number of model evaluations needed to build the surrogate as low as possible. Clearly, the impact of a new sample must be estimated from the knowledge gained during the process up to that point. Therefore, adaptive sampling strategies require a method to choose the coordinates of a new training point, which in turn requires a decision-making metric. This metric should target and quantify a potential/estimated improvement of the surrogate model prediction and/or a reduction of the associated uncertainty.

In the context of multi-fidelity models, the quest for a satisfying accuracy-to-cost balance involves the determination of the amount and distribution of the HiFi and LoFi samples. In fact, fewer HiFi simulations will reduce the prediction accuracy, while a larger number will increase the computational cost. Moreover, the choice of conditions in the design space for the HiFi subset of simulations has a crucial effect on the prediction accuracy of the resulting MF-ROM.

To this end, adaptive sampling strategies have been combined with multi-fidelity surrogate models, with the goal of deciding on a cost/effect balance basis whether an expensive HiFi sample is needed in place of a cheaper one. An augmented expected improvement for multi-fidelity Kriging was proposed by Huang et al. [42], by using multi-fidelity data to minimize the total evaluation cost by finding the overall optimum of the real system. Pellegrini et al. [18] used a sampling strategy that sequentially adds LoFi or both HiFi and LoFi samples depending on the prediction uncertainty and on a parameter based on their computational cost ratio. Cai et al. [43] used an algorithm to sample both HiFi and LoFi at every iteration based on the cross-validation error and a Voronoi partition. Liu et al. [14] applied a sampling method updating the dataset of the corresponding fidelity level that provides the maximum reduction of the MF uncertainty. Kimpton et al. [44] combined expected improvement with a repulsion function that ensures a filling design space, in a multi-level fashion.

The first incremental sampling method proposed in this work is rather a benchmark strategy that requires the availability of the full HiFi and LoFi datasets $\mathbf{X}$ and $\mathbf{Y}$, respectively. Starting from a baseline/minimal set of HiFi data and the full set of LoFi data, the choice of a new HiFi training point is driven by the calculated prediction accuracy improvement. Therefore, given the initial set of HiFi samples $D^0 \in \mathbf{X}$, the HiFi dataset is progressively augmented by adding the simulation data $\boldsymbol{x}'(\boldsymbol{d}')$ corresponding to the DoE coordinates $\boldsymbol{d}'$ such that, at the $i+1$-th increment, $\boldsymbol{x}'$ is:

$$\underset{\boldsymbol{x}' \in (\boldsymbol{X} - D^i)}{\operatorname{argmin}} \ \mathcal{E}(\boldsymbol{x}') := \{\boldsymbol{x}' | \forall \boldsymbol{x} : \mathcal{E}(\mathcal{M}[D^i + \boldsymbol{x}, \boldsymbol{Y}]) > \mathcal{E}(\mathcal{M}[D^i + \boldsymbol{x}', \boldsymbol{Y}])\}, \quad (16)$$

where $\mathcal{E}(\mathcal{M}[D^i + \boldsymbol{x}, \boldsymbol{Y}])$ is the NRMS error defined in Eq. (12) computed between the test vector and the data vector predicted by the model $\mathcal{M}$ obtained using the HiFi dataset D^i augmented with a new sample $\boldsymbol{x}$. Therefore, the chosen new HiFi sample is the one that maximizes the prediction accuracy improvement among the pool of available samples.

Clearly, at a given stage i, this strategy requires the generation of all possible models $\mathcal{M}$ built with the HiFi training dataset D^i plus each one of the remaining training data $\boldsymbol{x}$. Therefore, besides requiring the full HiFi dataset upfront, it encompasses the generation and evaluation of many models at each sampling increment, making it unfeasible in a practical context. This strategy, from now on is named as *error-based strategy*, will be used as a reference.

In a practical context, a different approach is needed since the full HiFi dataset is generally not available upfront because of its computational cost. At the same time, the full set of LoFi data is considered to be available, being it much cheaper, although this might not be the case in general. In this case, the HiFi sample coordinates of successive dataset increments should be identified using heuristic methods. The uncertainty of the CoKriging model might be a good indicator of missing information, i.e., of the need to add a sample in the proximity of the region where it is larger. In fact, the addition of a HiFi sample will improve the CoKriging discrepancy model where it has the largest uncertainty.

Therefore, a new sample $\boldsymbol{x}'(\boldsymbol{d}')$ is added corresponding to the DoE coordinates $\boldsymbol{d}'$ such that, at the $i+1$-th increment, it is:

$$\underset{\boldsymbol{d}_j}{\operatorname{argmax}} \left[\sum_{k}^{N_{modes}} \Sigma_k \, \sigma_{\alpha^k}(\boldsymbol{d}_j) \right] \qquad j = 1, \ldots, p - i, \quad (17)$$

where the overall prediction uncertainty in the DoE coordinates $\boldsymbol{d}_j$ is computed by summing the uncertainty associated with each POD coefficient k, weighted with the corresponding eigenvalue Σ_k, i.e., the POD coefficients' uncertainty are weighted with the importance of their mode. This strategy will be referred to as *uncertainty-based strategy*.

3. Datasets and workflow settings

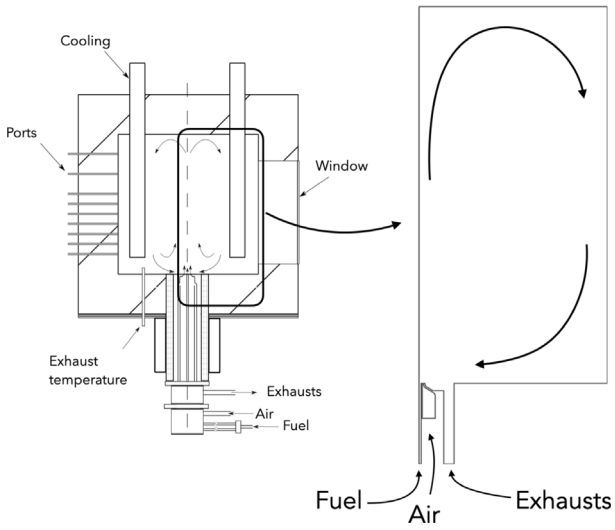


Fig. 3. Schematic of the 20 kW MILD combustion furnace (left), detail of the midsection with fuel/exhaust/EGR injectors, and recirculation pattern (right).

.

The HiFi and LoFi datasets in this work were assembled by 3D and 2D CFD simulations of the semi-industrial MILD combustion furnace with a nominal power of 20 kW, whose schematic is shown in Fig. 3. The experimental setup is flexible in terms of providing different injection profiles, air excess, fuel, air velocity, and internal load [45]. The numerical simulations were obtained using the commercial software ANSYS Fluent 19.1. Considering the symmetry of the setup, the numerical domain for the 3D simulations was chosen as a 45° angular sector of the full geometry. The validity of this assumption with respect to the full geometry was confirmed in a previous study by Ferrarotti et al. [45]. For the 2D simulations, leveraging the symmetry of the problem, an axisymmetric domain defined by the cross section of the cylindrical representation of the furnace was used. Ferrarotti et al. [45] conducted an extensive analysis to compare the performance of 3D and 2D simulation outcomes, revealing that the use of 3D geometry aids in capturing complex fluid dynamic flow patterns, resulting in more uniform species/temperature profiles. Additionally, solutions obtained with a 2D domain exhibited a consistent overestimation of temperature peaks by up to 10%. The setup for all the simulations used the standard k-ϵ turbulence model with Partially Stirred Reactor (PaSR) [46] model for turbulence-chemistry interactions. The selection of the chemical mechanism was based on the sensitivity study carried out by Aversano et al. [12], where the Kee mechanism (17 species and 58 reactions) was chosen as appropriate for this problem. The discrete ordinate (DO) model with the weighted-sum-of-gray-gases (WSGG) model was applied for the radiation modeling. As the introduction of hydrogen leads to notable changes in simulation outcomes, accounting for diffusion

effects, a non-unity Lewis number assumption was applied to both HiFi and LoFi datasets. This was addressed by accounting for binary diffusion coefficients calculated following the kinetic theory and a modification of the Chapman–Enskog formula. An effective diffusion coefficient of the species in the mixture was obtained by applying a mixing rule. Following the mesh sensitivity performed by Ferrarotti et al. [45], numerical domains of 216k and 29k cells for HiFi and LoFi simulations were used, respectively.

The design space, shown in Fig. 4, is composed of fuel composition, equivalence ratio, and injector diameter. These parameters range between 0%–100% for H_2 mole fraction, 0.7–1 for equivalence ratio ϕ, and 16 mm-20 mm-25 mm for the air injector diameter. The DoE was performed by using the same Latin Hypercube Sampling (LHS) employed by Aversano et al. [12], which features 45 points. A total of 45 simulations were thus available for both HiFi and LoFi models with the following features of interest: temperature, major species (CH_4, H_2, O_2, H_2O, OH), and minor species (CO and OH).

The training dataset, namely the 45 companion HiFi/LoFi simulations, needed to be split into train/test sets. Instead of making a-priori choices of the test simulations, a k-fold cross-validation is performed, given the limited amount of data samples available, which would otherwise lead to biases in the model performance evaluation. More specifically, a 30-fold cross-validation with random test sets containing 4 simulations was performed.

The total variance retained in the POD modes after dimensionality reduction was set to 99.9%, while the CoKriging model was built with radial basis functions as covariance, and a constant regression model for the scaling factor, the latter being estimated with a Bayesian optimization together with the covariance hyper-parameters.

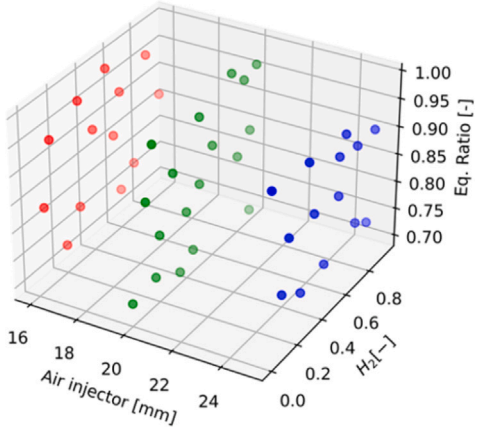


Fig. 4. Design space composed of fuel composition, equivalence ratio, and injector diameter. Points are colored by the injector diameter value.

4. Results

This section shows the performance of MF-ROMs built employing the two incremental sampling strategies described in Section 2.5, first the error-based and then the uncertainty-based. The predictive capabilities of the MF-ROMs against a reference full-high-fidelity model is compared, as described here below. The full-high-fidelity model delivers the most accurate predictions for the given DoE, and such comparison serves to quantify the biases (inaccuracies) introduced by the low-fidelity information in the MF-ROMs.

4.1. Reference high-fidelity digital twin

A high-fidelity digital twin of the semi-industrial MILD combustion chamber was already developed by Aversano et al. [12] using 3D simulations only, and then employed by Donato et al. [47] to assimilate experimental data. The POD method was used to decompose the physics-invariant information (POD modes) from the system-dependent information (POD coefficients), while the behavior of the POD coefficients over the design space was modeled with a Kriging regression. This strategy can be seen as the single-fidelity special case of the methodology presented in this paper, where the HiFi information contained in the POD coefficients $\mathbf{Z}$ always overrides the low-fidelity information contained in the dataset $\mathbf{Y}$, which therefore is not used.

The average NRMS error for the features of interest, computed as per Eq. (13) in the prediction of the test simulations is provided in Fig. 5. More specifically, the error statistics are reported (mean, standard deviation, and a kernel density estimation in the form of a violin plot) over the 30 random folds. The error for temperature and the main chemical species mass fractions and pollutants, except for CO, OH, CH_4, and H_2, are predicted around and below 10%. The higher errors for CO, OH, CH_4, and H_2 are expected since these species are more challenging to predict due to their much more localized distribution [12]. The 30-fold-mean NRMS error averaged over the QoIs is $\mu_{30}[E_{\text{NRMS}}] = 0.8$, where the contribution of the less accurate species is clearly dominant.

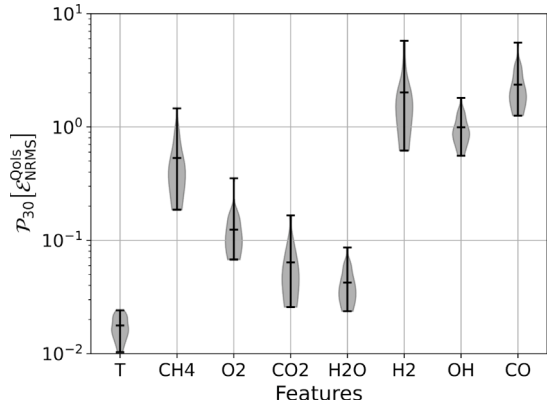


Fig. 5. 30-fold statistics (mean, standard deviation and kernel density estimation) of the QoIs NRMS errors $\mathcal{E}_{\text{NRMS}}^{QoIs}$ exhibited by the reference high-fidelity ROM built by using the full set of 41 HiFi simulations.

.

Fig. 6 shows the temperature and OH field comparison between the CFD and the HiFi ROM for a selected test point, which is an unseen design condition for the ROM. The field plots show that the HiFi ROM is able to predict accurately the true distribution of the features of interest without any noteworthy difference, except for a slight misplacement of the flame edge in the injector post-tip.

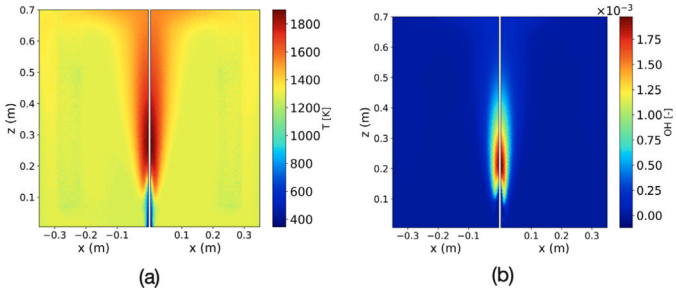


Fig. 6. Slice of the 3D flow field compared between the true field from the CFD simulation *(left half)* and the digital twin prediction *(right half)* in a selected test point for temperature (a) and OH mass fraction (b).

.

The full-HiFi error figures shown in Figs. 5 and 6 will represent the lower error bounds, i.e., the best achievable performance of any MF-ROMs, and will be reported in black in all the following figures.

4.2. Multi-fidelity digital twin

In this section two MF-ROMs were build by employing the incremental sampling strategies outlined in Section 2.5.

Both strategies required an initial HiFi training set that acts as a seed for the incremental sampling. To this end, 6 design conditions placed close to the corners of the parameter space were picked, leaving 39 simulations for the train/test sets, performing again 30-fold cross-validations with random test sets containing 4 simulations each. This choice was justified by the fact that Kriging-type methods are known to be inaccurate when used for extrapolation [19].

The performance of the MF-ROM utilizing the initial 6 HiFi and 41 LoFi simulations (referred to as Base MF-ROM), is compared with a SF-ROM constructed using the same 6 HiFi simulations and no LoFi simulations (referred to as Base SF-ROM), in order to show the effect of the LoFi data on predictions. This comparison is depicted in Fig. 7 for each feature of interest within the context of a 30-fold cross-validation scenario. The full HiFi ROM performance is reported as well for reference. As illustrated in the figure, the MF-ROM generally provides more accurate predictions with reduced bias across most features compared to the SF-ROM. Notably, for species characterized by localized and challenging-to-predict behaviors such as H_2, OH, and CO, the average performance of MF-ROM improves that of the SF-ROM. This outcome underlines the utility of employing MF-ROMs, as they demonstrate the capability to achieve better prediction accuracy compared to their SF-ROM counterparts constructed using the same amount of HiFi data.

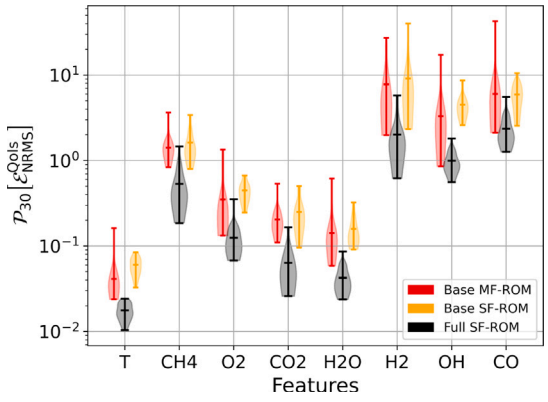


Fig. 7. 30-fold statistics (mean, standard deviation and kernel density estimation) of the QoI-averaged NRMS error E_{NRMS} for Base MF-ROM and SF-ROM compared to the SF-ROM with 41 HiFi simulations (Full SF-ROM).

4.2.1. Error-based incremental sampling method

Here, MF-ROMs were incrementally built employing the criterion in Eq. (16), which progressively adds the HiFi sample that provides the largest accuracy improvement among the pool of remaining HiFi simulations. The performance of the obtained models were evaluated by comparing their predictions against the test data in the 4 randomized test points, using 30-fold cross-validation. Fig. 8 shows the probability distribution of the NRMS error E_{NRMS} defined in Eq. (14) over the 30 folds at each increment. A decreasing trend is observed for the 30-fold-mean error and 30-fold-standard-deviation with increasing HiFi samples, which reveals that (i) the addition of HiFi simulations improves the prediction accuracy of the MF-ROM and (ii) the bias induced by the test set choice (represented by the standard deviation, or PDF spread) is less and less important because the left-out simulations are less impactful on prediction accuracy. Moreover, the PDFs are skewed towards lower errors, meaning that only a few random test case combinations have less accurate predictions, typically when test cases are clusterized towards the domain boundary. Further, it is observed that the MF-ROM accuracy largest improvement happens in the first ~ 20 increments, while it plateaus after ~30 HiFi samples are reached. When the full set of 41 HiFi samples is added, the accuracy of the MF-ROM recovers the performance of the high-fidelity ROM shown in Section 4.1, Fig. 5, as expected.

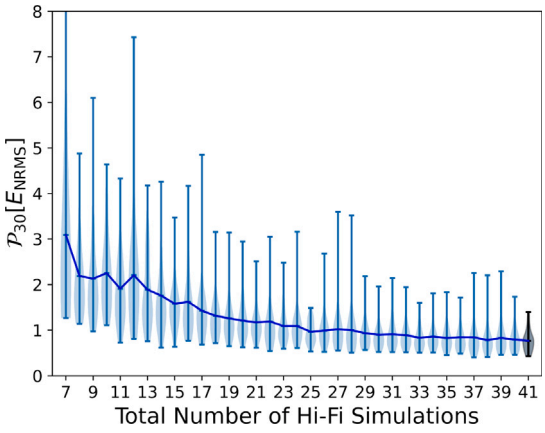


Fig. 8. 30-fold statistics (mean, standard deviation and kernel density estimation) of the QoI-averaged NRMS error E_{NRMS} with increasing HiFi samples employing the error-based strategy.

After a general understanding of the behavior of this sampling strategy obtained by looking into the QoIs-averaged error, it is beneficial to examine the error behavior feature by feature. Fig. 9 shows the QoIs errors (Eq. (13)) at 4 selected stages of the incremental sampling, namely 9, 18, 27, and 41 (full) HiFi samples. These error figures show that for every feature, the addition of HiFi simulations increases the performance of the MF-ROM, as expected. Very disparate prediction performances across the QoIs are observed, coherently with the full-high-fidelity model (black symbols, also shown in Fig. 5), with excellent accuracy for temperature and major products, and decreasing quality for reactants and radical species. It is noteworthy to remark that the MF-ROM with 9 HiFi samples already exhibits acceptable error figures compared to the full-high-fidelity ROM.

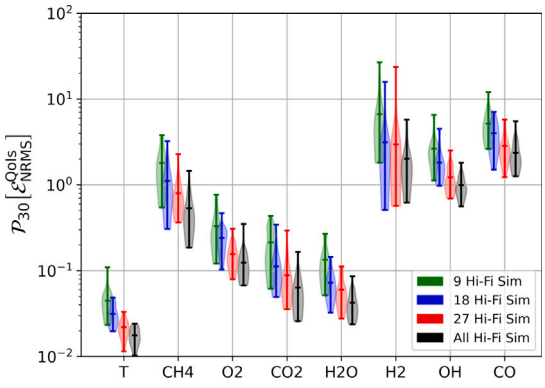
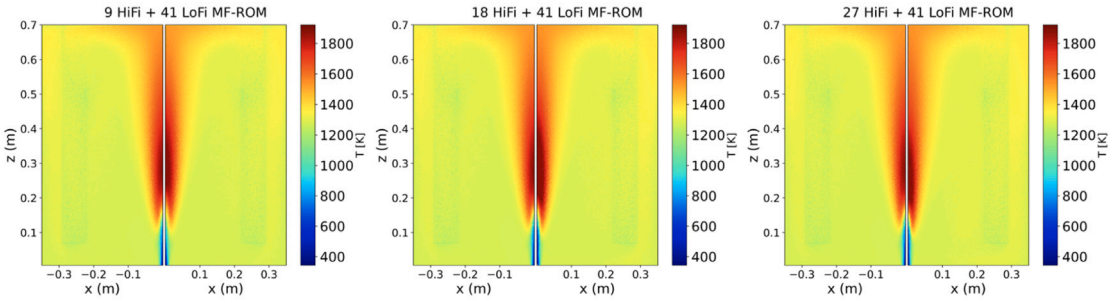


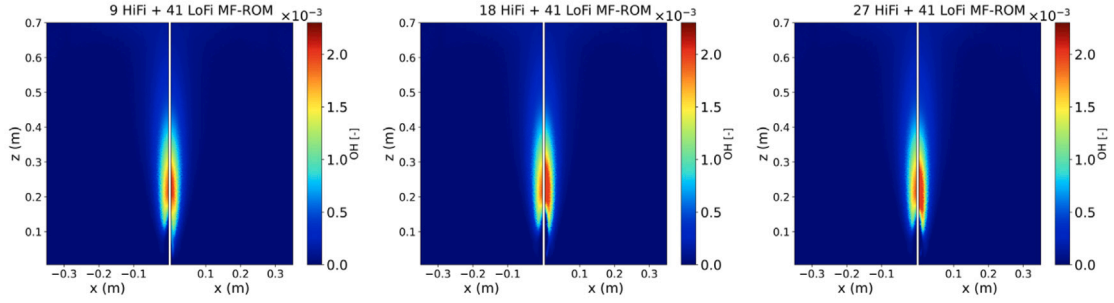
Fig. 9. 30-fold-statistics of the QoIs NRMS errors $\mathcal{E}^{QoIs}_{NRMS}$ exhibited by MF-ROMs built by using 9, 18, 27 and 41 HiFi simulations employing the error-based sampling strategy.

The temperature and OH predictions for the same test case are compared in Fig. 10 as the number of HiFi simulations increases, particularly in the ranges of 9-18-27 HiFi simulations, for the error-based method. The temperature exhibits an overestimation near the symmetry axis for small and medium number of HiFi samples. A similar overestimation trend is observed in OH predictions, accompanied by a less accurate depiction of the flame shape. Both predictions improve with increasing number of HiFi samples.

Lastly, topological observables were analyzed, namely the wall temperature (i.e., the average temperature predicted on the side wall $x = 0.35$ m), the peak OH value/position and the flame length (measured as the maximum height reached by the 20% OH iso-line). 30-fold-statistics on the relative errors ε_{obs} on such quantities (Eq. (15)) are shown in Fig. 11 at the same 4 selected stages of the incremental



(a) Temperature



(b) OH mass fraction

Fig. 10. Comparison between the MF-ROMs for the error-based method built with 9 (left), 18 (center), 27 (right) HiFi simulations and 41 LoFi, temperature (top row), and OH mass fraction (bottom row). Each plot compares a slice of the predicted 3D field (right half) against the same (mirrored) slice of the test simulation (left half).

sampling (9, 18, 27, full). The convergent trend towards the reference full-HiFi performance is here found again, except for the 30-fold-mean peak OH position whose error does not change appreciably. The error values confirm that the OH prediction in general is not very accurate, regardless of the number of low-fidelity data injected in the model, while the wall temperature (which is more relevant in a practical context) is very well predicted, even with 9 HiFi samples only.

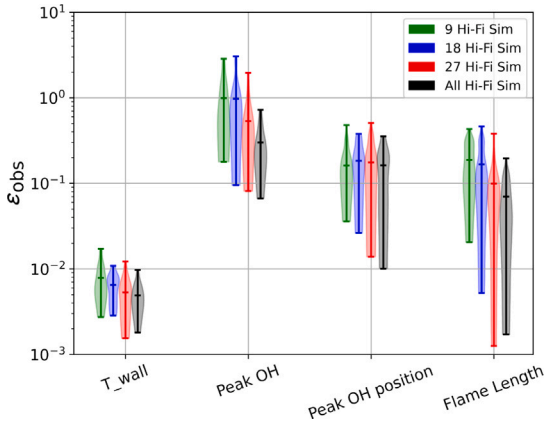


Fig. 11. 30-fold-statistics of the relative errors ε_{obs} on wall temperature, peak OH value, peak OH position, and flame length exhibited by MF-ROMs built by using 9, 18, 27, and 41 HiFi simulations employing the error-based sampling strategy.

4.2.2. Uncertainty-based incremental sampling method

In this section, MF-ROMs were incrementally built employing the criterion in Eq. (17), which adds the HiFi sample corresponding to the DoE location where the MF-ROM uncertainty is larger. Again, the performance of the obtained models were evaluated by comparing their predictions' average NRMSE distribution against the test data in the 4 randomized test points, using 30-fold cross-validation. Fig. 12 shows the 30-fold-statistics of the NRMS error E_{NRMS} defined in Eq. (14) at each increment. The trend of the 30-fold-mean of the error with an increasing number of HiFi simulations is similar to the error-based method. However, the number of increments needed to reach the plateau is lower. A sharper decrease in the error indicates a faster and steadier convergence to the full-high-fidelity NRMS error.

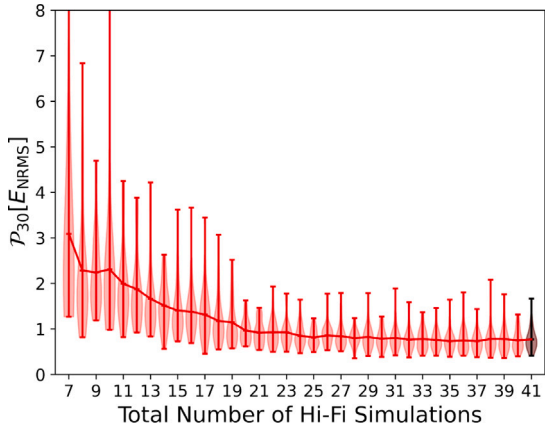


Fig. 12. 30-fold statistics (mean, standard deviation and kernel density estimation) of the NRMS error E_{NRMS} with increasing HiFi samples employing the uncertainty-based strategy.

.

The effect of additional HiFi simulations on the prediction accuracy of the QoIs is shown in Fig. 13, where only the mean over the 30-folds is shown for the sake of clarity. All the QoIs errors exhibit a

smooth decreasing trend with increasing HiFi samples, maintaining the disparities across the QoIs already observed in the previous sections.

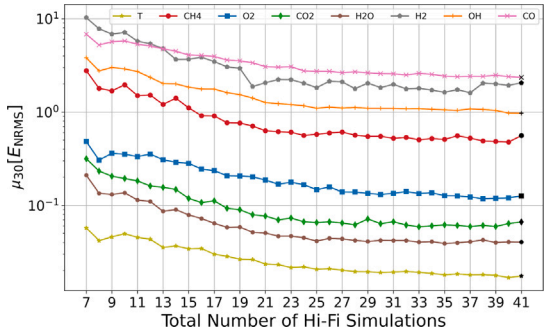


Fig. 13. 30-fold-mean of the QoIs NRMS errors $\mathcal{E}_{\text{NRMS}}^{QoIs}$ exhibited by MF-ROMs built employing the uncertainty-based sampling strategy.

.

Again, a more compact visualization of predictive performance is shown in Fig. 14, which depicts the QoIs errors (Eq. (13)) at 4 selected stages of the incremental sampling, namely 9, 18, 27, and 41 (full) HiFi samples. These error figures show that for every feature, the addition of HiFi simulations increases the performance of the MF-ROM. Moreover, the accuracy of the MF-ROM with 27 HiFi simulations and the high-fidelity ROM trained with the full set of HiFi simulations show very close accuracy metrics. Fig. 15 shows slices of the predicted flow fields of temperature and OH in the 9/18/27 HiFi stages of the incremental sampling, compared against the test simulation in a selected test point. The temperature field progressively improves with additional HiFi simulations, both in the flame location and in the peak value. The OH field topology improves as well, reaching a satisfying flame location with 27 HiFi simulations, with the peak OH value being over-predicted at 9 HiFi stage and under-predicted at 18 HiFi, finally reaching a more accurate value at 27.

This behavior is confirmed by Fig. 16, which shows the 30-fold-statistics of the topological relative errors at the same increment stages.

4.2.3. Performance comparison of incremental sampling methods

In this study, the performance of error-based and uncertainty-based sampling methods are compared, alongside an adaptive sampling methodology introduced by Guénot et al. [48]. The comparison is based on the NRMS errors exhibited by the MF-ROMs resulting from these three methods, as illustrated in Fig. 17. Guénot et al.'s adaptive sampling method was utilized with non-intrusive POD-based surrogate models. The findings reveal that all three methods exhibit similar

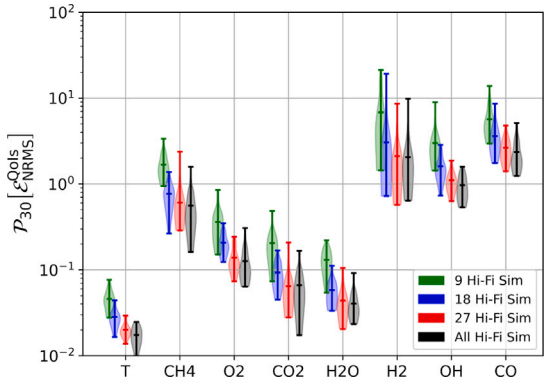


Fig. 14. 30-fold-statistics of the QoIs NRMS errors $\mathcal{E}_{\text{NRMS}}^{QoIs}$ exhibited by MF-ROMs built by using 9, 18, 27 and 41 HiFi simulations employing the uncertainty-based sampling strategy.

.

trends, with Guénot et al.'s method closely resembling the error-based approach presented in this study. Additionally, it is observed that the uncertainty-based strategy demonstrates faster convergence in prediction accuracy with increasing HiFi samples compared to the other two methods. This result may appear counter-intuitive at a first glance, because one would expect the error-based strategy to be the best possible strategy if the goal is to minimize the prediction error of the MF-ROM. However, it is important to consider that this sampling strategy can be classified as a *greedy* algorithm, meaning that at each increment it selects the *locally* optimal solution. In other words, this methodology does not guarantee to uncover the globally optimal set of training samples at each iteration. This would require a combinatorial search over the set of possible training samples, which becomes computationally intractable when the number of possible training samples grows larger than a few samples.

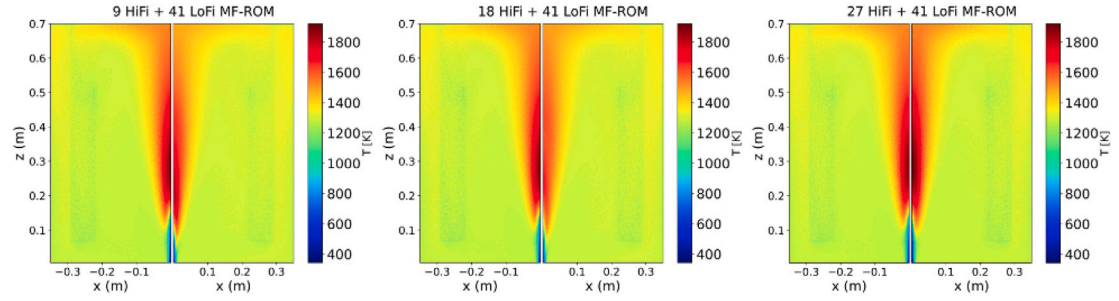
The uncertainty-based strategy can also be classified as a greedy algorithm, and thus it incurs the same local optimization problem. However, given that the uncertainty exhibited by the MF-ROM in a specific location is informed by the entire set of LoFi training samples as well, it is believed that the uncertainty-based method approximates a better global optimization method. By choosing the uncertainty-minimizing samples, this method selects HiFi samples that efficiently explore the design space based on the combined HiFi and LoFi information, and in the long run this results in more accurate models than the ones produced by the error-based strategy.

Such a result is remarkable since it demonstrates that the uncertainty-based strategy is a satisfying incremental sampling method for the problem under investigation, as well as being the only feasible one in a practical context. Moreover, it is demonstrated that an MF-ROM built with such a strategy employing ~20 HiFi simulations instead of 41, plus 41 LoFi simulations, delivers a model whose accuracy is comparable to the full-high-fidelity model while reducing by approximately one-half the number of degrees of freedom needed for the training.

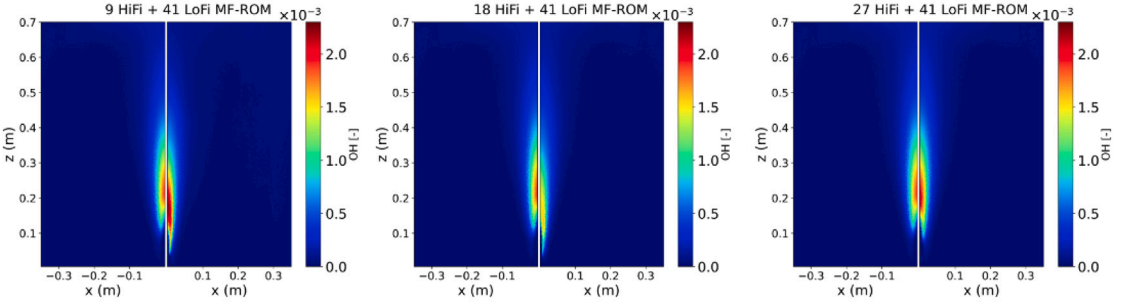
5. Summary and conclusions

In this work, an MF-ROM for a combustion furnace which operates on MILD conditions was developed and the effects of sampling strategies on the prediction accuracy of this model were investigated. This framework was based on the combination of the POD, the Procrustes manifold alignment, and the CoKriging regression model. The MF-ROM was used to predict the three-dimensional flowfields of temperature and chemical species for varying design parameters, namely air injector diameter, H_2 mole fraction, and the equivalence ratio. The training datasets included 45 two-dimensional (LoFi) and 45 three-dimensional (HiFi) CFD simulations of the same furnace for the same design conditions. The framework fused the LoFi and HiFi data with the goal of reducing the number of HiFi simulations required for a desired model accuracy. Preliminary studies indicated that the number and the design conditions of the chosen HiFi data have a crucial impact on the prediction accuracy. As a result, two incremental sampling strategies were implemented and their effectiveness was examined. The error-based strategy added the HiFi sample that provides the largest accuracy improvement among the pool of remaining HiFi simulations, while the uncertainty-based one added the HiFi sample corresponding to the DoE location where the MF-ROM uncertainty was larger. The latter is more suitable as well in a practical context since it does not require the availability of the HiFi dataset upfront.

The uncertainty-based method reached the full-high-fidelity model accuracy on the quantities of interest with ~20 HiFi and 41 LoFi simulations, instead of 41 HiFi simulations, therefore exhibiting a saving of approximately half the number of degrees of freedom needed for the training. If accuracy requirements on minor species were relaxed, the MF-ROM with ~9 HiFi and 41 LoFi simulations demonstrated good performance, leading to 65% saving.



(a) Temperature



(b) OH mass fraction

Fig. 15. Comparison between the MF-ROMs for the uncertainty-method built with 9 (left), 18 (center), 27 (right) HiFi simulations and 41 LoFi, temperature (top row), and OH mass fraction (bottom row). Each plot compares a slice of the predicted 3D field (right half) against the same (mirrored) slice of the test simulation (left half).

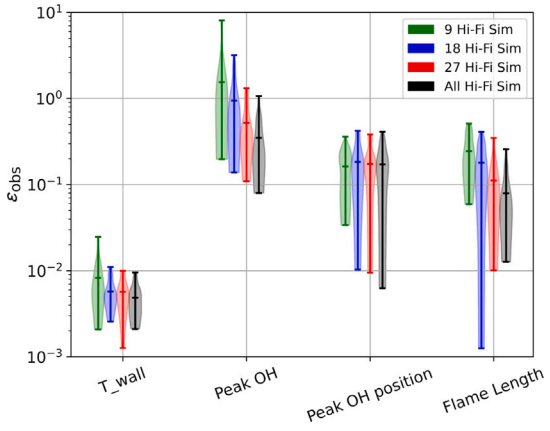


Fig. 16. 30-fold-statistics of the relative errors ε_{obs} on wall temperature, peak OH value, peak OH position, and flame length exhibited by MF-ROMs built by using 9, 18, 27, and 41 HiFi simulations employing the uncertainty-based sampling strategy.

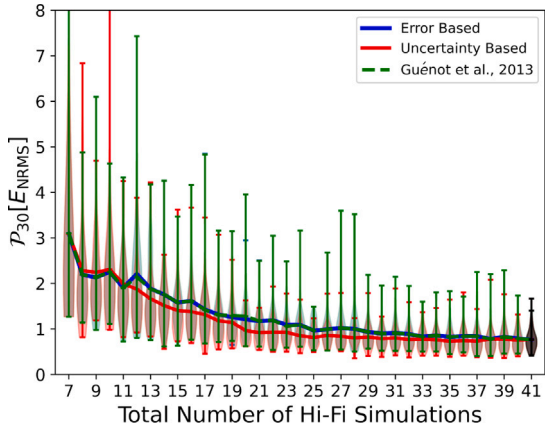


Fig. 17. Comparison between error-based and uncertainty-based incremental sampling with Guénot et al. [48]'s adaptive sampling strategies on the 30-fold-statistics of the NRMS error E_{NRMS}.

Future developments will seek further performance improvements in the MF-ROM accuracy, particularly in minor species, by calibrating the CoKriging hyper-parameters which have a strong impact on the model predictions. Moreover, a scenario in which the DoE exploration is performed on the fly during the incremental sampling will be addressed, instead of being constrained to a fixed LHS. Lastly, additional fidelity levels will be explored using lower fidelity simulationg (e.g. chemical reactor networks) and more detailed chemical kinetic mechanisms, to improve the predictions on minor species.

CRediT authorship contribution statement

A. Özden: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **A. Procacci:** Conceptualization, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **R. Malpica Galassi:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. **F. Contino:** Conceptualization, Methodology, Supervision, Writing – review & editing. **A. Parente:** Conceptualization, Formal

analysis, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

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

Data will be made available on request.

Acknowledgments

A. Özden acknowledges the financial support of the Fonds de la Recherche Scientifique de Belgique (F.R.S.-F.N.R.S) project METRIC, Belgium, under grant No. T.0175.21. A. Procacci acknowledges the financial support of the Fonds National de la Recherche Scientifique (F.R.S-F.N.R.S.), Belgium through an FRIA fellowship. R.Malpica-Galassi acknowledges the financial support from the Italian Ministry of University and Research, Italy in the framework of the project National Center for "High Performance Computing, Big Data, and Quantum Computing" funded by European Union – NextGenerationEU. This project has also received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No. 714605.

References

[1] UN, The sustainable development goals report, 2022, https://unstats.un.org/sdgs/report/2022.
[2] A. Cavaliere, M. de Joannon, Mild combustion, Prog. Energy Combust. Sci. 30 (4) (2004) 329–366, http://dx.doi.org/10.1016/j.pecs.2004.02.003.
[3] S. Kruse, B. Kerschgens, L. Berger, E. Varea, H. Pitsch, Experimental and numerical study of MILD combustion for gas turbine applications, Appl. Energy 148 (2015) 456–465, http://dx.doi.org/10.1016/j.apenergy.2015.03.054.
[4] N. Swaminathan, Physical insights on MILD combustion from DNS, Front. Mech. Eng. 5 (2019) http://dx.doi.org/10.3389/fmech.2019.00059.
[5] Z. Li, R.M. Galassi, P.P. Ciottoli, A. Parente, M. Valorani, Characterization of jet-in-hot-coflow flames using tangential stretching rate, Combust. Flame 208 (2019) 281–298, http://dx.doi.org/10.1016/j.combustflame.2019.06.023.
[6] D. Mira, E.J. Pérez-Sánchez, R. Borrell, G. Houzeaux, HPC-enabling technologies for high-fidelity combustion simulations, Proc. Combust. Inst. 39 (4) (2023) 5091–5125, http://dx.doi.org/10.1016/j.proci.2022.07.222.
[7] P. Benner, S. Gugercin, K. Willcox, A survey of projection-based model reduction methods for parametric dynamical systems, SIAM Rev. 57 (4) (2015) 483–531, http://dx.doi.org/10.1137/130932715.
[8] G. Berkooz, P. Holmes, J.L. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, Annu. Rev. Fluid Mech. 25 (1) (1993) 539–575, http://dx.doi.org/10.1146/annurev.fl.25.010193.002543.
[9] D.J. Lucia, P.S. Beran, W.A. Silva, Reduced-order modeling: new approaches for computational physics, Prog. Aerosp. Sci. 40 (1) (2004) 51–117, http://dx.doi.org/10.1016/j.paerosci.2003.12.001.
[10] C. Perron, D. Rajaram, D. Mavris, Development of a multi-fidelity reduced-order model based on manifold alignment, in: AIAA Aviation 2020 Forum, 2020, p. 3124, http://dx.doi.org/10.2514/6.2020-3124.
[11] P.G. Constantine, E. Dow, Q. Wang, Active subspace methods in theory and practice: Applications to kriging surfaces, SIAM J. Sci. Comput. 36 (4) (2014) A1500–A1524, http://dx.doi.org/10.1137/130916138.
[12] G. Aversano, M. Ferrarotti, A. Parente, Digital twin of a combustion furnace operating in flameless conditions: reduced-order model development from CFD simulations, Proc. Combust. Inst. 38 (4) (2021) 5373–5381, http://dx.doi.org/10.1016/j.proci.2020.06.045.
[13] P. Perdikaris, D. Venturi, J.O. Royset, G.E. Karniadakis, Multi-fidelity modelling via recursive co-kriging and Gaussian–Markov random fields, Proc. R. Soc. A: Math. Phys. Eng. Sci. 471 (2179) (2015) 20150018, http://dx.doi.org/10.1098/rspa.2015.0018.
[14] J. Li, W. Zhang, The performance of proper orthogonal decomposition in discontinuous flows, Theor. Appl. Mech. Lett. 6 (5) (2016) 236–243, http://dx.doi.org/10.1016/j.taml.2016.08.008.
[15] Y. Kuya, K. Takeda, X. Zhang, A.I. Forrester, Multifidelity surrogate modeling of experimental and computational aerodynamic data sets, AIAA J. 49 (2) (2011) 289–298, http://dx.doi.org/10.2514/1.J050384.
[16] L.W.-T. Ng, M. Eldred, Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation, in: 53rd Aiaa/Asme/Asce/Ahs/Asc Structures, Structural Dynamics and Materials Conference 20th Aiaa/Asme/Ahs Adaptive Structures Conference 14th Aiaa, 2012, p. 1852, http://dx.doi.org/10.2514/6.2012-1852.
[17] J. de Baar, S. Roberts, R. Dwight, B. Mallol, Uncertainty quantification for a sailing yacht hull, using multi-fidelity kriging, Comput. & Fluids 123 (2015) 185–201, http://dx.doi.org/10.1016/j.compfluid.2015.10.004.
[18] R. Pellegrini, U. Iemma, C. Leotardi, E.F. Campana, M. Diez, Multi-fidelity adaptive global metamodel of expensive computer simulations, in: 2016 IEEE Congress on Evolutionary Computation, CEC, 2016, pp. 4444–4451, http://dx.doi.org/10.1109/CEC.2016.7744355.
[19] Z.-H. Han, Zimmermann, S. Görtz, Alternative cokriging model for variable-fidelity surrogate modeling, AIAA J. 50 (5) (2012) 1205–1210, http://dx.doi.org/10.2514/1.J051243.
[20] G. Lamberti, C. Gorlé, A multi-fidelity machine learning framework to predict wind loads on buildings, J. Wind Eng. Ind. Aerodyn. 214 (2021) 104647, http://dx.doi.org/10.1016/j.jweia.2021.104647.
[21] B. Peherstorfer, K. Willcox, M. Gunzburger, Survey of multifidelity methods in uncertainty propagation, inference, and optimization, SIAM Rev. 60 (3) (2018) 550–591, http://dx.doi.org/10.1137/16M1082469.
[22] M. Gastpar, P.L. Dragotti, M. Vetterli, The distributed karhunen–loève transform, IEEE Trans. Inform. Theory 52 (12) (2006) 5177–5196, http://dx.doi.org/10.1109/TIT.2006.885449.
[23] A. Parente, J.C. Sutherland, Principal component analysis of turbulent combustion data: Data pre-processing and manifold sensitivity, Combust. Flame 160 (2) (2013) 340–350, http://dx.doi.org/10.1016/j.combustflame.2012.09.016.
[24] I.T. Jolliffe, Principal component analysis for special types of data, Principal Component Analysis, Springer New York, New York, NY, 2002, pp. 338–372, http://dx.doi.org/10.1007/0-387-22440-8_13.
[25] R. Pinnau, Model reduction via proper orthogonal decomposition, in: W.H.A. Schilders, H.A. van der Vorst, J. Rommes (Eds.), Model Order Reduction: Theory, Research Aspects and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 95–109, http://dx.doi.org/10.1007/978-3-540-78841-6_5.
[26] G. Aversano, A. Bellemans, Z. Li, A. Coussement, O. Gicquel, A. Parente, Application of reduced-order models based on PCA & kriging for the development of digital twins of reacting flow applications, Comput. Chem. Eng. 121 (2019) 422–441, http://dx.doi.org/10.1016/j.compchemeng.2018.09.022.
[27] C. Wang, S. Mahadevan, A general framework for manifold alignment, in: 2009 AAAI Fall Symposium Series, 2009.
[28] C. Wang, S. Mahadevan, Manifold alignment using procrustes analysis, in: Proceedings of the 25th International Conference on Machine Learning, ICML '08, Association for Computing Machinery, New York, NY, USA, 2008, pp. 1120–1127, http://dx.doi.org/10.1145/1390156.1390297.
[29] J. Ham, D. Lee, L. Saul, Semisupervised alignment of manifolds, in: International Workshop on Artificial Intelligence and Statistics, PMLR, 2005, pp. 120–127.
[30] R. Guerrero, C. Ledig, D. Rueckert, Manifold alignment and transfer learning for classification of alzheimer's disease, in: G. Wu, D. Zhang, L. Zhou (Eds.), Machine Learning in Medical Imaging, Springer International Publishing, Cham, 2014, pp. 77–84.
[31] F. Diaz, D. Metzler, Pseudo-aligned multilingual corpora., in: IJCAI, 2007, pp. 2727–2732.
[32] B. Luo, Feature matching with procrustes alignment and graph editing, in: IET Conference Proceedings, Institution of Engineering and Technology, 1999, pp. 72–76, (4).
[33] C.K. Williams, C.E. Rasmussen, Gaussian Processes for Machine Learning, vol. 2, MIT press Cambridge, MA, 2006.
[34] M. Kennedy, A. O'Hagan, Predicting the output from a complex computer code when fast approximations are available, Biometrika 87 (1) (2000) 1–13, http://dx.doi.org/10.1093/biomet/87.1.1.
[35] L.L. Gratiet, J. Garnier, Recursive co-kriging model for design of computer experiments with multiple levels of fidelity, Int. J. Uncertain. Quantif. 4 (5) (2014) 365–386.
[36] A.I. Forrester, A. Sóbester, A.J. Keane, Multi-fidelity optimization via surrogate modelling, Proc. R. Soc. A: Math. Phys. Eng. Sci. 463 (2088) (2007) 3251–3269, http://dx.doi.org/10.1098/rspa.2007.1900.
[37] L. Le Gratiet, Bayesian analysis of hierarchical multifidelity codes, SIAM/ASA J. Uncertain. Quantif. 1 (1) (2013) 244–269, http://dx.doi.org/10.1137/120884122.
[38] J.S. Gray, J.T. Hwang, J.R.R.A. Martins, K.T. Moore, B.A. Naylor, Openmdao: An open-source framework for multidisciplinary design, analysis, and optimization, Struct. Multidiscip. Optim. 59 (2019) 1075–1104, http://dx.doi.org/10.1007/s00158-019-02211-z.
[39] Openmeasure, python libraries for soft sensing applications, 2023, Accessed: 2023-10-07, https://github.com/albertoprocacci/OpenMEASURE.
[40] H. Liu, Y.-S. Ong, J. Cai, A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design, Struct. Multidiscip. Optim. 57 (2018) 393–416, http://dx.doi.org/10.1007/s00158-017-1739-8.

[41] A. Serani, R. Pellegrini, J. Wackers, C.-E. Jeanson, P. Queutey, M. Visonneau, M. Diez, Adaptive multi-fidelity sampling for CFD-based optimisation via radial basis function metamodels, Int. J. Comput. Fluid Dyn. 33 (6–7) (2019) 237–255, http://dx.doi.org/10.1080/10618562.2019.1683164.

[42] D. Huang, T.T. Allen, W.I. Notz, R.A. Miller, Sequential kriging optimization using multiple-fidelity evaluations, Struct. Multidiscip. Optim. 32 (2006) 369–382, http://dx.doi.org/10.1007/s00158-005-0587-0.

[43] X. Cai, H. Qiu, L. Gao, L. Wei, X. Shao, Adaptive radial-basis-function-based multifidelity metamodeling for expensive black-box problems, AIAA J. 55 (7) (2017) 2424–2436, http://dx.doi.org/10.2514/1.J055649.

[44] L. Kimpton, J. Salter, T. Dodwell, H. Mohammadi, P. Challenor, Cross-validation based adaptive sampling for multi-level Gaussian process models, 2023, arXiv:2307.09095.

[45] M. Ferrarotti, M. Fürst, E. Cresci, W. de Paepe, A. Parente, Key modeling aspects in the simulation of a quasi-industrial 20 kw moderate or intense low-oxygen dilution combustion chamber, Energy Fuels 32 (10) (2018) 10228–10241, http://dx.doi.org/10.1021/acs.energyfuels.8b01064.

[46] J. Chomiak, Combustion a study in theory, fact and application, 1990.

[47] L. Donato, C. Galletti, A. Parente, Self-updating digital twin of a hydrogen-powered furnace using data assimilation, Appl. Therm. Eng. 236 (2024) 121431, http://dx.doi.org/10.1016/j.applthermaleng.2023.121431.

[48] M. Guénot, I. Lepot, C. Sainvitu, J. Goblet, R.F. Coelho, Adaptive sampling strategies for non-intrusive POD-based surrogates, Eng. Comput. 30 (4) (2013) 521–547, http://dx.doi.org/10.1108/02644401311329352.