The 9th European Congress on Computational Methods in Applied Sciences and Engineering
ECCOMAS Congress 2024
3–7 June 2024, Lisboa, Portugal

COMPARISON BETWEEN LINEAR AND NON-LINEAR MULTIFIDELITY MODELS FOR TURBULENT FLOW PROBLEMS

Mikhail Glazunov[1†], Alistair Revell[1], Philipp Schlatter[2] and Saleh Rezaeiravesh[1]

[1] Department of Mechanical and Aerospace Engineering
The University of Manchester
Oxford Road, Manchester, M13 9PL, United Kingdom
[†] e-mail: mikhail.glazunov@manchester.ac.uk

[2] Institute of Fluid Mechanics (LSTM)
Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg
Cauerstraße 4, 91058 Erlangen, Germany

Key words: Multifidelity Modelling, Gaussian Process Regression, Turbulent Flows, Uncertainty Quantification.

Summary. This study compares two prominent multifidelity modelling approaches based on Gaussian Process Regression (GPR): linear co-kriging method and a non-linear autoregressive GP model. These methods are applied to a periodic hill flow case, to understand how variations in the hill geometry affect flow characteristics, particularly the height of the separation bubble. High- and low-fidelity turbulent flow data was obtained from Direct Numerical Simulations (DNS) and Reynolds-Averaged Navier-Stokes (RANS), respectively. The capabilities of the MFM methods were also evaluated for an uncertainty propagation problem and sensitivity analysis. The comparison of the two MFM approaches reveals that the non-linear method performed better than the linear model, but at a significantly higher computational cost. Moreover, both models provide similar satisfactory accuracy for the uncertainty propagation and global sensitivity analysis.

1 INTRODUCTION

The challenges engineers face often involve navigating complex design spaces, where an exhaustive exploration of all possibilities can be impractical due to time and cost constraints. The quest for optimal computational models has given rise to the innovative paradigm of applying machine learning and data science approaches to computational modelling problems to improve their cost effectiveness. Among them, multifidelity modelling (MFM) is a promising approach that aims to combine data of differing levels of fidelity to enhance the predictive accuracy while minimising overall computational cost [9, 15]. This study compares two prominent MFM approaches based on Gaussian Process Regression (GPR) [17]. The first method is the linear co-kriging model by Forrester *et al.* [5] based on Ref. [9], and the other method is a non-linear autoregressive Gaussian process (NARGP) model by Perdikaris *et al.* [16]. The methods will be applied to the turbulent flow over periodic hills [22, 18] with the overarching objective of

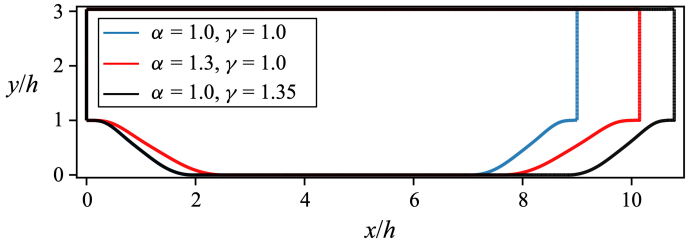

Figure 1: Variation of the geometry of the periodic hill problem due to the uncertain parameters α and γ [18].

studying how the flow is affected by changes in the geometry of the periodic hills due to two uncertain parameters, and to expand on previous studies of applying MFM to turbulent flow problems such as those by Mole *et al.* [13], Ahlfeld *et al.* [1], and de Baar *et al.* [3]. The surrogate of the quantities of interest (QoIs) is constructed in the space of design parameters using the co-kriging and NARGP models. Then using Polynomial Chaos Expansion (PCE) [25], the propagated uncertainty in the QoIs can be quantified. The QoI for this problem is considered to be the height of the separation bubble at some x/h which can vary greatly depending on the geometry of the hill. Moreover, the periodic hill case was chosen as a way to highlight the differences between low- (LF) and high-fidelity (HF) turbulent flow data. HF data are generated using Direct Numerical Simulation (DNS). Such methods give the most accurate data of turbulence, but at a very high computational cost, especially at higher Reynolds numbers. Therefore, lower-fidelity methods such as Reynolds-Averaged Navier-Stokes (RANS) [23], relying on a statistical interpretation of turbulence, are often used as a way to compute complex flow problems at a reasonable computational cost. Unfortunately, such methods come with a low accuracy especially for separated flows that MFM methods aim to improve.

The paper is structured as follows. First, the problem setup is described. Then, the theory of the GPR, co-kriging, and NARGP models are briefly explained. Next, we discuss the results of single-fidelity GPR to create a "ground-truth" for the comparison study. Finally, we present and discuss the results from the MFM models and how they compare to each other and the ground truth.

2 PROBLEM DESCRIPTION

To highlight the effectiveness of the MFM in an outer-loop problem, we have considered a turbulent flow with separation over periodic hills. This canonical case is considered here with geometrical uncertainties due to two parameters α and γ. These uncertain parameters are assumed to be independent with uniform distributions: $\alpha \sim \mathcal{U}[0.448, 1.552]$ and $\gamma \sim \mathcal{U}[0.356, 1.644]$ [18]. The length of the lower wall varies as $L_x/h = 3.858\alpha + 5.142\gamma$, where h is the height of the hill.

The QoI for this problem is taken to be the height of the separation bubble at $x/h = 2.5$, which can vary greatly depending on the geometry of the hill and the approach to simulation of turbulence. The HF data can be chosen from the 9 DNS at the Reynolds number of $Re = 5600$ created by Xiao *et al.* [24]. The LF data consisted of 25 samples created using a $k-\omega$ SST turbulence model [12] by Rezaeiravesh *et al.* [18] at the same Reynolds number. Efficient multifidelity solutions aim to use as few HF samples as possible. Therefore, two different sampling schemes

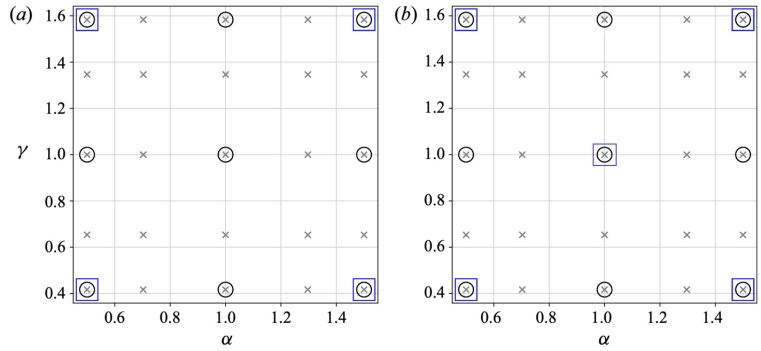

Figure 2: The two sampling schemes visualised [18]: LF samples - ×, total available HF samples (for constructing the ground truth) - ○, and selected HF samples - □. The latter is combined with the LF samples for multifidelity modelling. The sampling schemes are hereafter referred to as (left) the 25-4 and (right) 25-5 datasets.

combining 25 RANS simulations with 4 and 5 HF samples were considered to highlight the effect of HF samples on the overall accuracy of the multifidelity prediction. Figure 2 schematically shows the sampling schemes.

3 METHODS

3.1 Gaussian process regression (GPR)

GPR is a non-parametric probabilistic approach in machine learning used for predicting complex functions. This approach has the advantage of providing confidence intervals for its predictions, therefore, it is suitable for uncertainty quantification (UQ). A Gaussian process (GP) is "a collection of random variables, any finite number of which have a joint Gaussian distribution" [17]. Let $\mathbf{x}$ be the inputs or design/uncertain parameters and y represent the noisy observations. Adopting a noise-additive model, we have

$$y = f(\mathbf{x}) + \epsilon\,, \tag{1}$$

where $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$. A GP surrogate $f(\mathbf{x})$ is defined by its mean function $m(\mathbf{x})$ and the covariance function $k(\mathbf{x}, \mathbf{x}')$:

$$f(x) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))\,, \tag{2}$$

where,

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})]\,, \tag{3}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))]\,. \tag{4}$$

The mean and covariance functions in the prior GP (2) contain a set of hyperparameters θ which should be learned from a set of training data $\mathcal{D} = \{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^{n}$. The inference of θ can be based on deterministic and Bayesian approaches [17]. In the present study, we use the maximum likelihood estimator,

$$\hat{\theta}_{MLE} = \arg\max_{\theta} L(\theta) \tag{5}$$

in which, for an independent and identically distributed (iid) Gaussian noise, the likelihood function is defined as,

$$L(\theta) = \prod_{i=1}^{n} p(\epsilon_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_n} exp\left(-\frac{(y^{(i)} - f(\mathbf{x}^{(i)};\theta))^2}{2\sigma_n^2}\right). \tag{6}$$

The optimization in Eq. (5) was performed using the L-BFGS (gradient-based) method [11] in GPy [7] library. The resulting posterior of $f(\mathbf{x})$ is also a GP with the following mean and variance evaluated at test samples $\mathbf{X}^* = \{\mathbf{x}^{*^{(i)}}\}_{i=1}^{n^*}$:

$$\mathbb{E}[f^*] = K^{*^T}(K + \sigma_n^2 I)^{-1}\mathbf{Y}\,, \tag{7}$$

$$\mathbb{V}[f^*] = K^{**} - K^{*^T}(K + \sigma_n^2 I)^{-1}K^* \tag{8}$$

where, $\mathbf{Y}$ is the vector of training output values, $K^* = k(\mathbf{X}, \mathbf{X}^*)$ is the kernel covariance matrix of size $n \times n^*$, and $k(\cdot,\cdot)$ is the kernel covariance function. Similarly, $K = k(\mathbf{X}, \mathbf{X})$ and $K^{**} = k(\mathbf{X}^*, \mathbf{X}^*)$ represent the covariance function between the training-test and training-training samples, respectively.

3.2 Co-kriging MFM

In geostatistical science, the GPR that is used for interpolation of single-variate functions is called kriging, after Danie G. Krige [10]. **Co-kriging**, is an extension of Kriging for prediction of multifidelity and multi-output (multivariate) functions. In the context of MFM, each output can be attributed to a specific fidelity level. Without loss of generality, an overview to the formulation of the co-kriging MFM for two fidelities is presented here. The LF and HF parameter samples are denoted by $\mathbf{X}_L$ and $\mathbf{X}_H$ with the corresponding output values $\mathbf{Y}_L$ and $\mathbf{Y}_H$, respectively. Naturally, we deal with cases where $n_H < n_L$. The samples of parameters and their output values can be concatenated:

$$\mathbf{X} = \begin{bmatrix}\mathbf{X}_L \\ \mathbf{X}_H\end{bmatrix} = [\mathbf{x}_L^{(1)}, ..., \mathbf{x}_L^{(n_L)}, \mathbf{x}_H^{(1)}, ..., \mathbf{x}_H^{(n_H)}]^T\,, \tag{9}$$

$$\mathbf{Y} = \begin{bmatrix}\mathbf{Y}_L(\mathbf{X}_L) \\ \mathbf{Y}_H(\mathbf{X}_H)\end{bmatrix} = [\mathbf{Y}_L(\mathbf{x}_L^{(1)}), ..., \mathbf{Y}_L(\mathbf{x}_L^{(n_L)}), \mathbf{Y}_H(\mathbf{x}_H^{(1)}), ..., \mathbf{Y}_H(\mathbf{x}_H^{(n_H)})]^T\,. \tag{10}$$

Following the autoregressive model of Kennedy & O'Hagan [9]:

$$cov\{\mathbf{Y}_H(\mathbf{x}^{(i)}), Y_L(\mathbf{x})|\mathbf{Y}_L(\mathbf{x}^{(i)})\} = 0, \quad \forall \mathbf{x} \neq \mathbf{x}^{(i)} \tag{11}$$

The local features of the LF and HF models can be represented by GPs $f_L(\cdot)$ and $f_H(\cdot)$, respectively. Using the autoregressive model of [9], Forrester *et al.* [5] proposed the co-kriging formulation as,

$$f_H(\mathbf{x}) = \rho f_L(\mathbf{x}) + f_D(\mathbf{x})\,, \tag{12}$$

where the HF GP is approximated by the LF GP multiplied by a scaling factor ρ plus a GP $f_D(\cdot)$ that represents the model discrepancy. As a generalization of single-variate GPR, the covariance

matrix for co-kriging should describe all covariances between LF and HF data. Using the training data, this matrix is written as,

$$C = \begin{bmatrix} \sigma_L^2 k_L(\mathbf{X}_L, \mathbf{X}_L) & \rho\sigma_L^2 k_L(\mathbf{X}_L, \mathbf{X}_H) \\ \rho\sigma_L^2 k_L(\mathbf{X}_H, \mathbf{X}_L) & \rho^2\sigma_L^2 k_L(\mathbf{X}_H, \mathbf{X}_H) + \sigma_D^2 k_D(\mathbf{X}_H, \mathbf{X}_H) \end{bmatrix}, \tag{13}$$

where k_L and k_D are the kernel covariance functions for the GPs associated with the low-fidelity and model discrepancy, respectively. Any choice for these kernels contains hyperparameters that together with the scaling factors σ_L, σ_D, and ρ must be inferred. To this end, MLE can be used as detailed in [5]. With the hyperparameters estimated, the co-kriging prediction of the HF GP at a test sample $\mathbf{x}^*$ is given by:

$$\hat{y}_H(\mathbf{x}^*) = \hat{m}(\mathbf{x}^*) + c^T C^{-1}(\mathbf{Y} - \mathbf{1}\hat{m}(\mathbf{x}^*)), \tag{14}$$

$$c = \begin{bmatrix} \hat{\rho}\hat{\sigma}_L^2 k_L(\mathbf{X}_L, \mathbf{x}^*) \\ \hat{\rho}^2\hat{\sigma}_L^2 k_L(\mathbf{X}_H, \mathbf{x}^*) + \hat{\sigma}_D^2 k_D(\mathbf{X}_H, \mathbf{x}^*) \end{bmatrix}, \tag{15}$$

where $\hat{m}(\mathbf{x}^*) = \mathbf{1}^T C^{-1}\mathbf{Y}/\mathbf{1}^T C^{-1}\mathbf{1}$ and the notation $\mathbf{1}$ designates a matrix of ones, whilst the notation $k_L(\mathbf{X}_H, \mathbf{x}^*)$ denotes a column vector of correlations of the form k_L between the data $\mathbf{X}_H$ and the new point $\mathbf{x}^*$. In the present study, the implementation of the co-kriging model in Emukit library [14] that relies on GPy [7] has been used.

3.3 Non-linear (NARGP) MFM

As suggested by Eq. (12), the linear co-kriging MFM is accurate only when the LF and HF data have a strong linear correlation. Therefore, the use of this model for RANS and DNS data has been found challenging, see *e.g.* [1, 22]. A method put forth by Perdikaris *et al.* [16], uses a probabilistic framework based on GPR and non-linear autoregressive methods that are capable of learning complex non-linear structures as well as safeguarding against LF models that may provide wrong trends. The concept of non-linear autoregressive GPR (NARGP) is based on the work of Kennedy and O'Hagan [9], and largely similar to the linear co-kriging model. To mitigate the shortcoming of using the scaling multiplier ρ in the co-kriging model, the non-linear model defines the connection between the LF and HF models through

$$f_H(\mathbf{x}) = z_L(f_L(\mathbf{x})) + \delta_H(\mathbf{x}), \tag{16}$$

where $\delta_H(\mathbf{x})$ is the model discrepancy and z_L is an unknown function that maps the outputs of the LF model to those of the HF model. This unknown function has its own GP surrogate, creating a *deep GP* [2] in Eq. (16). However, such a generality comes at a significant cost, as the intractability of deep GP algorithms involves variational approximations in the training procedure, leading to far more complex implementation and computational costs than standard GPR. One way to improve a completely separate GP prior while maintaining analytical tractability and favourable algorithmic complexity is to replace the GP prior f_L with the posterior GP of the previous inference level $f_L^*(\mathbf{x})$. Then, it is possible to summarise the multifidelity prediction as the additive structure of Eq. (16) whilst maintaining the independence assumption [8] between the GPs z_L and δ_H:

$$f_H(\mathbf{x}) = g_H(\mathbf{x}, f_L^*(\mathbf{x})), \tag{17}$$

where $g_H \sim \mathcal{GP}(f_H|0, k_L((\mathbf{x}, f_L^*(\mathbf{x})), (\mathbf{x}', f_L^*(\mathbf{x}')); \theta_H))$ and θ_H denotes the hyperparameters of the HF covariance kernel function k_H. With the independence assumption, δ_H is replaced by g_H, leading to an equivalent Markov property that the lower-fidelity posterior $z_L(f_L^*(\mathbf{x}))$ is not able to learn anything more about $f_H(\mathbf{x})$ from the model output $z_L(f_L^*(\mathbf{x}'))$, $for\ \mathbf{x} \neq \mathbf{x}'$. To introduce a more structured approach to the prior, the covariance kernel of g_H can be decomposed as [16]:

$$k_{H_g} = k_{H_\rho}(\mathbf{x}, \mathbf{x}'; \theta_{H_\rho} \cdot k_{H_f}(f_L^*(\mathbf{x}), f_L^*(\mathbf{x}'); \theta_{H_f}) + k_{H_\delta}(\mathbf{x}, \mathbf{x}'; \theta_{H_\delta}) \tag{18}$$

Here $k_{H_\rho}, k_{H_f}, k_{H_\delta}$ are valid covariance functions such as Exponentiated-Quadratic [6] or Matern [4] with associated hyperparameters $\theta_{H_\rho}, \theta_{H_f}, \theta_{H_\delta}$. The hyperparameters in the NARGP model are estimated like conventional GPs but via an MLE recursive scheme [8]. However, unlike co-kriging, the non-linear method needs to learn $2d+3$ hyperparameters for every input dimension d instead of $d+3$ due to the additional covariance kernel within g_H.

The first level of the non-linear model corresponds to the standard GP model using standard mean and covariance functions. However, on subsequent recursive levels, when the posterior distribution is no longer Gaussian, a prediction must be made given a test sample $(\mathbf{X}^*, f_L^*(\mathbf{X}^*))$. Consequently, for problems where the prediction needs to be made on uncertain inputs, the uncertainty is propagated along each recursive step. Thus, the posterior distribution is given by [16]:

$$\begin{aligned} p(f_H^*(\mathbf{X}^*)) &= p(f_H(\mathbf{X}^*, f_L^*(\mathbf{X}^*))|f_L^*, \mathbf{X}^*, \mathbf{Y}_H, \mathbf{X}_H) \\ &= \int p(f_H(\mathbf{X}^*, f_L^*(\mathbf{X}^*))|\mathbf{Y}_H, \mathbf{X}_H, \mathbf{X}^*) p(f_L^*(\mathbf{X}^*))\, d\mathbf{X}^* \,. \end{aligned} \tag{19}$$

To estimate the posterior $p(f_H^*(\mathbf{X}^*))$, the integrals are computed using a Monte Carlo method, and from there associated mean and covariance of the posterior are obtained.

4 RESULTS AND DISCUSSION

In this section, first, the single-fidelity model is applied to the LF and HF data. Then, the co-kriging and NARGP results will be analysed and compared against each other and the ground truth using the two sampling schemes represented in Figure 3. Furthermore, using different single- and multifidelity models, uncertainty propagation and global sensitivity analysis are conducted.

4.1 Single-fidelity GPR predictions

A ground truth is established using single-fidelity GPR on all 9 HF data samples shown in Figure 3 to create a baseline to compare the MFM models to. Additionally, a single-fidelity GPR is created using only LF data (25 samples) as a way to showcase the significant improving impact of MFM. The single-fidelity GPR was created using GPy [7] library with a Matern 5/2 covariance kernel [6]. The optimization of the hyperparameters was performed using the L-BFGS gradient descent algorithm [11]. Figure 3 shows how the QoI, that is the height of the separation bubble in the periodic hill flow at $x/h = 2.5$, varies with α and γ in the single-fidelity cases.

As expected, there is a significant difference between the LF and HF predictions due to the disparity in the RANS and DNS data. The prediction of the LF model does not follow the trends established by the ground truth due to the inaccuracy of the RANS $k-\omega$ SST model in capturing the flow separation. To ensure that learning of hyperparameters using an MLE [8] has

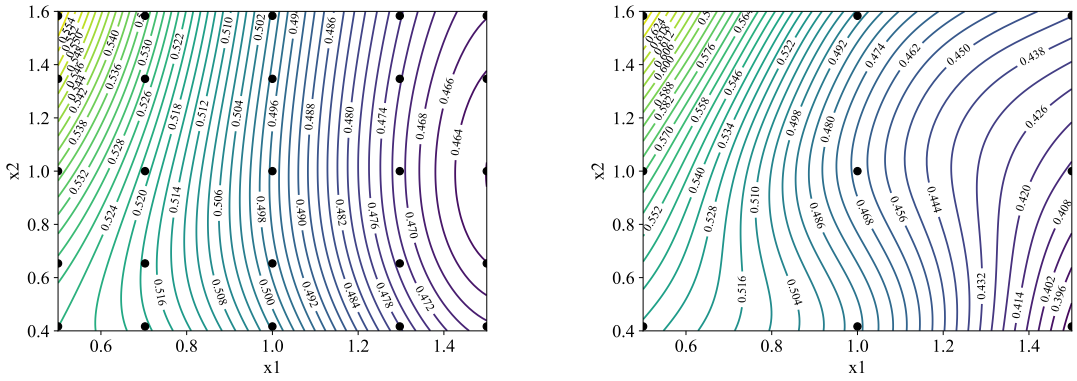

Figure 3: Comparison of the response surface of the mean QoI predicted by single-fidelity GPR with (left) only LF data and (right) all DNS samples (ground truth). Parameters $x1$ and $x2$ correspond α and γ, respectively.

been done accurately, the GPR isolines were confirmed by the results of non-intrusive polynomial chaos expansion (PCE) [25].

4.2 Multifidelity GPR predictions

Two multifidelity models were created using 25 LF(RANS) samples with 4 and 5 HF(DNS) samples to highlight the variation between sampling schemes on the accuracy of the predictions. The models used identical covariance kernel functions to have an unbiased comparison for the sampling schemes. A product of the Matern 5/2 [6] and Exponentiated-Quadratic [4] kernels was adopted to most accurately capture the correlation between the LF and HF data. The optimal kernel hyperparameters were achieved through applying the L-BFGS method [11] to associated MLE. Figure 4 shows the response surfaces of the mean QoI predicted by the co-kriging and NARGP models. Neither of the models could accurately create the isolines of the ground truth, although the NARGP showed better accuracy. The latter is not surprising, recalling that the co-kriging model relies mainly only on the linear scaling factor to account for for all correlations. Since there are significant systematic differences between RANS and DNS data, finding a scalar correlation between the two sample sets proves to be challenging. Nevertheless, the results showcased a viable prediction when using co-kriging for turbulent-flow problems and the effect of sampling when using tis method. The NARGP model, in contrast, relies on a function over the parameter space to capture the correlation between the two fidelities. This feature leads to more accurate predictions, particularly when using the 25-5 dataset. Even for the 25-4 dataset, the isolines of the mean predictions by the NARGP model are more similar to the ground truth compared to predictions by the co-kriging model.

The response surfaces provide a pointwise picture of the variation of the mean predictions over the parameter space. Alternatively, the predictions of the OoIs by different models can be represented by probability density functions (PDF). Figure 5 shows the joint PDF between the predictions of the co-kriging and NARGP models and the ground-truth. Moreover, Figure 6 provides a comparison between the PDFs for predictions by various single- and multifidelity models. In Figure 5, the use of the 25-5 dataset improves the predictions of the NARGP model

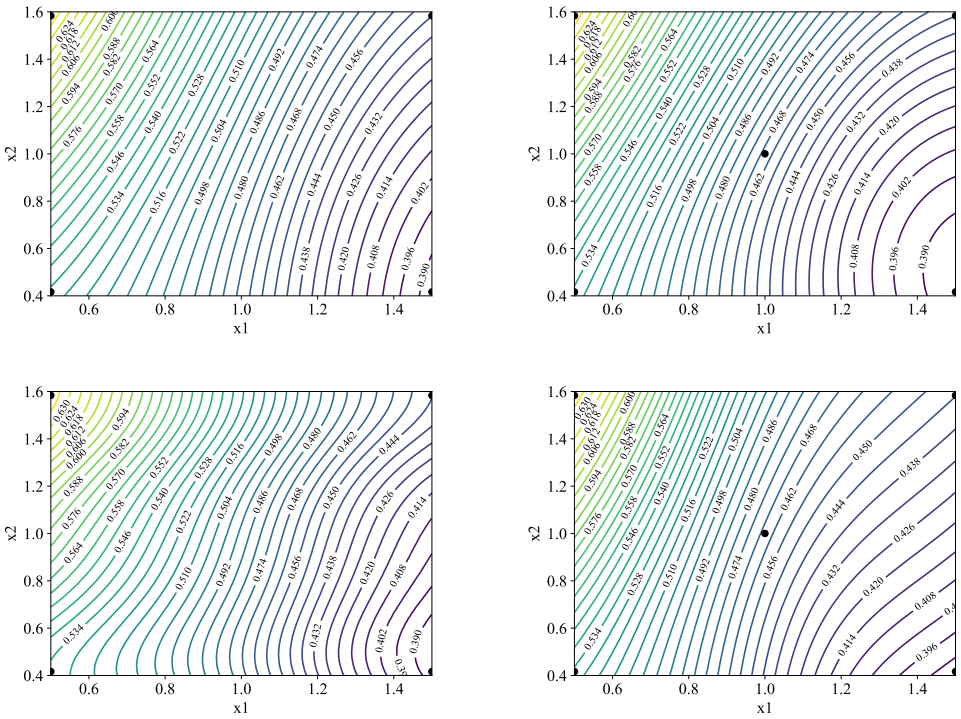

Figure 4: Response surface of the mean predicted QoI in the $\alpha - \gamma$ space using (top) co-kriging and (bottom) NARGP models with the (left) 25-4 and (right) 25-5 datasets.

by aligning the JPDF more with the reference median line. Surprisingly, adding an additional HF data sample for the co-kriging model is less effective and even deteriorating. There are two likely possibilities for this observation. The first is that the additional HF sample was misinterpreted by the model and an incorrect correlation was inferred that propagated throughout the predictions and skewed the JPDF. Conversely, the other reason would be a poor hyperparameter optimisation caused by gradient descent algorithms struggling to find the global maximum of the likelihood estimator. However, this issue is universal for all models in the present study, although its impact could be different. As suggested in Ref. [18], the solution is to adopt Bayesian approaches which can be pursued in future studies. According to Figure 6, the PDFs of the predicted QoI by single-fidelity GPRs using only LF and HF data are completely different. The goal of MFM is to combine the LF and HF data in such a way that the resulting PDF becomes close to the ground truth PDF that is bimodal. Clearly, this is achieved only by using the NARGP model with the 25-5 dataset, although some deviations persist. The deterioration of the PDF of the co-kriging model with the 25-5 dataset is also visible in this plot.

Sometimes in a UQ forward problem, we are only interested in specific stochastic moments of the QoI with respect to the uncertain parameters, rather than the full PDFs. Here, for the given training samples, we adopt the non-intrusive generalised polynomial chaos expansion (gPCE) [25] to estimate the stochastic mean and variance of the QoI. Assuming α and γ are random variables with uniform distribution over their specified range, Lagrange polynomial basis functions are used in gPCE. As a complement to the UQ forward problem, a global sensitivity

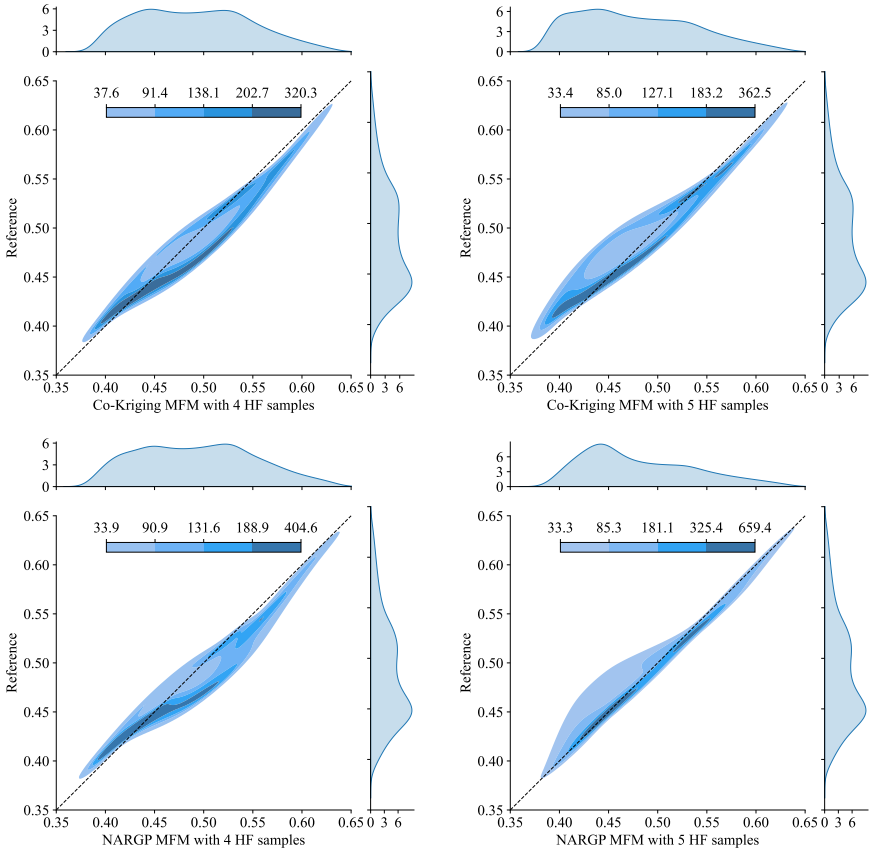

Figure 5: Joint PDFs between the mean QoI predicted by the multifidelity models and the ground truth values: (top) co-kriging and (bottom) NARGP models trained by the (left) 25-5 and (right) 25-4 datasets.

analysis has also be done. The stochastic moments and Sobol sensitivity indices [21] estimated by single- and multifidelity models using UQit [20] are summarised in Table 1. Comparatively, it is the single-fidelity model with LF data that has the most inaccurate estimations, especially for the standard deviation of the QoI and Sobol indices. The overall differences between the estimations by the single-fidelity GPR with HF data and multifidelity models is relatively small. However, the NARGP with the 25-5 dataset has the closest estimations to the ground truth, especially for the Sobol indices. A cautious conclusion, similar to [19, 18] is that the choice of the MFM and its LF-HF training data may have a small impact on the stochastic moments and Sobol indices of the QoI. If this holds, then the choice of the MFM strategy can be based on the overall computational cost and complexity. In the test case considered in the present study, the co-kriging model required, on average, 20 seconds on a single processor for training and prediction. For the same setup, the NARGP model was found to be nearly 300 times slower, showing the additional accuracy comes at a considerable computational cost.

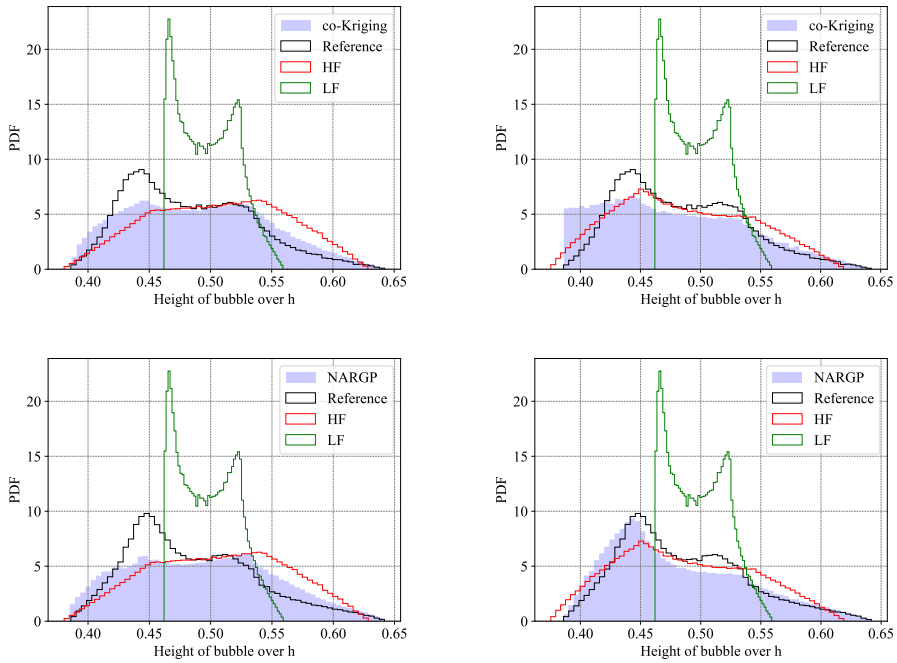

Figure 6: PDFs of the mean QoI predicted by the single-fidelity GPR and multifidelity models compared to the ground truth: (top) co-kriging and (bottom) NARGP models trained by the (left) 25-5 and (right) 25-4 datasets.

5 CONCLUSIONS

The linear co-kriging multifidelity model by Forrester *et al.* [5] and the non-linear autoregressive Gaussian process (NARGP) method by Perdikaris *et al.* [16] have been applied to the periodic hill flow problem [22, 18]. The aim was to quantify the impact of the variation of the hills geometry controlled by two uncertain parameters on the flow quantities. The QoI was particularly chosen to be the height of the separation bubble at a specific location. To compare the performance of different single- and multifidelity surrogates, the response surfaces, probability density functions (PDFs), stochastic moments and Sobol sensitivity indices of the QoI were studied. Two training datasets combining 25 RANS simulations with 4 and 5 DNS data samples were utilized. Compared to single fidelity Gaussian process regression (GPR), both multifidelity models showed improved predictions especially for the 25-5 dataset. However, the NARGP model outperformed the co-kriging model for accurate prediction of response surfaces and PDFs. This highlights the crucial role of constructing a functional rather than a scalar for the correlations between the two fidelities. Nevertheless, the increased accuracy of the NARGP comes at a significant computational cost, as this model was found to be 300 times slower to train and predict compared to the co-kriging model. Similar to Ref. [19], where the co-kriging model was compared to a hierarchical Bayesian model, the impact of the multifidelity model on the stochastic moments and Sobol indices was found to be smaller than the pointwise indicators such as response surface and PDFs. As brought up by Ref. [18], a main concern about the performance of multifidelity modelling strategies relying on GPR is the use of point estimators

Table 1: Stochastic moments and total Sobol indices of the QoI with respect to α and γ for various models and datasets.

	Stochastic Moments		Total Sobol Indices of QoI	
Model	$\mathbb{E}[QoI]$	$\mathbb{S}[QoI]$	α	γ
Reference	0.48495	0.05151	0.99405	0.12084
LF	0.49813	0.02386	0.99620	0.03300
The 25-4 dataset				
Single-Fidelity HF	0.50635	0.05485	0.99969	0.19564
NARGP	0.49356	0.05822	0.99890	0.16264
co-kriging	0.49130	0.05663	0.99891	0.14851
The 25-5 dataset				
Single-fidelity HF	0.48643	0.05542	0.99982	0.17560
NARGP	0.48129	0.05637	0.99898	0.15060
co-kriging	0.47849	0.05935	0.99916	0.16910

for learning the hyperparameters. An extension of the present study will be based on adopting Bayesian methods for such purpose to remove any potential bias in the interpretation of the models performance.

REFERENCES

[1] Ahlfeld, Laizet, G. Geraci, G. Iaccarino, and F. Montomoli. Multi-fidelity uncertainty quantification using RANS and DNS. *Proceedings of the CTR Stanford Summer Program*, 2016.

[2] A. Damianou and N. D. Lawrence. Semi-described and semi-supervised learning with Gaussian processes. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015.

[3] J. de Baar, S. Roberts, R. Dwight, and B. Mallol. Uncertainty quantification for a sailing yacht hull, using multi-fidelity kriging. *Computers & Fluids*, 123:185–201, 2015.

[4] D. K. Duvenaud. *The kernel cookbook: Advice on covariance functions.* Cambridge University Press, 2014.

[5] A. I. Forrester, A. Sóbester, and A. J. Keane. Multi-fidelity optimization via surrogate modelling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2088):3251–3269, 2007.

[6] M. G. Genton. Classes of kernels for machine learning: a statistics perspective. *J. Mach. Learn. Res.*, 2:299–312, mar 2002.

[7] GPy. GPy: A Gaussian process framework in Python. `http://github.com/SheffieldML/GPy`, since 2012.

[8] L. L. Gratiet and J. Garnier. Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4(5):365–386, 2014.

[9] M. Kennedy and A. O'Hagan. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13, 03 2000.

[10] D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6):119–139, 1951.

[11] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical programming*, 45(1-3):503–528, 1989.

[12] F. Menter. Zonal two equation $k-\omega$ turbulence models for aerodynamic flows. In *23rd fluid dynamics, plasmadynamics, and lasers conference*, page 2906, 1993.

[13] A. Mole, A. Skillen, and A. Revell. Multi-fidelity surrogate modelling of wall mounted cubes. *Flow, Turbulence and Combustion*, 110(4):835–853, 2022.

[14] A. Paleyes, M. Pullin, M. Mahsereci, C. McCollum, N. Lawrence, and J. González. Emulation of physical processes with Emukit. In *Second Workshop on Machine Learning and the Physical Sciences, NeurIPS*, 2019.

[15] B. Peherstorfer, K. Willcox, and M. Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60(3):550–591, 2018.

[16] P. Perdikaris, M. Raissi, A. Damianou, N. D. Lawrence, and G. E. Karniadakis. Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2198):20160751, 2017.

[17] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2005.

[18] S. Rezaeiravesh, T. Mukha, and P. Schlatter. Efficient prediction of turbulent flow quantities using a Bayesian hierarchical multifidelity model. *Journal of Fluid Mechanics*, 964, May 2023.

[19] S. Rezaeiravesh, T. Mukha, and P. Schlatter. Comparison between linear and hierarchical multifidelity models for uncertainty quantification in turbulent flows. In R. Örlü, A. Talamelli, J. Peinke, and M. Oberlack, editors, *Progress in Turbulence X*, pages 297–302, Cham, 2024. Springer Nature Switzerland.

[20] S. Rezaeiravesh, R. Vinuesa, and P. Schlatter. UQit: A Python package for uncertainty quantification (UQ) in computational fluid dynamics (CFD). *Journal of Open Source Software*, 6(60):2871, Apr. 2021.

[21] I. M. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 55(1-3):271–280, 2001.

[22] L. J. Voet, R. Ahlfeld, A. Gaymann, S. Laizet, and F. Montomoli. A hybrid approach combining DNS and RANS simulations to quantify uncertainties in turbulence modelling. *Applied Mathematical Modelling*, 89:885–906, 2021.

[23] D. Wilcox. *Turbulence Modeling for CFD*. Dcw Industries, 3 edition, Jan. 2006.

[24] H. Xiao, J.-L. Wu, S. Laizet, and L. Duan. Flows over periodic hills of parameterized geometries: A dataset for data-driven turbulence modeling from direct simulations. *Computers Fluids*, 200:104431, 2020.

[25] D. Xiu and G. E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2):619–644, 2002.