AIAA JOURNAL
Vol. 56, No. 12, December 2018

Multifidelity Surrogate Based on Single Linear Regression

Yiming Zhang,* Nam H. Kim,† Chanyoung Park,‡ and Raphael T. Haftka§
University of Florida, Gainesville, Florida 32611

DOI: 10.2514/1.J057299

Multifidelity surrogates (MFS) combine low-fidelity models with few high-fidelity samples to infer the response of the high-fidelity model for design optimization or uncertainty quantification. Most publications in MFS focus on Bayesian frameworks based on Gaussian process. Other types of surrogates might be preferred for some applications. In this paper, a simple and yet powerful MFS based on single linear regression is proposed, termed as *linear regression multifidelity surrogate* (LR-MFS), especially for fitting high-fidelity data with noise. The LR-MFS considers the low-fidelity model as a basis function and identifies unknown coefficients of both the low-fidelity model and the discrepancy function using a single linear regression. Because the proposed LR-MFS is obtained from standard linear regression, it can take advantage of established regression techniques such as prediction variance, D-optimal design, and inference. The LR-MFS is first compared with three Bayesian frameworks using a benchmark dataset from the simulations of a fluidized-bed process. The LR-MFS showed a comparable accuracy with the best Bayesian frameworks. The effect of combining multiple low-fidelity models was also discussed. Then the LR-MFS is evaluated using an algebraic function with different sampling plans. The LR-MFS bested co-kriging for 55 ~ 63% cases with an increasing number of high-fidelity (HF) samples. The sources of uncertainty with an increasing number of samples were also discussed. For both examples, the LR-MFS proved to be better than fitting only HF samples and robust with noisy data.

Nomenclature

$\boldsymbol{B}$	=	parameter vector for multifidelity surrogate
b_i	=	coefficient of $X_i(x)$
e	=	residual error vector between prediction and the value of high-fidelity samples
$f_H(\boldsymbol{x})$	=	high-fidelity model
$\hat{f}_H(\boldsymbol{x})$	=	multifidelity surrogate prediction
$f_L(\boldsymbol{x})$	=	low-fidelity model
$\hat{f}_L(\boldsymbol{x})$	=	low-fidelity surrogate prediction
$M(\boldsymbol{x})$	=	design vector at a point $\boldsymbol{x}$
$\boldsymbol{X}$	=	design matrix for multifidelity surrogate at samples
$X_i(\boldsymbol{x})$	=	ith polynomial term for the discrepancy function with p terms
$\boldsymbol{X}_H$	=	input matrix of n high-fidelity samples
$\boldsymbol{x}_H^{(j)}$	=	the input of the jth high-fidelity sample ($j = 1$ to n)
$\boldsymbol{Y}$	=	response vector for linear regression
$\boldsymbol{y}_H$	=	the vector for response of high-fidelity samples
$y_H^{(j)}$	=	the response value of the jth high-fidelity sample ($j = 1$ to n)
ρ	=	scale factor for multifidelity surrogate
$\hat{\delta}(\boldsymbol{x})$	=	surrogate fitting to the discrepancy between low- and high-fidelity samples

I. Introduction

SURROGATE models have been applied for various engineering design optimization and uncertainty quantification problems, which require many physical tests or simulations [1,2]. Because physical tests and high-fidelity simulations are usually time-consuming or expensive, surrogate models can be developed based on a small number of samples to infer the system response [3,4]. However, performing a number of tests or simulations needed for fitting an accurate surrogate is often too expensive. Multifidelity surrogates (MFS) can provide a solution for this problem by combining a small number of high-fidelity simulations or tests with lower accuracy models. Multifidelity models are commonly encountered for engineering analysis, such as finite element simulations with different resolutions and numerical simulations combined with physical tests.

Various frameworks have been proposed to predict mechanical system responses by combining data from different fidelities for design optimization (e.g., see reviews by Fernández-Godino et al. [5] and Peherstorfer et al. [6]). Efforts are emerging to extend MFS for various engineering analysis. Ng and Eldred [7] extended MFS with polynomial chaos expansion for uncertainty quantification (UQ). The multifidelity UQ process could converge more rapidly than a single-fidelity UQ. Aside from MFS for black-box dataset, other schemes have been proposed for error compensation to incorporate prior information from the dataset. Drohmann and Carlberg [8] proposed the reduced order modeling (ROM) error surrogates to map the physics-based error indicator of ROM to the true output errors using surrogate. The developed error surrogate is based on a single input (i.e., error indicator) and serves to compensate for the additive error.

MFS usually introduces a discrepancy surrogate to model the difference between low- and high-fidelity samples [9]. The popular Bayesian discrepancy-based MFS was introduced in [10–13] and demonstrated its effectiveness for the approximation of multifidelity datasets. Part of its success compared with earlier MFS was because of scaling the low-fidelity model in addition to introducing the discrepancy [14]. Model calibration associated with surrogate is another approach to combine multifidelity datasets. The physical parameters required by the low-fidelity model are first optimized and fitted by a surrogate in the design space. Then the approximated parameters are used to improve agreement between low-fidelity predictions and high-fidelity samples [15–17]. A comprehensive Bayesian MFS model that uses both calibration and discrepancy was proposed in [15], offering greater flexibility, although this is the most complex framework. However, the Bayesian framework requires a model for the uncertainty structure, and has been applied almost exclusively for Gaussian process (GP). As demonstrated by Viana et al. [18], there are situations where GP is not the most accurate surrogate, and so there is a merit to having an MFS arsenal that includes other surrogates.

Received 14 March 2018; revision received 1 July 2018; accepted for publication 5 August 2018; published online 12 October 2018.

*Graduate Student, Department of Mechanical and Aerospace Engineering. Student Member AIAA.

†Professor, Department of Mechanical and Aerospace Engineering. Associate Fellow AIAA.

‡Research Scientist, Department of Mechanical and Aerospace Engineering. Member AIAA.

§Distinguished Professor, Department of Mechanical and Aerospace Engineering. Fellow AIAA.

We propose a simple and yet powerful MFS based on regression, or least-squares fit, which is called linear regression multifidelity surrogate (LR-MFS). The high-fidelity behavior is approximated by a linear combination of low-fidelity predictions and a discrepancy function. In this paper, the discrepancy function is represented by a linear combination of monomial basis functions, but the approach is also applicable to other basis functions such as radial basis functions. The key idea is to consider the low-fidelity model as one additional basis function in the multifidelity model with the scale factor as a regression coefficient. Because a low-fidelity model is considered as a basis, including multiple low-fidelity models is straightforward in this approach. The design matrix in the regress equation consists of both the low-fidelity models and discrepancy basis functions. Then the scale factors and coefficients of the basis functions are obtained simultaneously by solving the regression set of linear equations. LR-MFS is expected to easily use various tools available for regression fits, such as prediction variance and D-optimal designs. An effective LR-MFS could serve as the basis for uncertainty propagation [19,20] and design optimization [21]. Compared with the Bayesian frameworks, LR-MFS has a closed-form solution for the optimum surrogate parameters and the training cost is essentially a few matrix multiplications. The Bayesian framework relies on expensive global optimization to find the surrogate parameters and the cost increases exponentially with sample size and dimensionality. The training of LR-MFS could be faster than the Bayesian framework by several orders of magnitude especially for large sample size in high-dimensional space.

The proposed LR-MFS assumes that the discrepancy could be captured by the scale factor and the polynomial form of additive discrepancy. This assumption has been also used by co-kriging and proved to be effective for various engineering analysis [5,22]. With a small number of high-fidelity samples, it would make sense to use a simple functional form for the discrepancy. Model validation schemes, such as cross-validation, could be used to evaluate the effectiveness of the MFS predictions and avoid potential misusage.

The paper is organized as follows. Section II briefs the Bayesian MFS and presents the proposed LR-MFS using single linear regression. Section III evaluates the proposed LR-MFS using a benchmark dataset from a fluidized-bed process. The LR-MFS is compared with three Bayesian frameworks. In Sec. IV, the LR-MFS is further evaluated using an algebraic test function with different levels of noise and different sampling plans. Section V concludes the technical contributions and major observations of this paper.

II. Linear Regression Multifidelity Surrogate

The basics of Bayesian MFS is discussed first to understand the major components of an MFS and potential limitations. Then the LR-MFS is proposed, which also includes a scale factor and a discrepancy function as with the Bayesian MFS.

A. Bayesian Multifidelity Surrogate Frameworks

Kriging [23] or Gaussian process is one of the most popular surrogate models for design optimization. The kriging is a stochastic process (a collection of random variables), such that every finite collection of those random variables has a multivariate normal distribution; that is, every finite linear combination of the variables is normally distributed. Kriging naturally provides the prediction variance at an unsampled point. The Bayesian multifidelity surrogate developed by Kennedy and O'Hagan [10] introduces the correlation between multifidelity dataset. The low-fidelity dataset is denoted as $\boldsymbol{X}_L = \{\boldsymbol{x}_L^{(1)}, \ldots, \boldsymbol{x}_L^{(n_L)}\}$ from the low-fidelity model $f_L(\boldsymbol{x})$, and the high-fidelity dataset is denoted as $\boldsymbol{X}_H = \{\boldsymbol{x}_H^{(1)}, \ldots, \boldsymbol{x}_H^{(n_H)}\}$ from the high-fidelity model $f_H(\boldsymbol{x})$. The corresponding function values are $\boldsymbol{y}_L = \{y_L^{(1)}, \ldots, y_L^{(n_L)}\}$ and $\boldsymbol{y}_H = \{y_H^{(1)}, \ldots, y_H^{(n_h)}\}$, respectively. The Bayesian multifidelity is made of two sets of correlated kriging models. The kriging surrogate $\hat{f}_L(\boldsymbol{x})$ is first constructed based on $(\boldsymbol{X}_L, \boldsymbol{y}_L)$. The second kriging $\hat{\delta}(\boldsymbol{x})$ is then built based on the discrepancies $\boldsymbol{y}_H - \rho\hat{f}_L(\boldsymbol{X}_H)$ as the discrepancy function. The scale factor ρ is estimated from maximum likelihood estimation as part of $\hat{\delta}(\boldsymbol{x})$. Variations of the Bayesian multifidelity surrogate are developed to improve accuracy and computational efficiency [11,24].

Since Kennedy and O'Hagan [10] introduced the GP-based Bayesian multifidelity surrogate, it has become popular and was found to do well [22]. However, the use of GP model also embraces technical difficulties regarding hyper parameter estimation. Finding the hyperparameters of the GP model is equivalent to finding a global optimum solution for a highly nonlinear likelihood function [25,26]. In addition to the computational burden of likelihood function evaluations, likelihood functions are often plagued with numerical instability due to covariance matrix inverse operation. Forrester et al. introduced co-kriging that is featured with better computational efficiency [27]. Qian and Wu proposed the use of Marko chain Monte Carlo (MCMC) and sample average approximation algorithm for hyperparameter estimation [11]. Le Gratiet proposed an approach to reduce the computational burden by simplifying the covariance matrix inversion operation [24]. For data with noise, GP-based surrogates tend to underestimate the noise by overly smoothing the prediction [28].

B. Proposed Linear Regression Multifidelity Surrogate

Because linear regression surrogates are sometimes more accurate than GP surrogates [16], it makes sense to have in our toolbox MFS based on linear regression surrogates. Similar to the comprehensive Bayesian framework; in this paper the following form of MFS is used:

$$\hat{f}_H(\boldsymbol{x}) = \rho f_L(\boldsymbol{x}) + \hat{\delta}(\boldsymbol{x}) \tag{1}$$

where ρ is the scale factor for the low-fidelity model, and $\hat{\delta}(\boldsymbol{x})$ is the discrepancy function. Different from the Bayesian frameworks, we do not use any approximation on the low-fidelity model. It will be shown later that it is unnecessary to have an explicit expression of the low-fidelity model. It should be enough that $f_L(\boldsymbol{x})$ can be evaluated at high-fidelity samples $\boldsymbol{X}_H$ and at prediction points.

Without loss of generality, in this paper, we assume that the discrepancy function is represented in the form of polynomial response surface (PRS) as

$$\hat{\delta}(\boldsymbol{x}) = \sum_{j=1}^{p} \xi_j(\boldsymbol{x}) b_j \tag{2}$$

where $\xi_j(\boldsymbol{x})$ denotes the jth monomial basis, and b_j is the unknown coefficient of $\xi_j(\boldsymbol{x})$. Although we explain our approach using a PRS, the proposed approach will work for any surrogates that can use linear regression. Because the number of high-fidelity samples is often limited, the number of terms in Eq. (2) is limited. We often use a constant or low-order polynomials for approximating the discrepancy function.

Because the model forms in Eqs. (1) and (2) are fixed, the only remaining task is to determine the unknown coefficients. The main observation in this paper is that we view the low-fidelity model as another basis function with the scale factor ρ as an unknown coefficient. Then, all unknown coefficients, ρ and b_j, $j = 1, \ldots, p$ can be determined using linear regression. Such seemingly natural observation turns out to be extremely simple MFS framework that can be easily applicable with single linear regression. Because we view the low-fidelity model as a basis, it is also possible to include multiple low-fidelity models together with different scale factors.

As with the standard least-squares method, the errors between MFS predictions and high-fidelity samples are defined as

$$\begin{aligned} e^{(i)} &= y_H^{(i)} - \hat{f}_H(\boldsymbol{x}_H^{(i)}) \\ &= y_H^{(i)} - \rho f_L(\boldsymbol{x}_H^{(i)}) - \sum_{j=1}^{p} \xi_j(\boldsymbol{x}_H^{(i)}) b_j, i = 1, \ldots, n_H \end{aligned} \tag{3}$$

Or, in the vector form, the vector of errors at high-fidelity samples can be written as

$$\boldsymbol{e} = \boldsymbol{Y} - \boldsymbol{XB} \tag{4}$$

where

$$\boldsymbol{e} = \begin{Bmatrix} e^{(1)} \\ \vdots \\ e^{(n_H)} \end{Bmatrix}, \quad \boldsymbol{Y} = \begin{Bmatrix} y_H^{(1)} \\ \vdots \\ y_H^{(n_H)} \end{Bmatrix},$$

$$\boldsymbol{X} = \begin{bmatrix} f_L(\boldsymbol{x}_H^{(1)}) & \xi_1(\boldsymbol{x}_H^{(1)}) & \cdots & \xi_p(\boldsymbol{x}_H^{(1)}) \\ \vdots & \vdots & \ddots & \vdots \\ f_L(\boldsymbol{x}_H^{(n_H)}) & \xi_1(\boldsymbol{x}_H^{(n_H)}) & \cdots & \xi_p(\boldsymbol{x}_H^{(n_H)}) \end{bmatrix}, \quad \boldsymbol{B} = \begin{Bmatrix} \rho \\ b_1 \\ \vdots \\ b_p \end{Bmatrix} \tag{5}$$

In the above equation, $\boldsymbol{e}$ is the vector of residual errors, $\boldsymbol{Y}$ is the vector of high-fidelity samples, $\boldsymbol{X}$ is the augmented design matrix, and $\boldsymbol{B}$ is the vector of unknown coefficients. By augmenting the design matrix, the scale parameter and unknown coefficients of the discrepancy functions are estimated simultaneously. It is clear from Eq. (5) that the low-fidelity model is included in the design matrix $\boldsymbol{X}$ with other basis functions of the discrepancy function.

The unknown coefficients in LR-MFS are obtained by minimizing the square sum of errors as

$$\underset{B}{\text{minimize}}\, \boldsymbol{e}^T\boldsymbol{e} = (\boldsymbol{Y} - \boldsymbol{XB})^T(\boldsymbol{Y} - \boldsymbol{XB}) \tag{6}$$

where the standard regression technique can yield the following form of unknown coefficients:

$$\boldsymbol{B} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \tag{7}$$

Linear regression assumes that the residual errors are normal, independent, and identically distributed. The standard error is estimated by

$$\hat{\sigma} = \sqrt{\frac{\boldsymbol{e}^T\boldsymbol{e}}{n_H - (p+1)}} \tag{8}$$

and the prediction variance at a point $\boldsymbol{x}$ is obtained as

$$\text{Var}\left[\hat{f}_H(\boldsymbol{x})\right] = \hat{\sigma}^2\boldsymbol{M}(\boldsymbol{x})^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{M}(\boldsymbol{x}) \tag{9}$$

where $\boldsymbol{M}(\boldsymbol{x}) = \{ f_L(\boldsymbol{x}) \quad \xi_1(\boldsymbol{x}) \quad \cdots \quad \xi_p(\boldsymbol{x}) \}^T$ denotes the design vector of linear regression at prediction point $\boldsymbol{x}$.

The scale factor ρ implies the level of trend similarity between $\hat{f}_H(\boldsymbol{x})$ and $f_L(\boldsymbol{x})$, and plays a critical role to approximate multifidelity data. Negative values or extremely large values of ρ indicate a risky prediction, which is likely to be associated with undesirable low-fidelity models, inappropriate surrogate forms, or inadequate samples. The discrepancy function $\hat{\delta}(\boldsymbol{x})$ is likely to be modeled well by a low-order PRS when $f_L(\boldsymbol{x})$ had a similar trend as $f_H(\boldsymbol{x})$.

There are nice properties of LR-MFS. First, the LR-MFS could be easily applied to more than two-fidelity models by augmenting the design matrix with multiple low-fidelity models with multiple scale factors as

$$\hat{f}_H(\boldsymbol{x}) = \sum_{j=1}^{k} \rho_j f_{Lj}(\boldsymbol{x}) + \sum_{j=1}^{p} \xi_j(\boldsymbol{x}) b_j \tag{10}$$

where k is the number of low-fidelity models. Second, it is possible to use other basis functions beyond monomials to approximate the discrepancy function. Third, the established LR-MFS can have several advantages, such as 1) handling noise from different distributions; 2) estimating confidence intervals, prediction intervals, and tolerance limits; 3) reducing model-form uncertainty via stepwise regression; and 4) obtaining optimal design of experiments (e.g., D-optimal designs).

III. Approximation of the Fluidized-Bed Process Using LR-MFS

In this section, the proposed LR-MFS was compared with three Bayesian multifidelity surrogates from the literature using the dataset from fluidized-bed process. The effect of different low-fidelity models was also discussed.

A. Benchmark Dataset: Fluidized-Bed Process

A comparison is made between the proposed LR-MFS and three Bayesian multifidelity surrogates developed by Kennedy and O'Hagan [10], Qian and Wu [11], and Le Gratiet [24]. This case was selected because the results for the Bayesian implementation are given in Le Gratiet [24]. Using these results from the literature removes possible implementation bias in comparison. The multifidelity surrogates are used to approximate the simulation of a fluidized-bed process [29]. The quantity of interest is the temperature of the steady-state thermodynamic operation point for a fluidized-bed process. There are six variables affecting the quantity of interest: humidity (H_r), room temperature (T_r), temperature of the air from the pump (T_a), flow rate of the coating solution (R_f), pressure of atomized air (P_a), and fluid velocity of the fluidization air (V_f). Twenty-eight different process conditions were considered with coating solution used for distilled water at room temperature. Four-fidelity outputs are available: $\boldsymbol{T}_{\text{exp}}$, $\boldsymbol{T}_3$, $\boldsymbol{T}_2$, and $\boldsymbol{T}_1$, with decreasing fidelity. $\boldsymbol{T}_{\text{exp}}$ is the experimental response denoted as high-fidelity data, $\boldsymbol{T}_3$ is the most accurate simulation considering the adjustments for heat losses and inlet airflow. $\boldsymbol{T}_2$ provides medium accuracy considering adjustment for heat losses. $\boldsymbol{T}_1$ has lowest accuracy without adjustment for either heat losses or inlet airflow.

B. Comparison with the Bayesian MFS Using Two-Fidelity Dataset

We focus on the prediction of $\boldsymbol{T}_{\text{exp}}$ (high-fidelity) assisted by $\boldsymbol{T}_2$ (low-fidelity) using two-level multifidelity surrogates as in Le Gratiet [24]. The dataset is collected at 28 points with 6 design variables for both $\boldsymbol{T}_{\text{exp}}$ and $\boldsymbol{T}_2$ as visualized in Fig. 1. $\boldsymbol{T}_2$ is highly correlated with $\boldsymbol{T}_{\text{exp}}$ having the correlation coefficient 0.99 and $\boldsymbol{T}_2$ overpredicts the value of $\boldsymbol{T}_{\text{exp}}$.

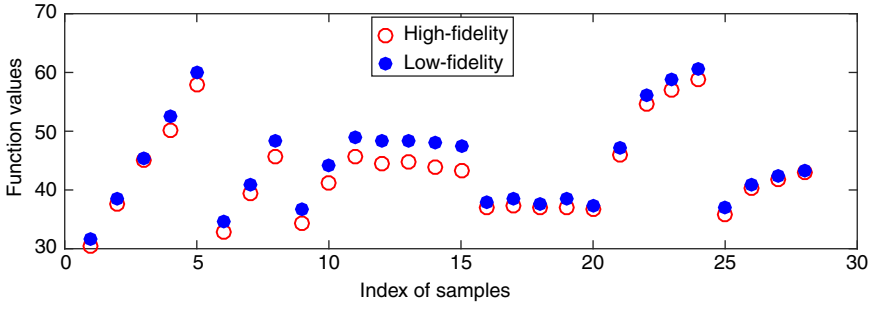


Fig. 1 Function values of the $\boldsymbol{T}_{\text{exp}}$ (high-fidelity) and $\boldsymbol{T}_2$ (low-fidelity) from fluidized-bed process [29].

Table 1 Evaluation plan for the approximation of multifidelity dataset from fluidized-bed process following Le Gratiet [24]

Steps	Procedures
1	Randomly select 20 out of 28 LF samples and then select 10 out of 20 HF samples as training points.
2	Use the other 18 HF samples as test set.
3	Generate 100 different combinations of training and test sets.
4	Use RMSE as prediction metric.

The evaluation is performed according to the plan proposed by Le Gratiet [24,30] as shown in Table 1. Twenty samples were randomly selected from $\boldsymbol{T}_2$ as LF samples as shown in Fig. 2. Then 10 high-fidelity (HF) samples were randomly selected such that the HF samples were nested into the LF samples. The 18 points left out from $\boldsymbol{T}_{\text{exp}}$ were used as the test set $(\boldsymbol{x}_T, \boldsymbol{y}_T)$. Root mean square error (RMSE) with the test set was adopted to evaluate prediction accuracy as

$$\text{RMSE} = \sqrt{\frac{1}{18}\sum_{j=1}^{18}\left(\hat{f}_H(\boldsymbol{x}_T^{(j)}) - \boldsymbol{y}_T^{(j)}\right)^2} \tag{11}$$

For comparison, 100 different combinations of training samples and test sets were produced. The model from Kennedy and O'Hagan [10] (KO) was implemented using R CRAN package "approximator"; the model from Qian and Wu [11] (QW) was implemented using the WinBugs software; and the model from Le Gratiet [24] (LG) was implemented using the R CRAN package "MuFiCokriging." Note that KO and LG are based on the same GP models and their theoretical characteristics are identical. However, they may give different results for the same problem because of the way of implementations and the performance of their optimizers to estimate hyper parameters.

It is noted that we have not produced any of the results reported below besides the LR-MFS, and instead took them from Le Gratiet [21].

The median RMSEs for the three Bayesian approaches are taken from Le Gratiet [24,30] as shown in Fig. 3. The comparison was based on 100 combinations of training samples and test sets. The compared approaches were PRS based on only HF samples, LR-MFS (LR), the model from LG, the model from KO, and the model from QW. For the construction of LR-MFS, a linear PRS was fitted to the LF samples and the LR-MFS was developed with a constant discrepancy. A single-fidelity linear PRS was developed based on only HF samples to study the effect of introducing LF models. The LR-MFS was more accurate than the PRS by introducing the LF samples. From Fig. 3, it is clear that LR-MFS and LG were the most accurate and were significantly better than the KO and QW models. The results for LR-MFS and PRS were both produced by the authors and highlighted in red. The results of the three Bayesian schemes (LG, KO, QW) were cited directly from Le Gratiet [24,30] for a fair comparison. The small difference between LG and LR-MFS is partially from the uncertainty from resampling. The scale factor ρ of LR-MFS is shown in Fig. 3b and varied between 0.73 and 1.15, which implied the similar trend between HF samples and LF samples. The median of ρ was 0.9, which was consistent with the observation that LF model overpredicted HF model.

C. Effect of Noise for High-Fidelity Dataset

High-fidelity samples might be obtained from noisy physical tests or simulations with noise from numerical errors. Therefore, LR-MFS was further evaluated by perturbing HF samples with synthetic noise from a normal distribution $N(0, 2^2)$. The noise level is comparable with the difference between HF and LF samples. The evaluation was repeated based on Table 1 with the perturbed HF samples. One realization of the HF samples with noise is shown in Fig. 4. RMSE of LR-MFS to approximate HF samples with noise is shown in Fig. 5 (based on 100 sets of samples). The noise level was estimated from Eq. (8). It is seen that LR-MFS was robust with respect to noise. The median RMSE only increased from 2.07 to 2.24 while the standard deviation of the noise was 2.

D. LR-MFS with Multiple Low-Fidelity Datasets

We compared the effect of different low-fidelity datasets using LR-MFS. $\boldsymbol{T}_{\text{exp}}$ was the high-fidelity dataset for prediction. $\boldsymbol{T}_3$, $\boldsymbol{T}_2$, and $\boldsymbol{T}_1$ were used as the low-fidelity datasets to build single-LF LR-MFS prediction, and then, all the $\boldsymbol{T}_3$, $\boldsymbol{T}_2$, and $\boldsymbol{T}_1$ were used to build a three-LF LR-MFS prediction. The evaluation was performed according to Table 1 using 100 sets of design of experiments. The RMSEs of all the LR-MFS's are shown in Fig. 6 based on 100 combinations of training samples and test sets. All the single-LF LR-MFS had similar accuracy around 2. The accuracy of LR-MFS improved along with the accuracy of low-fidelity model. One interesting observation is that the three-LF LR-MFS using all the $\boldsymbol{T}_i$ had worst accuracy and largest variation. The scale factors ρ are shown in Fig. 7 for further investigation of the results. For the two-fidelity LR-MFS, ρ was about

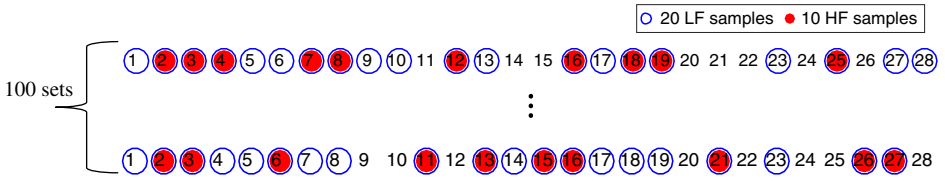


Fig. 2 Illustration of the resampling procedure to generate training and test data for fluidized bed process.

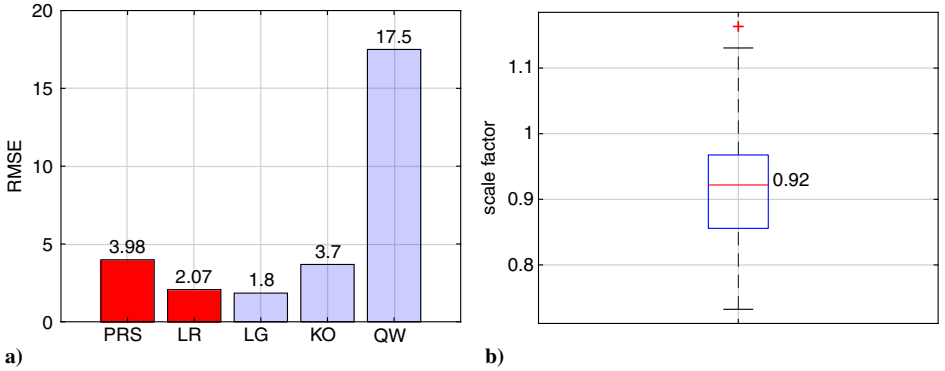


Fig. 3 Performance of the surrogate models. a) Median RMSE from surrogates. b) Scale factors ρ of LR-MFS.

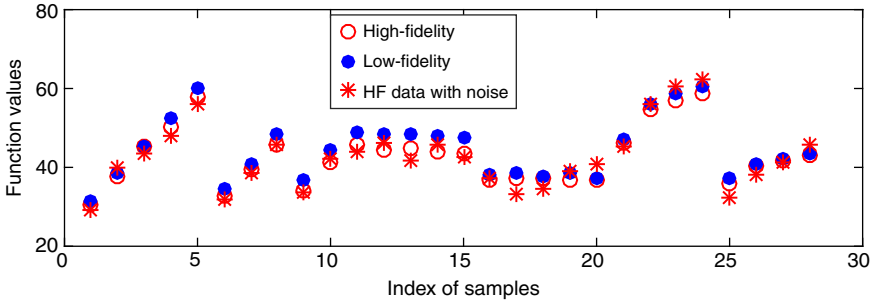


Fig. 4 Function values of the T_{exp} (high-fidelity) with one realization of synthetic noise from normal distribution, fluidized bed problem.

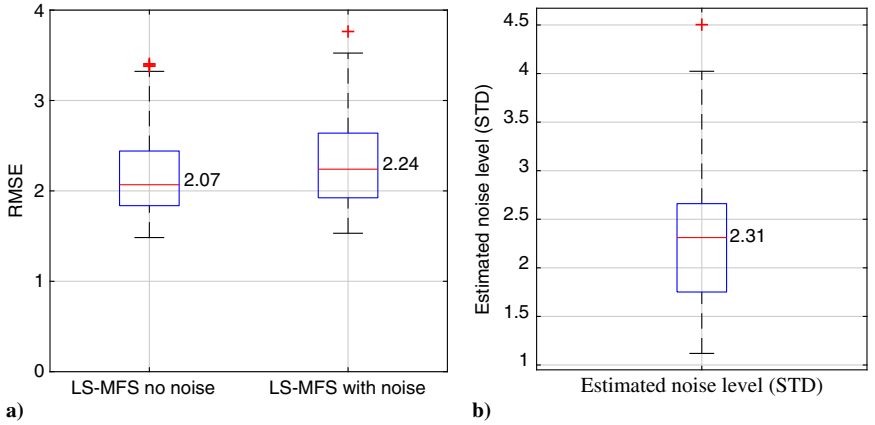


Fig. 5 RMSE of LR-MFS to approximate HF samples with noise, fluidized bed problem. a) RMSE of LR-MFS. b) Estimated noise level from LR-MFS.

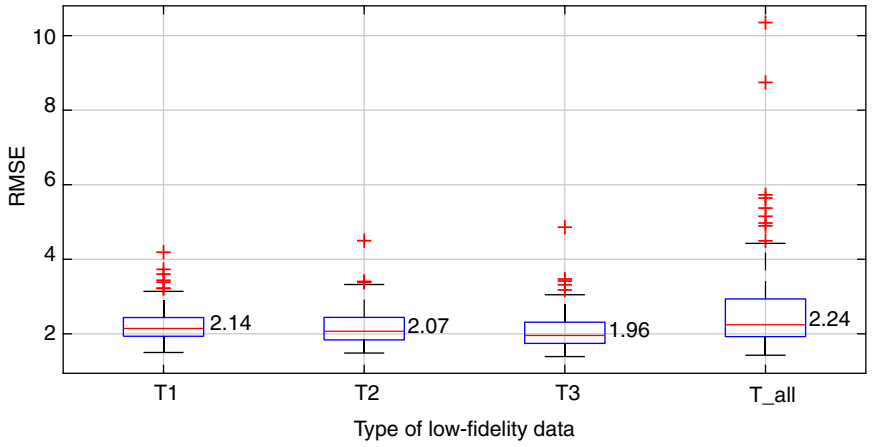


Fig. 6 RMSE of LR-MFS with different low-fidelity data sets and all the low-fidelity data sets together, fluidized bed problem.

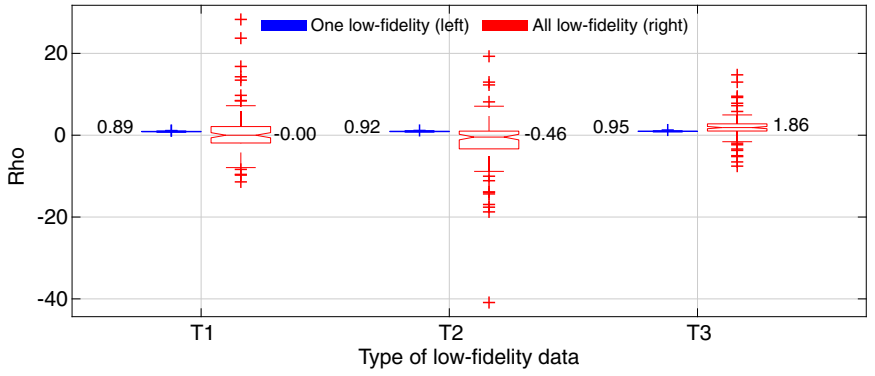


Fig. 7 Scale factors ρ of LR-MFS with different low-fidelity data sets and all the low-fidelity data sets together, fluidized bed problem.

Table 2 Major settings for the modified Currin test function

Input variables	Range of $f_H(\boldsymbol{x})$	Range of $f_L(\boldsymbol{x})$	Correlation coefficient between $f_H(\boldsymbol{x})$ and $f_L(\boldsymbol{x})$
$x_1, x_2 \in [0, 1]$	[1.1804, 13.7692]	[−0.1867, 4.9498]	0.99

0.9 with small variability. While for the three-LF LR-MFS, all the ρ were associated with significant variability and sometime negative. The undesirable behavior of ρ was due to the increasing number of parameters for surrogate training and possible correlation between low-fidelity models. Considering the limited number of high-fidelity samples, more parameters would suffer higher risk of overfitting. Regarding the median value, $\boldsymbol{T}_3$ had the highest fidelity and largest ρ, which implied significant contribution to the LR-MFS prediction. For $\boldsymbol{T}_2$ and $\boldsymbol{T}_1$, the ρ was close to 0 and LR-MFS assigned larger weighs to the dataset with higher fidelity.

Introducing more samples or adding a regularization term are effective techniques for handling overfitting. The prediction of the three-LF LR-MFS is likely to be improved after reducing the potential overfitting. The improved three-LF LR-MFS could be very close to the single-LF LR-MFS based on $\boldsymbol{T}_3$ considering that the trends between the multifidelity dataset are highly correlated. The discrepancy regression is a low-order PRS and could be estimated reasonably with introducing one set of low-fidelity dataset.

IV. Approximation of the Modified Currin Function Using LR-MFS

In this section, the proposed LR-MFS was evaluated using the modified Currin function that has a known expression. The effect of different level of noise and sampling plans were discussed.

A. Modified Currin Function

The Currin function [31,32] with two variables was adopted for further investigation of the proposed LR-MFS approach. The high-fidelity model $f_H(\boldsymbol{x})$ is given in Eq. (12). We modified the original low-fidelity model [33] with a larger scale factor and added a quadratic function as shown in Eq. (13). These variations make the multifidelity modeling more challenging. Major settings of the test function are summarized in Table 2. The function values of $f_H(\boldsymbol{x})$ vary between 1.2 and 13.8 with a range of 12.6, whereas the range of low-fidelity function is less than half of that of the high-fidelity function. However, the $f_H(\boldsymbol{x})$ and $f_L(\boldsymbol{x})$ are highly correlated with a correlation coefficient 0.99. The responses of $f_H(\boldsymbol{x})$ without noise and $f_L(\boldsymbol{x})$ are shown in Fig. 8.

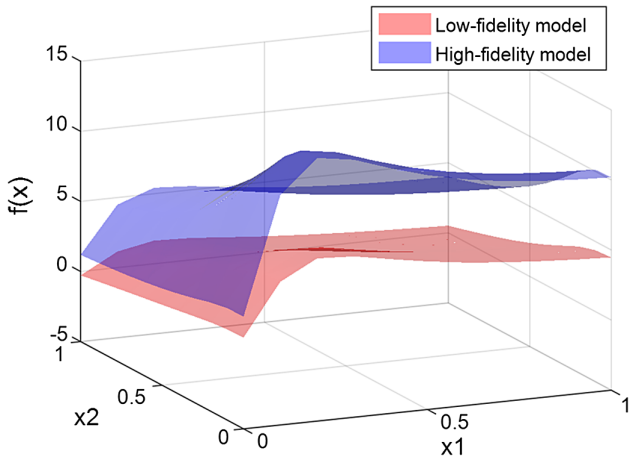


Fig. 8 High-fidelity and low-fidelity versions of the modified Currin function.

Table 3 Key factors of the evaluation plan for modified Currin function

Key factors	Plan
Effect of low-fidelity samples	Compare LR-MFS with PRS fitted to only high-fidelity samples.
Comparison with Bayesian MFS	Compare LR-MFS with co-kriging from Kennedy and O'Hagan [10].
Effect of noise for HF	Add synthetic noise to HF samples from $N(0, 0.1^2)$, $N(0, 0.2^2)$, $N(0, 0.3^2)$.
Effect of sample size	Investigate surrogates for 3 HF samples and 10 HF samples.
Prediction metric	RMSE
Fitting error to only LF samples	Low-fidelity sample was from $f_L(\boldsymbol{x})$, instead of a surrogate.
Sampling plan	Latin hypercube sampling, repeated 100 times with different samples.

$$f_H(\boldsymbol{x}) = \left[1 - \exp\left(-\frac{1}{2x_2}\right)\right]\frac{2300x_1^3 + 1900x_1^2 + 2092x_1 + 60}{100x_1^3 + 500x_1^2 + 4x_1 + 20} \tag{12}$$

$$\begin{aligned} f_L(\boldsymbol{x}) = {} & \frac{1}{8}\left[f_H(x_1 + 0.05, x_2 + 0.05) + f_H(x_1 + 0.05, \max(0, x_2 - 0.05))\right] \\ & + \frac{1}{8}\left[f_H(x_1 - 0.05, x_2 + 0.05) + f_H(x_1 - 0.05, \max(0, x_2 - 0.05))\right] \\ & + \frac{1}{8}\left(-5x_1 - 7x_2^2\right) \end{aligned} \tag{13}$$

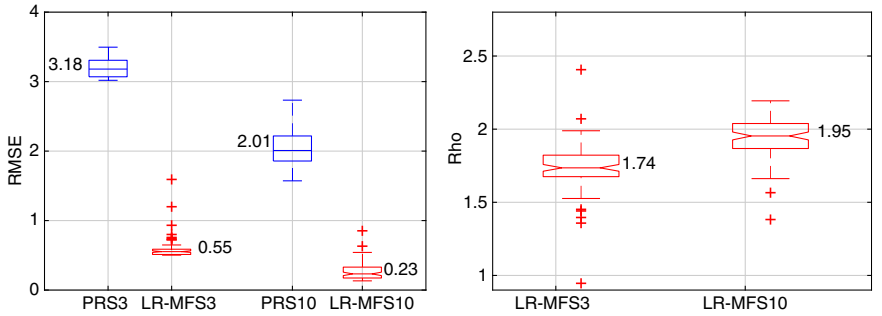
We have evaluated multiple aspects of the LR-MFS as summarized in Table 3. A PRS was fitted to only high-fidelity samples as the baseline and compared with LR-MFS in order to check on the usefulness of the low-fidelity samples. Then the LR-MFS was compared with the co-kriging from Kennedy and O'Hagan [10]. The co-kriging model was implemented through the ooDACE Toolbox [34]. The effect of sample size was analyzed by comparing the surrogates from 3 and 10 HF samples. Low-fidelity samples were produced at the same location as HF samples to build LR-MFS. To study the effect of noise, synthetic noise was added to HF samples with levels of noise $N(0, 0.1^2)$, $N(0, 0.2^2)$, and $N(0, 0.3^2)$. The prediction accuracy was evaluated from RMSE at 100×100 test grid from $f_H(\boldsymbol{x})$, and 100 sets of training samples was generated using Latin hypercube sampling (LHS). For the MFS models, the LF samples were obtained directly from $f_L(\boldsymbol{x})$ instead of a surrogate to avoid approximation error in building the low-fidelity surrogate. LF samples were generated at the same location of HF samples and test grid. It is also possible to build an $\hat{f}_L(\boldsymbol{x})$ based on only low-fidelity data for repeated calls of LR-MFS in practical applications.

B. Effect of Introducing Low-Fidelity Dataset

We first investigated the effect of introducing low-fidelity samples without noise. With 3 HF samples, a constant PRS and LR-MFS with a constant discrepancy were fitted. With 10 HF samples, a quadratic PRS and LR-MFS with a quadratic discrepancy were fitted. Accuracy of the PRS and LR-MFS is shown in Fig. 9 based on 100 sets of samples; 3 and 10 are the number of HF samples. The median RMSE of the 100 repetitions was used to represent the overall performance. LR-MFS was more accurate than PRS by the order of magnitude, and both LR-MFS and PRS improved noticeably from 3 to 10 HF samples. The median scale factors of LR-MFS increased from 1.74 to 1.95 with more samples. The high values reflect the fact that the low-fidelity function is less than half of the high-fidelity function.

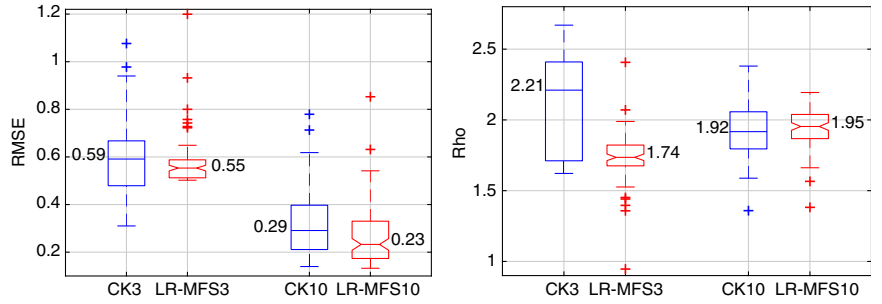
C. Comparison with Co-Kriging

Then the LR-MFS was compared with co-kriging without noise based on 100 sets of samples. It is seen in Fig. 10 that LR-MFS was



a) RMSE of surrogates b) Scale factors of LR-MFS

Fig. 9 Effect of introducing low-fidelity samples for prediction of modified Currin function without noise. a) RMSE of surrogates. b) Scale factors of LR-MFS.



a) RMSE of surrogates b) Scale factors for MFS

Fig. 10 Comparison of co-kriging (CK) with LR-MFS, for 3 and 10 HF samples, modified Currin function. a) Median RMSE of MFS. b) Median scale factors of MFS.

Table 4 Relative performance of LR-MFS and co-kriging for the approximation of the modified Currin function using 100 sets of HF samples

Number of HF samples	3 HF samples	10 HF samples
Cases better by LR-MFS	55	63
Cases better by co-kriging	45	37

slightly better than co-kriging for the median RMSE and has smaller variation. Co-kriging had larger scale factors for the approximation of 3 HF samples and similar scale factors for the approximation of 10 HF samples with LR-MFS. The relative performance between LR-MFS and co-kriging was summarized in Table 4. Among the 100 sets of HF samples, LR-MFS bested co-kriging for 55 cases and 63 cases while having 3 HF samples and 10 HF samples, respectively.

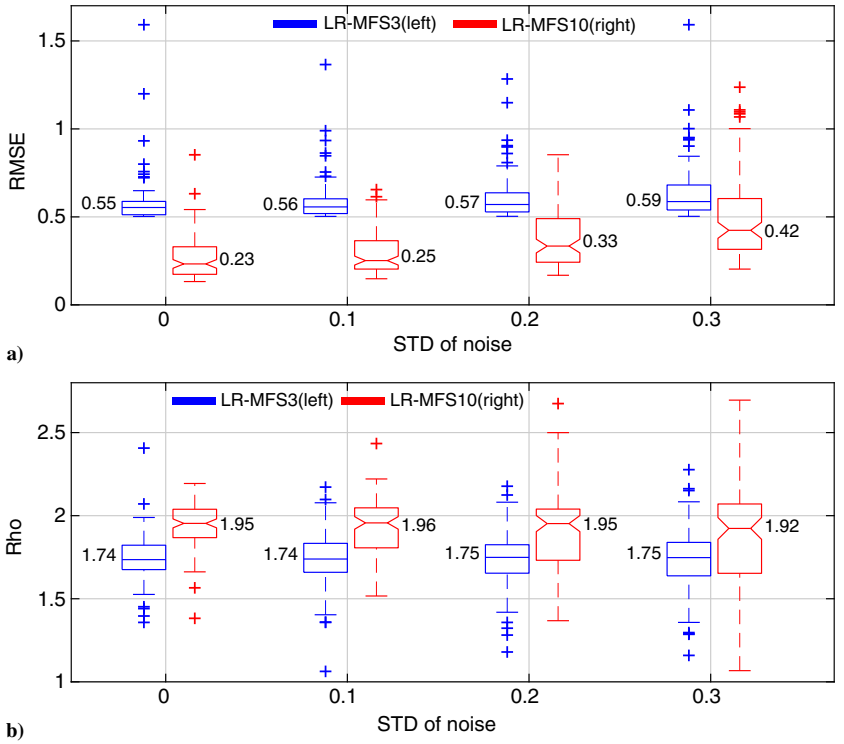


Fig. 11 Effect of noise on LR-MFS with 3 and 10 HF samples, modified Currin function. a) RMSE of the LR-MFS predictions. b) ρ of the LR-MFS predictions.

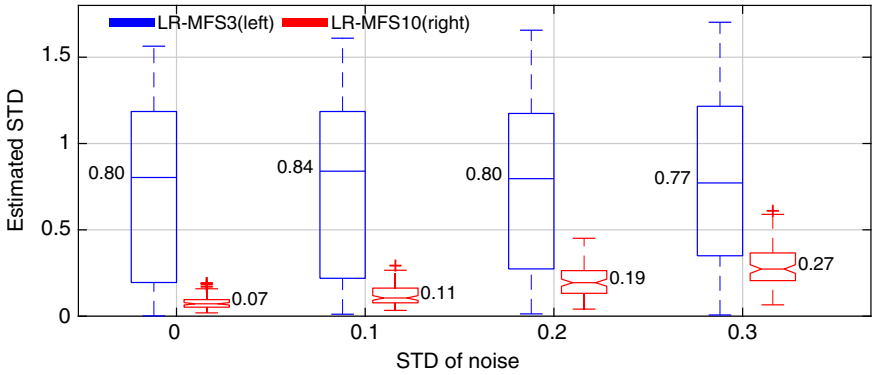


Fig. 12 Estimated noise level for the modified Currin function.

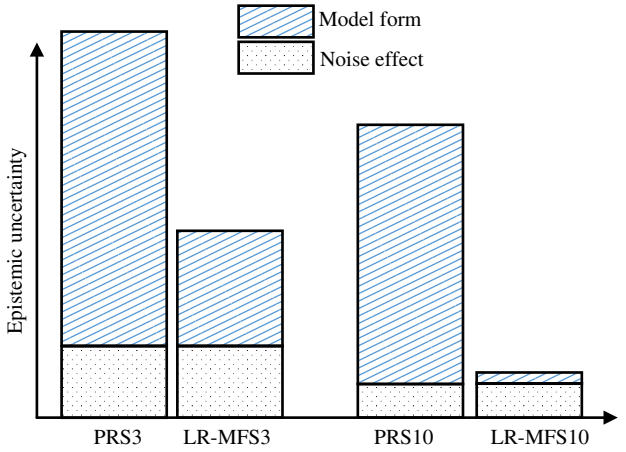


Fig. 13 Conceptual illustration for the major source of epistemic uncertainty to approximate modified Currin function with LR-MFS.

D. Fitting High-Fidelity Dataset with Different Level of Noise Using LR-MFS

HF samples might be associated with experimental variation. The HF samples were perturbed with synthetic noise from normal distributions $N(0, 0.1^2)$, $N(0, 0.2^2)$, $N(0, 0.3^2)$. The accuracy of LR-MFS for HF samples with noise is shown in Fig. 11. The computed RMSE in Fig. 11 is based on 100 sets of samples. The LR-MFS provided reasonable predictions for all the cases. The LR-MFS was expected to be robust with noise as inherited from PRS. LR-MFS with 3 high-fidelity samples (LR-MFS3) did not change much with increasing noise due to the large model-form uncertainty. In Fig. 11b, we can see that the low-fidelity samples contributed more on LR-MFS with 10 high-fidelity samples (LR-MFS10) than that on LR-MFS3. The estimated noise was computed according to Eq. (8) and summarized in Fig. 12. The estimated noise of modified Currin function is based on 100 sets of samples, with 3 and 10 HF samples, respectively. For LR-MFS3, the estimated noise level was significantly contaminated by model-form uncertainty. For LR-MFS10, the estimated noise level was close to the true value. The sources of epistemic uncertainty for LR-MFS prediction were conceptually illustrated in Fig. 13. The contribution of uncertainties varied with conditions (e.g., number of samples). The hazard of model form and noise effect decreased while introducing more samples. The effect of model form was minuscule for LR-MFS10 and therefore the estimated noise level was close to the true noise.

V. Conclusions

In this paper, the linear regression multifidelity surrogate (LR-MFS) was proposed to combine datasets with different fidelities, especially for high-fidelity with noise. The linear regression is commonly used, balanced between accuracy, cost, and simplicity. The LR-MFS is less likely to overfit noise by limiting the number of parameters. LR-MFS is derived from standard linear regression and therefore can use available tools, such as prediction variance and optimal DoE. The LR-MFS was demonstrated using polynomial response surface in this paper. The proposed LR-MFS was first compared with three Bayesian frameworks using a benchmark dataset from the simulations of a fluidized-bed process. The accuracy of Bayesian frameworks varied significantly, and the LR-MFS was comparable to the best GP-based approach. The effect of multiple low-fidelity dataset was also examined. For the case examined, it appears that introducing multiple low-fidelity models runs a higher risk of overfitting for limited number of high-fidelity data. Then the LR-MFS was evaluated to approximate a nonlinear numerical test function with different sampling plans. The LR-MFS was more accurate than co-kriging for 55–63% of cases with increasing number of HF samples. For both examples, the LR-MFS proved to be robust for HF samples with different levels of noise.

One major challenge for polynomial-based linear regression is the selection of basis terms. Various schemes exist, such as stepwise regression, to select the basis terms. For full-order PRS, the number of basis terms increases exponentially with dimensionality. Schemes to select the basis function or reduce dimension might be used for the approximation of sparse dataset in high-dimensional space. In general, multifidelity surrogates are used when the low-fidelity function is much cheaper than the high-fidelity function. In such a case, it is natural to assume that only a small number of high-fidelity samples, and thus discrepancy samples, are available. Therefore, it would not make sense to use a complex functional form for the discrepancy surrogate with many unknown parameters. Therefore, a lower-order polynomial or even a constant function can often be used as a discrepancy surrogate.

Compared with the Bayesian frameworks, LR-MFS has a closed-form solution for the optimum surrogate parameters and the training cost is essentially a few matrix multiplications. The Bayesian framework relies on expensive global optimization to find the surrogate parameters and the cost increases exponentially with sample size and dimensionality. The training of LR-MFS could be faster than the Bayesian framework by several orders of magnitude especially for large sample size in high-dimensional space.

Acknowledgment

This work was supported by the U.S. Department of Energy, National Nuclear Security Administration, Advanced Simulation and Computing Program, as a Cooperative Agreement under the Predictive Science Academic Alliance Program, under Contract No. DE-NA0002378.

References

[1] Forrester, A. I., and Keane, A. J., "Recent Advances in Surrogate-Based Optimization," *Progress in Aerospace Sciences*, Vol. 45, No. 1, 2009, pp. 50–79. doi:10.1016/j.paerosci.2008.11.001

[2] Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P. K., "Surrogate-Based Analysis and Optimization," *Progress in Aerospace Sciences*, Vol. 41, No. 1, 2005, pp. 1–28. doi:10.1016/j.paerosci.2005.02.001

[3] Frangos, M., Marzouk, Y., Willcox, K., and van Bloemen Waanders, B., Surrogate and Reduced-Order Modeling: A Comparison of Approaches for Large-Scale Statistical Inverse Problems, *Large-Scale Inverse Problems in Quantification of Uncertainty*, edited by L. Biegler, et al., Wiley, Chichester, U.K., 2010, Chap. 7, pp. 123–149.

[4] Zhang, Y., Meeker, J., Schutte, J., Kim, N., and Haftka, R., "On Approaches to Combine Experimental Strength and Simulation with Application to Open-Hole-Tension Configuration," *Proceedings of the American Society for Composites: Thirty-First Technical Conference*, DEStech Publ., Lancaster, PA, 2016, pp. 2205–2207.

[5] Fernández-Godino, M. G., Park, C., Kim, N.-H., and Haftka, R. T., "Review of Multi-Fidelity Models," arXiv preprint arXiv:1609.07196, 2016.

[6] Peherstorfer, B., Willcox, K., and Gunzburger, M., "Survey of Multifidelity Methods in Uncertainty Propagation, Inference and Optimization," Dept. of Aeronautics and Astronautics, Aerospace Computational Design Laboratory, Massachusetts Institute of Technology TR-16-1, Cambridge, MA, 2016.

[7] Ng, L. W.-T., and Eldred, M., "Multifidelity Uncertainty Quantification Using Non-Intrusive Polynomial Chaos and Stochastic Collocation," *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, AIAA Paper 2012-1852, 2012.

[8] Drohmann, M., and Carlberg, K., "The ROMES Method for Statistical Modeling of Reduced-Order-Model Error," *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 3, No. 1, 2015, pp. 116–145. doi:10.1137/140969841

[9] Balabanov, V., Grossman, B., Watson, L., Mason, W., and Haftka, R., "Multifidelity Response Surface Model for HSCT Wing Bending Material Weight," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA Paper 1998-4804, 1998.

[10] Kennedy, M. C., and O'Hagan, A., "Predicting the Output from a Complex Computer Code When Fast Approximations Are Available," *Biometrika*, Vol. 87, No. 1, 2000, pp. 1–13. doi:10.1093/biomet/87.1.1

[11] Qian, P. Z., and Wu, C. J., "Bayesian Hierarchical Modeling for Integrating Low-Accuracy and High-Accuracy Experiments," *Technometrics*, Vol. 50, No. 2, 2008, pp. 192–204. doi:10.1198/004017008000000082

[12] Han, Z., Zimmerman, R., and Görtz, S., "Alternative Cokriging Method for Variable-Fidelity Surrogate Modeling," *AIAA Journal*, Vol. 50, No. 5, 2012, pp. 1205–1210. doi:10.2514/1.J051243

[13] Kuya, Y., Takeda, K., Zhang, X., and Forrester, A. I., "Multifidelity Surrogate Modeling of Experimental and Computational Aerodynamic Data Sets," *AIAA Journal*, Vol. 49, No. 2, 2011, pp. 289–298. doi:10.2514/1.J050384

[14] Park, C., Kim, N. H., and Haftka, R. T., "Including ρ in Multi-Fidelity Surrogate Prediction Can Make Discrepancy Extrapolation Accurate by Reducing Bumpiness," *2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA Paper 2018-0915, 2018.

[15] Kennedy, M. C., and O'Hagan, A., "Supplementary Details on Bayesian Calibration of Computer Models," Internal Rept., 2001, http://www.shef.ac.uk/~st1ao/ps/calsup.ps.

[16] Higdon, D., Kennedy, M., Cavendish, J. C., Cafeo, J. A., and Ryne, R. D., "Combining Field Data and Computer Simulations for Calibration and Prediction," *SIAM Journal on Scientific Computing*, Vol. 26, No. 2, 2004, pp. 448–466. doi:10.1137/S1064827503426693

[17] Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. A., Cavendish, J., Lin, C.-H., and Tu, J., "A Framework for Validation of Computer Models," *Technometrics*, Vol. 49, No. 2, 2007, pp. 138–154. doi:10.1198/004017007000000092

[18] Viana, F. A., Haftka, R. T., and Steffen, V., Jr., "Multiple Surrogates: How Cross-Validation Errors Can Help Us to Obtain the Best Predictor," *Structural and Multidisciplinary Optimization*, Vol. 39, No. 4, 2009, pp. 439–457. doi:10.1007/s00158-008-0338-0

[19] Wang, L., Fang, X., Subramaniyan, A., Jothiprasad, G., Gardner, M., Kale, A., Akkaram, S., Beeson, D., Wiggs, G., and Nelson, J., "Challenges in Uncertainty, Calibration, Validation and Predictability of Engineering Analysis Models," *ASME 2011 Turbo Expo: Turbine Technical Conference and Exposition*, American Soc. of Mechanical Engineers, New York, June 2011, pp. 747–758.

[20] Sankararaman, S., Ling, Y., and Mahadevan, S., "Uncertainty Quantification and Model Validation of Fatigue Crack Growth Prediction," *Engineering Fracture Mechanics*, Vol. 78, No. 7, 2011, pp. 1487–1504. doi:10.1016/j.engfracmech.2011.02.017

[21] Jin, R., Du, X., and Chen, W., "The Use of Metamodeling Techniques for Optimization Under Uncertainty," *Structural and Multidisciplinary Optimization*, Vol. 25, No. 2, 2003, pp. 99–116. doi:10.1007/s00158-002-0277-0

[22] Park, C., Haftka, R. T., and Kim, N. H., "Remarks on Multi-Fidelity Surrogates," *Structural and Multidisciplinary Optimization*, Vol. 55, No. 3, 2017, pp. 1029–1050. doi:10.1007/s00158-016-1550-y

[23] Viana, F., *SURROGATES Toolbox User's Guide*, Ver. 3.0, Univ. of Florida, Gainesville, FL, 2011, https://sites.google.com/site/srgtstoolbox/.

[24] Le Gratiet, L., *Multi-Fidelity Gaussian Process Regression for Computer Experiments*, Université Paris-Diderot-Paris VII, Paris, France, 2013, pp. 114–120.

[25] Campbell, K., "Exploring Bayesian Model Calibration: A Guide to Intuition," Los Alamos TR LA-UR-02-7175, Albuquerque, NM, 2002.

[26] Swiler, L. P., and Trucano, T. G., "Calibration Under Uncertainty," Sandia National Lab. (SNL-NM) SAND2005-14981498, Albuquerque, NM, 2005.

[27] Forrester, A. I., Sóbester, A., and Keane, A. J., "Multi-Fidelity Optimization via Surrogate Modelling," *Proceedings of the Royal Society of London: A Mathematical, Physical and Engineering Sciences*, Vol. 463, The Royal Soc., London, U.K., 2007, pp. 3251–3269.

[28] Matsumura, T., Haftka, R. T., and Kim, N. H., "Accurate Predictions from Noisy Data: Replication Versus Exploration with Applications to Structural Failure," *Structural and Multidisciplinary Optimization*, Vol. 51, No. 1, 2015, pp. 23–40. doi:10.1007/s00158-014-1115-x

[29] Dewettinck, K., De Visscher, A., Deroo, L., and Huyghebaert, A., "Modeling the Steady-State Thermodynamic Operation Point of Top-Spray Fluidized Bed Processing," *Journal of Food Engineering*, Vol. 39, No. 2, 1999, pp. 131–143. doi:10.1016/S0260-8774(98)00144-7

[30] Le Gratiet, L., "Bayesian Analysis of Hierarchical Multifidelity Codes," *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 1, No. 1, 2013, pp. 244–269. doi:10.1137/120884122

[31] Currin, C., Mitchell, T., Morris, M., and Ylvisaker, D., "Bayesian Prediction of Deterministic Functions, with Applications to the Design and Analysis of Computer Experiments," *Journal of the American Statistical Association*, Vol. 86, No. 416, 1991, pp. 953–963. doi:10.1080/01621459.1991.10475138

[32] Bingham, S. S. D., "Virtual Library of Simulation Experiments: Test Functions and Datasets," Vol. 2016, 2016, http://www.sfu.ca/~ssurjano/index.html.

[33] Xiong, S., Qian, P. Z., and Wu, C. J., "Sequential Design and Analysis of High-Accuracy and Low-Accuracy Computer Codes," *Technometrics*, Vol. 55, No. 1, 2013, pp. 37–46. doi:10.1080/00401706.2012.723572

[34] Couckuyt, I., Dhaene, T., and Demeester, P., "ooDACE Toolbox: A Flexible Object-Oriented Kriging Implementation," *Journal of Machine Learning Research*, Vol. 15, No. 1, 2014, pp. 3183–3186.

R. Ghanem
Associate Editor