



# A multi-fidelity surrogate model based on moving least squares: fusing different fidelity data for engineering design

Shuo Wang<sup>1</sup> · Yin Liu<sup>1</sup> · Qi Zhou<sup>2</sup> · Yongliang Yuan<sup>1</sup> · Liye Lv<sup>1</sup> · Xueguan Song<sup>1</sup>

Received: 28 June 2020 / Revised: 6 May 2021 / Accepted: 4 August 2021 / Published online: 18 August 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

In numerical simulations, a high-fidelity (HF) simulation is generally more accurate than a low-fidelity (LF) simulation, while the latter is generally more computationally efficient than the former. To take advantages of both HF and LF simulations, a multi-fidelity surrogate (MFS) model based on moving least squares (MLS), termed as adaptive MFS-MLS, is proposed. The MFS-MLS calculates the LF scaling factors and the unknown coefficients of the discrepancy function simultaneously using an extended MLS model. In the proposed method, HF samples are not regarded as equally important in the process of constructing MFS-MLS models, and adaptive weightings are given to different HF samples. Moreover, both the size of the influence domain and the scaling factors can be determined adaptively according to the training samples. The MFS-MLS model is compared with three state-of-the-art MFS models and three single-fidelity surrogate models in terms of the prediction accuracy through multiple benchmark numerical cases and an engineering problem. In addition, the effects of key factors on the performance of the MFS-MLS model, such as the correlation between HF and LF models, the cost ratio of HF to LF samples, and the combination of HF and LF samples, are also investigated. The results show that MFS-MLS is able to provide competitive performance with high computational efficiency.

**Keywords** Multi-fidelity surrogate · Moving least squares · Adaptive influence domain · Adaptive scaling factor

## 1 Introduction

Numerical simulations have been widely used in the engineering design and optimization to facilitate quick exploration of design alternatives and obtain the optimal design. However, it is still challenging to deal with complex systems relying exclusively on numerical simulations due to the fact that the computational cost of high-fidelity (HF) simulations can be tremendous in spite of advances in computer capacity and speed nowadays (Viana et al. 2014). For example, it can take weeks to obtain the desired rollover crashworthiness and lightweight bus structure by applying an optimization algorithm directly to the simulations (Bai

et al. 2019). In fact, due to the lengthy running times of HF simulations, almost any optimization algorithm applied directly to the simulations will be slow (Forrester et al. 2008). An efficient way to speed up the design optimization process is to employ inexpensive surrogate models to replace the time-consuming HF simulations. Surrogate models are built, using data drawn from a small number of simulations, to provide fast approximations of the relationship between system inputs and outputs. There are many popular surrogate modeling techniques, which can be divided into two categories according to whether the surrogate model passes through the sample points. If the surrogate models pass through all of the sample points, these techniques fall into the category of interpolation, such as radial basis function (RBF) (Fang and Horstemeyer 2006; Majdisova and Skala 2017) and kriging (KRG) (Hao et al. 2018; Sacks et al. 1989). Otherwise, they will be regression, for example, polynomial response surface (PRS) (Kleijnen 2008; Myers et al. 2016), moving least squares (MLS) (Lancaster and Salkauskas 1981; Wang 2015), support vector regression (SVR) (Smola and Schölkopf 2004; Clarke et al. 2005) and artificial neural networks (ANN) (Basheer and Hajmeer

---

Responsible Editor: Vassili Toropov

✉ Xueguan Song  
sxg@dlut.edu.cn

<sup>1</sup> School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China

<sup>2</sup> School of Aerospace Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

2000; Napier et al. 2020). Each of these surrogate techniques has its own advantages and disadvantages, and it has been shown that there is no single surrogate technique that was found to be the most effective for all problems (Goel et al. 2007). More details and comparison of these techniques can be found in the literature (Jin et al. 2001; Wang and Shan 2007; Krishnamurthy 2005). Among the above-mentioned surrogate techniques, MLS is reported to be a very successful approximation scheme with advantages of high accuracy and low computational cost (Krishnamurthy 2005; Li et al. 2012), on the basis of which, it has been widely used in the engineering fields, such as structure reliability analysis (Lee et al. 2011; Lü et al. 2017), optimization for metal forming process (Breitkopf et al. 2005), and solar radiation estimation (Kaplan et al. 2020). Although a surrogate model saves an amount of computational resources, it still necessitates sufficient HF simulations to build the surrogate and ensure its accuracy, which tends to be unaffordable especially with increasing problem size. To address this challenge, multi-fidelity surrogate (MFS) models were proposed and have drawn much attention in the last two decades as they hold the promise of achieving the desired accuracy at a lower cost (Fernández-Godino et al. 2019).

Multi-fidelity surrogate is constructed by fusing many low-fidelity (LF) samples and a few HF samples. It is assumed that HF samples are more accurate than LF samples, but at a higher computational cost so that multiple evaluations of HF samples often cannot be afforded. LF samples, on the contrary, are cheaper than HF samples, but not as accurate as HF samples, though they could reflect the primary characteristics of the physical system. Moreover, it is worth noting that there is no clear boundary between the HF and LF samples. Whether a sample is HF or LF depends on the problem, and it can be determined based on the accuracy and cost against other fidelities (Fernández-Godino et al. 2019). For instance, the results obtained from experiments can be considered as HF samples, while the results from simulations as LF samples; on the other hand, results from simulations could also be regarded as HF as long as LF samples have even lower accuracy and cost. There are several ways to obtain LF samples, such as employing a simpler physical model, adopting a finite element model with coarse meshes, and reducing the dimensionality of the problem, etc.

The most popular approaches used for MFS are called correction-based approaches (Han and Görtz 2012; Fernández-Godino et al. 2016) or scaling function-based approaches (Zhou et al. 2017; Hao et al. 2020) which can be divided into three types: multiplicative, additive, and comprehensive approaches. MFS based on the multiplicative approach is constructed via multiplying the LF surrogate by a correction function, which denotes the ratio between HF and LF samples at the same locations. Multiplicative approach was first proposed by Haftka (1991) to develop

the global–local approximations (GLA) method with the aim of combining the advantages of both global and local approximations. Later, this approach was termed the variable complexity modeling technique and adopted by Hutchison et al. (1994), suggesting the use of a constant correction function for an approximation, and the constant correction function was calculated using a Taylor series expansion. Alexandov et al. (2001) constructed an MFS model using multiplicative corrections to solve aerodynamic optimization problems. Furthermore, Liu et al. (2014) adopted a kriging model, instead of a constant, as the correction function to calibrate the LF model. It should be noted, however, that the multiplicative approach might be ineffective if the values of the LF model are close to zero at some locations. This property, to some extent, has limited its application to design optimization, especially for constrained problems where a solution often locates at the constraint boundary, that is, the value of the constraint function (or surrogate) will be zero. Then, the additive approach was proposed to avoid the problem of division by zero occurred in multiplicative approach.

Multi-fidelity surrogate based on the additive approach is constructed by combining an LF surrogate with a discrepancy function, which models the difference between LF and HF samples. Eldred et al. (2004a, b) compared the performance of additive and multiplicative corrections in multi-fidelity surrogate-based optimization and showed that additive corrections were preferable to multiplicative corrections. Sun et al. (2011) employed an MFS using the additive approach, in which an MLS surrogate was built for the LF model and a PRS surrogate for the discrepancy function, to optimize sheet metal forming process. Zhou et al. (2015) used two SVR surrogates, representing the LF model and the discrepancy function, respectively, to build an MFS model. Since the additive approach has a simple form and is more robust than the multiplicative approach, it has been widely used in engineering optimization (Berci et al. 2014; Absi et al. 2019; Batra et al. 2019). A significant improvement in accuracy was achieved by combining the additive and multiplicative approaches, which leads to one of the most popular frameworks for MFS, namely the comprehensive approach. It is this approach that constitutes one pillar for the construction of co-kriging (Forrester et al. 2007). The other pillar of the co-kriging, an effective MFS method, is the correlated Gaussian process-based approximation which contains the information of LF and HF samples. Han et al. (2013) introduced the gradient information to co-kriging. Apart from kriging-based MFS, other MFS techniques under the comprehensive framework have been attracting interest as well. Mainini and Maggiore (2012) constructed multiple LF and HF models, and selected the best ones to build an MFS model. Zhang et al. (2018) proposed an MFS model based on linear regression (LR), named LR-MFS, by considering the LF model as an additional monomial in the MFS with the

scaling factor as a regression coefficient. Tao et al. (2019) introduced a deep learning-based MFS to robust aerodynamic design optimization. Durantin et al. (2017) proposed an MFS model based on RBF, which optimized the parameters by minimizing leave-one-out (LOO) cross-validation (CV) error. Song et al. (2019) combined the scaled LF model with a discrepancy function using two RBF surrogates in an MFS model and obtained a closed-form solution for the coefficients. Although efforts have been made to develop the MFS models, there is still space for investigating other MFS models to expand the arsenal of MFS models.

In this work, we proposed a simple and yet powerful MFS based on moving least squares, which is called adaptive MFS-MLS. The proposed MFS-MLS model is constructed using the comprehensive approach, which includes an LF scaling factor and a discrepancy function. In MFS-MLS, the scaling factor is a function of location and is multiplied with the LF model. The discrepancy function, modeling the difference between HF responses and the scaled LF model, is represented by an MLS model which consists of a linear combination of monomial basis functions. To compute the scaling factor and the coefficients of the basis functions, the predictions of the LF model and the basis functions are integrated into a matrix. Then, the scaling factor and the coefficients are correspondingly integrated into a coefficient vector and calculated by weighted least squares minimizing the error between the HF response and the prediction of the MFS-MLS. In addition, a new strategy was proposed to determine the size of the influence domain automatically. The MFS-MLS model allocates different weightings to HF samples within the influence domain, which distinguishes itself from global MFS models.

The remainder of this paper is organized as follows. Section 2 presents the details of the proposed MFS-MLS model. Comparisons between the proposed model and three MFS and three single-fidelity models on some benchmark numerical examples are given in Sect. 3. In Sect. 4, the MFS-MLS is applied to an engineering problem to further verify its effectiveness and applicability in dealing with practical problems. Conclusions and future work are drawn in Sect. 5.

## 2 The adaptive MFS-MLS methodology

The proposed MFS-MLS model is constructed using the comprehensive approach. The MFS-MLS model is adaptive because it can not only determine the size of influence domain automatically according to sample points, but also can determine the varying scaling factors at the prediction sites.

### 2.1 Adaptive MFS-MLS model

The comprehensive approach, possessing the advantages of additive and multiplicative approaches, has more flexibility and higher accuracy. Therefore, the comprehensive approach is employed in this paper and can be expressed as follows:

$$y_H(\mathbf{x}) = \rho y_L(\mathbf{x}) + \mathbf{d}(\mathbf{x}) \quad (1)$$

where  $\mathbf{x}$  represents design variables in the design space, and  $y_H(\mathbf{x})$  and  $y_L(\mathbf{x})$  denote the responses of HF and LF models, respectively.  $\rho$  is a scaling factor and plays an important role in approximating multi-fidelity data,  $\mathbf{d}(\mathbf{x})$  represents the difference between the scaled LF responses and the HF responses, called discrepancy function. The discrepancy function could be smoothed by selecting an appropriate scaling factor. The details about how a scaling factor improves multi-fidelity prediction can be found in the literature (Park et al. 2018).

The proposed MFS-MLS model is based on the MLS, thus, we give a brief introduction about MLS, more details about MLS, such as the derivation of equations, can be found in the literature (Lee et al. 2011; Lü et al. 2017; Breitung et al. 2005).

Moving least squares is a successful approximation scheme with advantages of both high accuracy and low computational cost (Krishnamurthy 2005; Li et al. 2012). In essence, MLS is an extension of the polynomial regression; however, there exist two significant differences from the traditional polynomial regression: (1) MLS recognizes that all sample points may not be equally important in estimating the regression coefficients. Therefore, each squared residuals are given a weighting when constructing the loss function. In addition, the weightings are varied depending upon the distance between the point to be predicted and each observed data point. (2) Unlike the traditional polynomial regression, the coefficients of an MLS model are not constant anymore, they are functions of input  $\mathbf{x}$ . At each prediction point  $\mathbf{x}_{\text{new}}$ , the coefficients are calculated using sample points within the neighborhood of the point  $\mathbf{x}_{\text{new}}$ . This neighborhood is referred to as the influence domain of the point  $\mathbf{x}_{\text{new}}$ , and the samples outside the influence domain are not considered. This method is termed moving least squares because the influence domain is “moving” as the prediction point changes.

Inspired by MLS, the MFS-MLS model based on the comprehensive correction can be built by the following equation:

$$\begin{aligned} \hat{y}_H(\mathbf{x}) &= a_0(\mathbf{x})y_L(\mathbf{x}) + \sum_{i=1}^m a_i(\mathbf{x})p_i(\mathbf{x}) \\ &= \mathbf{a}(\mathbf{x})^T \mathbf{p}(\mathbf{x}) \end{aligned} \quad (2)$$

where  $a_0(\mathbf{x})$  is the scaling function for the LF model  $y_L(\mathbf{x})$ , note that  $a_0(\mathbf{x})$  is not a constant, but rather a function of design variables  $\mathbf{x}$ , which provides more flexibility to the MFS-MLS model.  $p_i(\mathbf{x})$  is the monomial basis function,  $a_i(\mathbf{x})$  the coefficient of the basis function, and  $m$  is the number of terms in the basis.

In the MFS-MLS model, the conventional basis function vector of an MLS model is augmented as an integrated vector  $\mathbf{p}(\mathbf{x}) = [y_L(\mathbf{x})p_1(\mathbf{x}) \dots p_m(\mathbf{x})]^T$ .  $\mathbf{a}(\mathbf{x})$  is an augmented coefficient vector constituted by  $a_0(\mathbf{x})$  and  $a_i(\mathbf{x}), i = 1, 2, \dots, m$ . Note that  $a_0(\mathbf{x})$  is the counterpart of  $\rho$  in Eq. (1) and the term  $\sum_{i=1}^m a_i(\mathbf{x})p_i(\mathbf{x})$  of the MFS-MLS model is equivalent to the discrepancy function  $\mathbf{d}(\mathbf{x})$  of Eq. (1).

In the absence of specific knowledge about the characteristics of the real function, linear and quadratic monomials are often employed as the basis functions, e.g., a full quadratic basis in a two-dimensional (2D) space is of the form:  $\tilde{\mathbf{p}}(\mathbf{x}) = [1x_1x_2x_1^2x_1x_2x_2^2]^T$ . Therefore, the integrated vector  $\mathbf{p}(\mathbf{x})$  can be represented by  $\mathbf{p}(\mathbf{x}) = [y_L(\mathbf{x})\tilde{\mathbf{p}}(\mathbf{x})]$ .

To compute coefficient vector  $\mathbf{a}(\mathbf{x})$ , a cost function  $\mathbf{J}(\mathbf{a})$  that is a sum of weighted discrete  $L_2$  norms should be minimized:

$$\begin{aligned} \mathbf{J}(\mathbf{a}) &= \sum_{j=1}^{n_H} w(\mathbf{x} - \mathbf{x}_j) [\hat{y}_H(\mathbf{x}_j) - y_H(\mathbf{x}_j)]^2 \\ &= \sum_{j=1}^{n_H} w(\mathbf{x} - \mathbf{x}_j) [\mathbf{a}(\mathbf{x})^T \mathbf{p}(\mathbf{x}) - y_H(\mathbf{x}_j)]^2 \end{aligned} \tag{3}$$

where  $\mathbf{x}_j(j = 1, 2, \dots, n_H)$  are the  $n_H$  HF sample points in the neighborhood of the evaluation point  $\mathbf{x}$ ,  $w(\mathbf{x} - \mathbf{x}_j)$  is the weight function of the HF samples. The commonly used weight functions are Gaussian function, the cubic spline, the exponential function, and the quartic spline. The details about weight functions and identification of the size of the influence domain will be discussed in Sect. 2.2.

Essentially, Eq. (3) is a quadratic form so that it can be rewritten in the matrix form as follows:

$$\mathbf{J}(\mathbf{a}) = (\mathbf{Pa} - \mathbf{y}_H)^T \mathbf{W}(\mathbf{x})(\mathbf{Pa} - \mathbf{y}_H) \tag{4}$$

where

$$\mathbf{y}_H = [y_H(\mathbf{x}_1)y_H(\mathbf{x}_2) \dots y_H(\mathbf{x}_{n_H})]^T \tag{5}$$

$$\mathbf{P} = \begin{bmatrix} y_L(\mathbf{x}_1) & p_1(\mathbf{x}_1) & \dots & p_m(\mathbf{x}_1) \\ y_L(\mathbf{x}_2) & p_1(\mathbf{x}_2) & \dots & p_m(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ y_L(\mathbf{x}_{n_H}) & p_1(\mathbf{x}_{n_H}) & \dots & p_m(\mathbf{x}_{n_H}) \end{bmatrix} \tag{6}$$

and

$$\mathbf{W}(\mathbf{x}) = \begin{bmatrix} w(\mathbf{x} - \mathbf{x}_1) & 0 & \dots & 0 \\ 0 & w(\mathbf{x} - \mathbf{x}_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w(\mathbf{x} - \mathbf{x}_{n_H}) \end{bmatrix} \tag{7}$$

Taking the derivatives of Eq. (4) w.r.t  $\mathbf{a}(\mathbf{x})$  and setting to zero, we have

$$\mathbf{A}(\mathbf{x})\mathbf{a}(\mathbf{x}) = \mathbf{B}(\mathbf{x})\mathbf{y}_H \tag{8}$$

where matrices  $\mathbf{A}(\mathbf{x})$  and  $\mathbf{B}(\mathbf{x})$  are

$$\mathbf{A}(\mathbf{x}) = \mathbf{P}^T \mathbf{W}(\mathbf{x}) \mathbf{P} \tag{9}$$

$$\mathbf{B}(\mathbf{x}) = \mathbf{P}^T \mathbf{W}(\mathbf{x}) \tag{10}$$

Hence, we obtain

$$\mathbf{a}(\mathbf{x}) = \mathbf{A}^{-1}(\mathbf{x})\mathbf{B}(\mathbf{x})\mathbf{y}_H \tag{11}$$

Substituting  $\mathbf{a}(\mathbf{x})$  into Eq. (2), the MFS-MLS approximation  $\hat{y}_H(\mathbf{x})$ , can be obtained as

$$\hat{y}_H(\mathbf{x}) = \mathbf{p}^T(\mathbf{x})\mathbf{A}^{-1}(\mathbf{x})\mathbf{B}(\mathbf{x})\mathbf{y}_H \tag{12}$$

### 2.2 Weight function and the size of influence domain

Both the weight function and the size of influence domain have crucial impacts on the performance of the MLS (Lü et al. 2017), since the MFS-MLS is based on the MLS, it is reasonable to choose the weight function and the size of influence domain carefully in the process of constructing an MFS-MLS model. In this study, we selected the popular exponential function as the weight function and proposed a straightforward strategy to identify the size of the influence domain automatically.

The exponential function is adopted as the weight function and can be expressed by:

$$w(\bar{s}) = \begin{cases} e^{-(2\bar{s})^2}, & \bar{s} \leq 1 \\ 0, & \bar{s} > 1 \end{cases} \tag{13}$$

The size of the influence domain is determined by Euclidean distance  $\bar{s}$  which depends on the number of terms in the basis function. Let  $s = \|\mathbf{x} - \mathbf{x}_j\|$ , which denotes the Euclidean distance between the evaluation point  $\mathbf{x}$  and  $j$ th HF sample point  $\mathbf{x}_j$ .  $\bar{s} = s/s_k$ , let  $s_1, s_2, \dots, s_k, \dots, s_{n_H}$  be the list of Euclidean distances between the evaluation point  $\mathbf{x}$  and all the HF sample points, sorted in ascending order. In this way,  $s_k$  represents the  $k$ th largest Euclidean distance. The subscript  $k$ , representing the number of terms in the basis function, is given in Table 1. Therefore, the size of influence domain is determined by Euclidean distance  $\bar{s}$  and  $\bar{s}$  depends

**Table 1** Basis functions and the number of HF samples within the influence domain

Basis function	$k$
linear	$2 + n$
quadratic	$1 + \frac{(n+1)(n+2)}{2}$

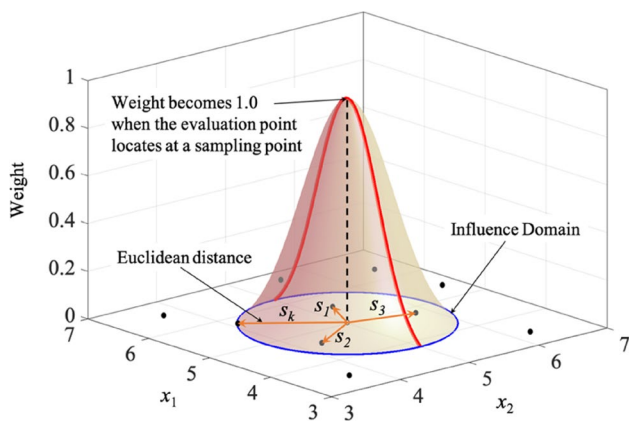
$n$  is the number of design variables

on  $k$ . Taking a 2D problem as an example, from Fig. 1, it can be observed that the influence domain is centered at the evaluation point, and the radius of the influence domain is the  $k$ th largest Euclidean distances  $s_k$ .

The basis functions in an MFS-MLS model consists of  $n + 1$  linear monomials or  $(n + 1)(n + 2)/2$  quadratic monomials. The number of unknown coefficients of an MFS-MLS model are  $n + 2$  or  $(n + 1)(n + 2)/2$  if the linear or quadratic monomials are employed as basis functions, respectively. Therefore,  $n + 2$  or  $1 + (n + 1)(n + 2)/2$  HF samples are sufficient to identify the coefficients of an MFS-MLS model, the radius of the influence domain can be set as the  $k$ th largest Euclidean distance between the evaluation point  $\mathbf{x}$  and all the HF sample points. The HF samples outside the influence domain are not involved in the calculation. However, if the number of HF samples is so scarce that  $n_H < k$ , then all the HF samples are included in the influence domain. This implies that it needs more and more samples for high-dimensional problems, otherwise the proposed strategy will turn the MFS-MLS model from local to global, which will incur a decline in the performance.

### 3 Numerical examples

To evaluate the performance of the MFS-MLS model, the MFS-MLS model is compared with three state-of-the-art benchmark MFS models (i.e., CoRBF proposed by Durantin



**Fig. 1** Illustration for weight function and influence domain

et al. (2017), LR-MFS proposed by Zhang et al. (2018), and MFS-RBF proposed by Song et al. (2019)) and three single-fidelity surrogate models (i.e., PRS, RBF, and MLS) on a number of widely used numerical test functions and one engineering problem.

Among the three benchmark MFS models, however, it should be noted that since the source code of CoRBF is not available, some specific parameters may be different from those in the original one. For the LR-MFS model, a first-order PRS is used to approximate the discrepancy function. For the three single-fidelity models, the SURROGATES Toolbox (Viana 2010) is employed to conduct the comparative experiments.

### 3.1 Design of experiments

Design of experiments (DOE) is the sampling plan in design space, which is generally the first step in the process of building a surrogate model. Among many available DOE techniques, Latin hypercube sampling (LHS) is chosen in this paper to generate samples due to its great capability of generating near-random samples uniformly. More specifically, for all surrogate models in this paper, the *lhsdesign* function, a Matlab built-in sampling function, is adopted to generate samples.

In this paper, it is assumed that the number of HF samples, used for building a single-fidelity surrogate model, is  $m \times n$ , where  $n$  is the dimension of the problem and  $m$  is a user-defined value. To compare the performance of MFS models and single-fidelity surrogate models fairly, the total computational budget of samples for building these two kinds of surrogate models is supposed to be equal. Therefore, to build an MFS model, the number of HF samples is set to  $k \times n$  ( $k < m$ ), and the remaining  $(m - k) \times n$  HF budget is replaced by more LF samples via cost ratio  $\theta$ . The cost ratio of HF samples to LF samples means that the cost of evaluating  $\theta$  LF samples is tantamount to that of evaluating one HF sample. Taking a 2D problem as an example, if  $m$  is set to 5 and cost ratio  $\theta$  set to 20, then the total budget to build a surrogate model is 10 HF samples. Thus, we can use either 10 HF samples, 200 LF samples, or any combinations as shown in Table 2.

**Table 2** Combination of HF and LF samples (for a 2D problem)

Total budget	Cost ratio	Combination
10 HF	20	4HF/120LF, 6HF/80HF, 8HF/40LF,
	10	4HF/60LF, 6HF/40HF, 8HF/20LF
	5	4HF/30LF, 6HF/20HF, 8HF/10LF

### 3.2 Performance criteria

To measure the performance of surrogate models, the coefficient of determination  $R^2$ , a global performance metric, is selected, and the formula of  $R^2$  is given in Eq. (14):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{14}$$

where  $n$  is the number of testing samples;  $y_i$  and  $\hat{y}_i$  represent true responses and predictions, respectively, at the testing points;  $\bar{y}$  is the mean of the true responses. The surrogate model is more accurate if  $R^2$  is closer to 1.

The Pearson correlation coefficient (PCC), also referred to as Pearson's  $r$ , is a measure of the correlation between two random variables. In this paper, we adopt the squared Pearson's  $r$ , denoted as  $r^2$ , to represent the correlation between HF and LF functions, which is inspired by Toal (2015), as shown in Eq. (15):

$$r^2 = \left( \frac{\sum_{i=1}^n (y_{hi} - \bar{y}_h)(y_{li} - \bar{y}_l)}{\sqrt{\sum_{i=1}^n (y_{hi} - \bar{y}_h)^2} \sqrt{\sum_{i=1}^n (y_{li} - \bar{y}_l)^2}} \right)^2 \tag{15}$$

where  $y_h$  and  $y_l$  are a set of  $n$  observations, respectively, of the HF and LF functions for identical inputs;  $\bar{y}_h$  and  $\bar{y}_l$  represent the mean of  $y_h$  and  $y_l$ , respectively. The correlation is in proportion to the value of  $r^2$ , which ranges from 0 to 1.

### 3.3 Test function 1

The HF function (Eq. (16)) is a 2D test function derived from Cai et al. (2017). Rather than having a single LF response, we consider a range of different LF responses given by Eq. (17).

HF function:

$$y_h(\mathbf{x}) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4 \tag{16}$$

LF function:

$$y_l(\mathbf{x}) = Ay_h(0.85\mathbf{x}) + x_1x_2 - 65 \tag{17}$$

where  $x_1, x_2 \in [-2, 2]$ , the parameter  $A$  varies from 0 to 1 and effectively controls the correlation of HF and LF responses.

A total computational budget of  $5n$  ( $n = 2$ ) HF samples are used to construct single-fidelity and MFS models. Here, 80% of the total budget is used to generate the HF samples and the remaining 20% budget to generate the LF samples, and the cost ratio is set to 10 so that the number of HF and LF samples are 8 and 20, respectively. For the three single-fidelity models, the total budget of the training samples is the same as that for constructing MFS models. Therefore,

10 HF samples are used to construct the three single-fidelity models. In addition, another  $1000n$  testing samples from the HF function are used for the validation of the single-fidelity and MFS models. All the samples are generated by LHS. To eliminate the effect of random sampling plan on the performance of surrogate models, all the results are averaged over 30 random DOEs. To alleviate the outliers in 30 DOEs affecting the results, in this work,  $R^2$  was set to 0 if  $R^2 \leq 0$ . Setting  $R^2 = 0$  if  $R^2 \leq 0$  has two merits. Firstly,  $R^2 < 0$  and  $R^2 = 0$  all represent that the surrogate model cannot capture the relationship between design variables and responses. Secondly, setting  $R^2 = 0$  if  $R^2 \leq 0$  will avoid the large negative value deteriorating the averaged results.

The reason why  $5n$  HF samples are selected as the total budget is stated as follows: A rule of thumb for choosing the number of training samples to construct a single-fidelity model is  $10n$  (Jones et al. 1998; Forrester and Keane 2009). The purpose of constructing MFS models is to achieve the desired accuracy at a lower cost. Thus, it is reasonable that the total budget of building an MFS model is no more than  $10n$  HF samples. In addition, it holds little promise of improving the performance of MFS models using too few samples. Thus,  $5n$  HF samples are chosen to build surrogate models.

The effect of the correlation between HF and LF function on the performance of the MFS-MLS model is studied by Figs. 2 and 3. Figure 2 compares the performance of the MFS-MLS model with that of three single-fidelity models (i.e., MLS, PRS, and RBF). The three single-fidelity and the MFS-MLS model are constructed by 10 HF samples and validated by the  $1000n$  ( $n = 2$ ) HF testing samples. In Fig. 2, the upper and lower subplot show the mean and standard deviation of  $R^2$  over 30 DOE. In the upper subplot,

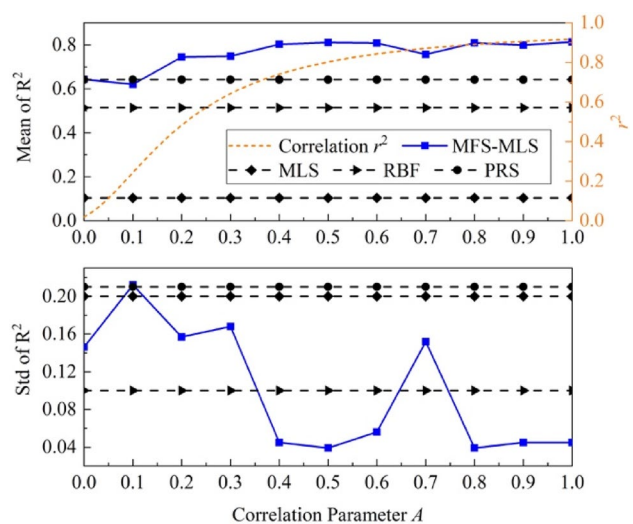


Fig. 2 Comparison between MFS-MLS and single-fidelity surrogate models for test function 1

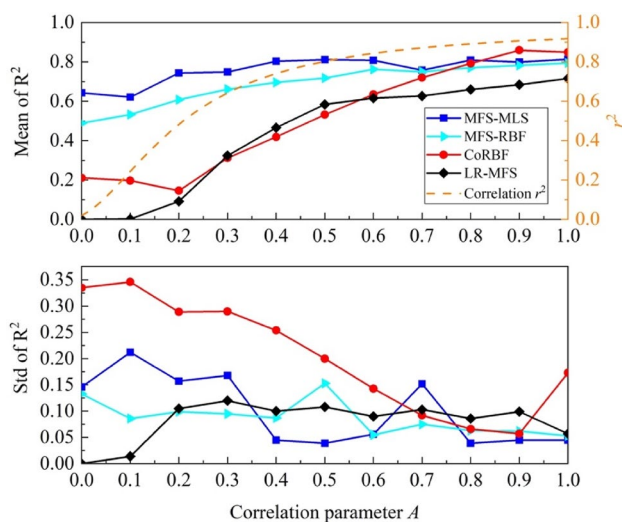


Fig. 3 Comparison of the MFS models for test function 1

the left-hand y-axis represents the prediction accuracy by the coefficient of determination  $R^2$ , the right-hand y-axis represents the correlation  $r^2$  between HF and LF functions. The x-axis represents the parameter  $A$  which controls the correlation  $r^2$ . As it can be seen from Eqs. (16–17), the LF function is a variant of the HF function and the correlation between the HF and LF function will be changed by varying the parameter  $A$ . To investigate the effect of the correlation between HF and LF function on the performance of the MFS-MLS model, parameter  $A$  is chosen from 0 to 1 spaced with 0.1, which means that total 11 LF functions and one HF function are used to form 11 pairs of HF and LF functions with diverse correlations. The red dashed line in Fig. 2 shows the relationship between the correlation  $r^2$  and the parameter  $A$  for test function 1. It is observed that the correlation  $r^2$  is monotonically increased from 0.02 to 0.92 as  $A$  increases, and the tendency of the prediction accuracy of the MFS-MLS, in general, matches the tendency of the correlation  $r^2$ . The performance of the MFS-MLS is much better than that of the three single-fidelity surrogate models when the correlation  $r^2$  is greater than 0.2.

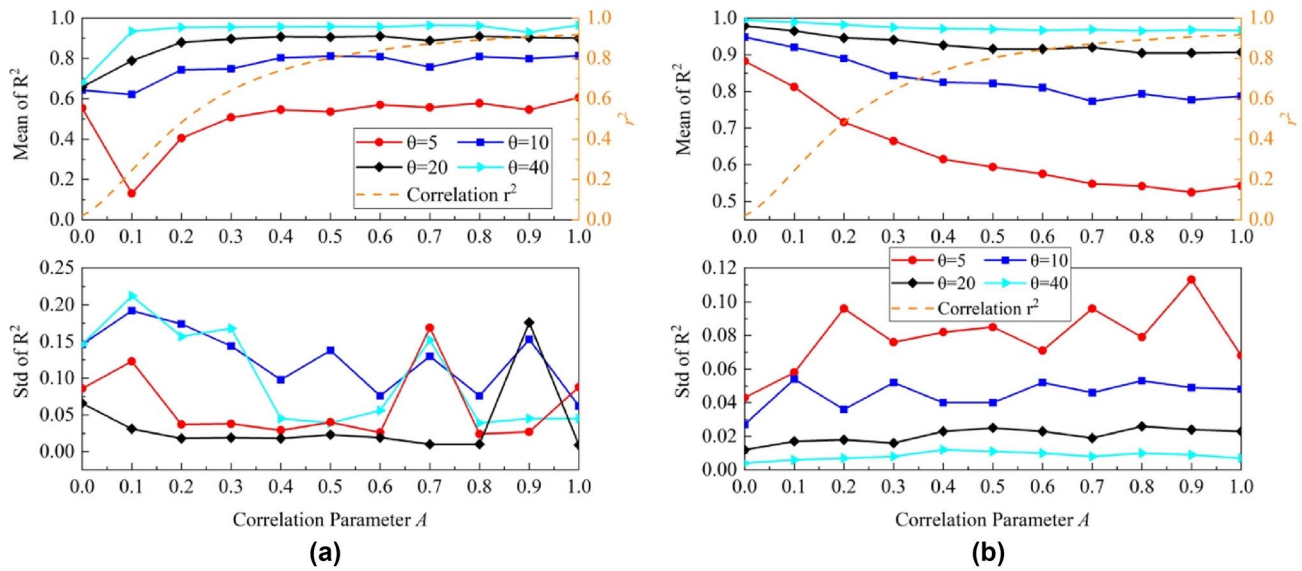
Figure 3 compares the performance of the MFS-MLS model with those of the LR-MFS, CoRBF, and MFS-RBF models in terms of  $R^2$  for test function 1. The four MFS models, MFS-MLS, LR-MFS, CoRBF, and MFS-RBF, are trained on the same training samples (i.e., 8 HF and 20LF samples) and validated on the same 1000n HF testing samples. It is shown that the MFS-MLS model performs much better than other MFS models when  $r^2$ , the correlation between HF and LF functions, is less than 0.9. When  $r^2 \geq 0.9$ , the performance of MFS-MLS is slightly worse than that of CoRBF, but still better than those of MFS-RBF and LR-MFS. It is worth noting that the tendency of the performance of all of the four MFS models is consistent with the

tendency of correlation  $r^2$ , while the MFS-MLS is less sensitive to the correlation. This is caused by the fact that, for the MFS-MLS model, only the HF samples within the influence domain were used to compute the loss, the HF samples far away from the evaluation point have less or even no influence on the prediction. An MFS model is more robust if its performance is less sensitive to the correlation between HF and LF models. This would be useful when dealing with practical engineering problems because the correlation of HF and LF models in the practical engineering problems may be inaccurate or unknown due to the scarce HF samples.

The effect of the cost ratio of HF to LF samples on the performance of MFS-MLS models is studied in Fig. 4. Assuming the total budget is the cost of  $5n$  HF samples, 80% of the total budget is used to generate the HF samples, the remaining 20% budget is used to generate the LF samples for building MFS-MLS models. The cost ratios are set to 5, 10, 20, and 40, respectively. The total budget and the number of the HF samples are fixed, the number of the LF samples is determined by the cost ratio. Then the number of training samples for the MFS-MLS model is shown in Table 3.

Figure 4 illustrates the effect of cost ratios on the performance of the MFS-MLS model under different correlations between HF and LF function. The correlation can be adjusted by parameter  $A$  of the LF function. Parameter  $A$  is chosen from 0 to 1 spaced with 0.1, hence, the correlation has 11 different values. It can be observed from Fig. 4a, when parameter  $A$  is fixed, that the correlation  $r^2$  will be constant, the performance of the MFS-MLS model gets improved as the cost ratio increases. For example, if  $A = 0.4$ , then the correlation  $r^2$  is 0.75, the  $R^2$  of the MFS-MLS models are 0.55, 0.80, 0.91, and 0.96, respectively, when cost ratios are 5, 10, 20, and 40. When the cost ratio increases, more LF samples are used to construct the LF model so that the LF model becomes more accurate, see Fig. 4b, which leads to a better performance of the MFS-MLS model. When  $r^2 > 0.5$ , with correlation increases, the performance of the MFS-MLS model reaches a plateau due to the decreasing performance of LF surrogate models. This also implies that the discrepancy function becomes easier to fit as the correlation increases. From Fig. 4a and b, we can also see that the performance of the MFS-MLS will be enhanced as that of the LF surrogate model gets improved if the correlation is not at extreme low values. Overall, both the cost ratio of HF and LF models and the performance of LF surrogate models have an important impact on the performance of MFS-MLS.

Assuming the cost ratio  $\theta$  is 10, three different combinations, under the total budget of  $5n$  HF samples, are used to investigate the effect of combinations of HF and LF samples on the performance of the MFS-MLS model. These three combinations are “ $4n-10n$ ”, “ $3n-20n$ ”, “ $2n-30n$ ”, which means that 80%, 60%, and 40% of the total budget are used to generate the HF samples, respectively; 20%, 40%,



**Fig. 4** Effect of the cost ratio on the performance of **a** the MFS-MLS model and **b** the LF model

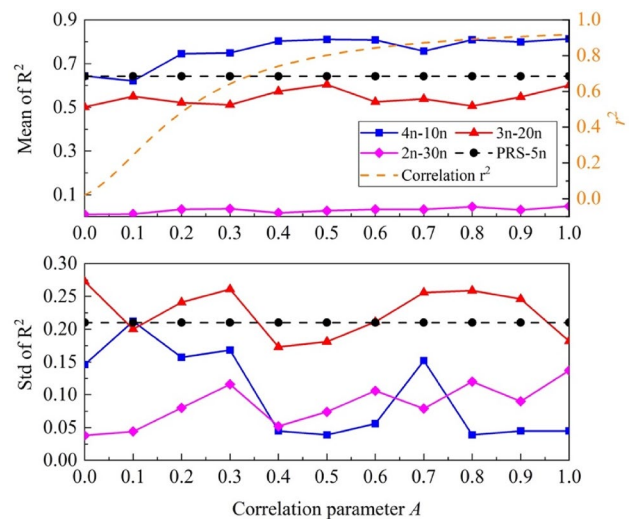
**Table 3** The number of training samples for the MFS-MLS model

Total budget	Cost ratio	No. HF	No. LF
10 HF	5	8	10
	10	8	20
	20	8	40
	40	8	80

and 60% of the total budget are used, correspondingly, to generate the LF samples. Taking “4n–10n” as an example, “4n–10n” means 4n HF samples, and 10n LF samples are used to build an MFS-MLS model. From Fig. 5, it can be observed that for these three combinations, the MFS-MLS model in the case of “4n–10n” performs better than the other two cases. In addition, only the case of “4n–10n” performs better than the single-fidelity PRS model, which exhibits the best performance among the above-mentioned three single-fidelity models, in terms of both the mean and standard deviation of  $R^2$ . Therefore, it is suggested that approximately 80% of the total budget should be allocated to generate the HF samples when using MFS-MLS models.

### 3.4 Test function 2

The test function 2 is directly derived from Toal (2015), in which the HF function (Eq. (18)) is the “Trid function” of ten variables. It is often used to test the performance of surrogate models on high-dimensional problems. It is worth noting that the allocations of HF and LF samples for test



**Fig. 5** Effect of the combination of HF and LF samples on the performance of MFS-MLS for test function 1

function 2 are identical to that of test function 1 throughout the follow-up numerical experiments.

HF function:

$$y_h = \sum_{i=1}^{10} (x_i - 1)^2 - \sum_{i=2}^{10} x_i x_{i-1} \tag{18}$$

LF function:

$$y_l = \sum_{i=1}^{10} (x_i - A)^2 - (A - 0.65) \sum_{i=2}^{10} i x_i x_{i-1} \tag{19}$$



where  $x_i \in [-100,100], i = 1,2, \dots, 10$ . The parameter  $A$  varies from 0 to 1.

Figure 6 compares the performance of the MFS-MLS model with that of three single-fidelity models (i.e., MLS, PRS, and RBF). For the three single-fidelity models,  $5n$  (i.e., 50) HF samples are used to construct the three single-fidelity models. In addition, another  $1000n$  testing samples from the HF function are used for the validation of the single-fidelity and the MFS-MLS models. All the samples are generated by LHS. It is observed that the performance of MFS-MLS is not very sensitive to the correlation of HF and LF functions in this case. The performance of MFS-MLS is much

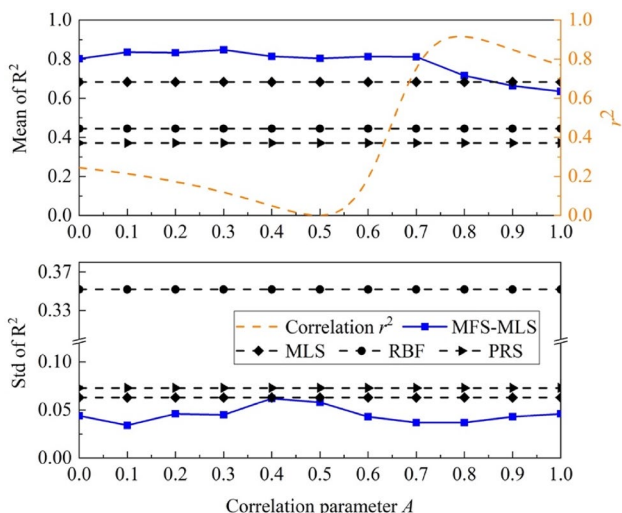
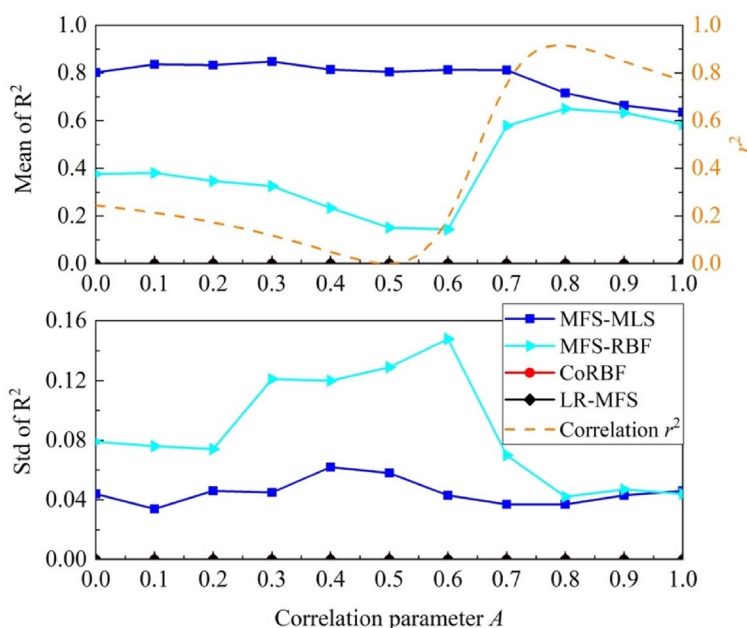


Fig. 6 Comparison between MFS-MLS and single-fidelity surrogate models for the test function 2

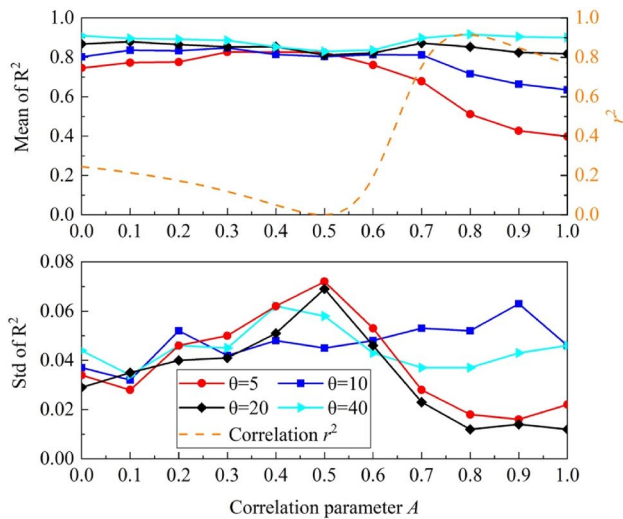
Fig. 7 Comparison of MFS models for test function 2



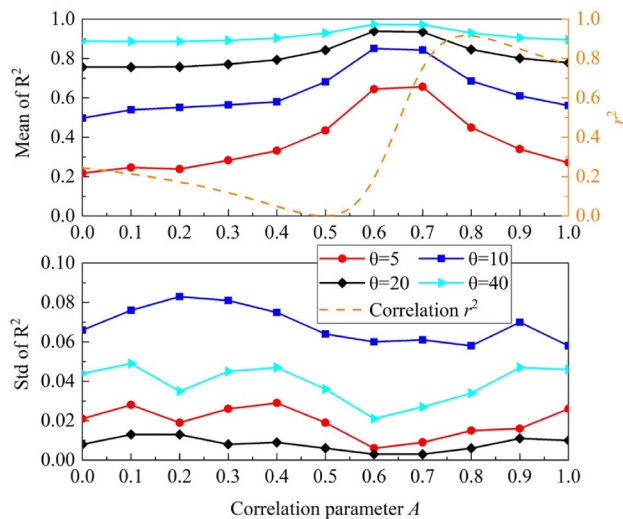
better than that of the three single-fidelity surrogate models in terms of both the mean and standard deviation of  $R^2$  when the correlation parameter  $A$  is no more than 0.8.

Figure 7 compares the performance of MFS-MLS with those of LR-MFS, CoRBF, and MFS-RBF for test function 2. It is shown that the prediction accuracy of both CoRBF and LR-MFS are quite low, which means that the CoRBF and LR-MFS models cannot fit the test function 2 at all. The reason why the LR-MFS model cannot fit the test function 2 is probably caused by the fact that the discrepancy of the HF and LF function in test function 2 is a quadratic function; however, it is fitted by a first-order polynomial in the LR-MFS model. As for the CoRBF model, the model parameter obtained by optimizing leave-one-out error is probably not very stable for high-dimensional problems. The CoRBF model tends not to be good at dealing with high-dimensional problems with small samples available (Durantin et al. 2017). The MFS-MLS model performs much better than the other MFS models in terms of both the mean and standard deviation of  $R^2$ . The tendency of the performance of the MFS-RBF is highly consistent with the tendency of the correlation  $r^2$ . Therefore, the performance of the MFS-RBF is very sensitive to the correlation  $r^2$ . On the contrary, the performance of the MFS-MLS is not sensitive to the correlation for this case, and it is also better than that of the other three MFS models. As a matter of fact, the performance of the MFS-MLS model is still related to the correlation to some extent, which is not evident in Fig. 7, but it can be observed later on in Figs. 8 and 9.

The effect of the cost ratio of HF to LF samples on the performance of MFS-MLS models is studied by Fig. 8. it is assumed that the total budget is the cost of  $5n$  ( $n=10$ ) HF



(a)

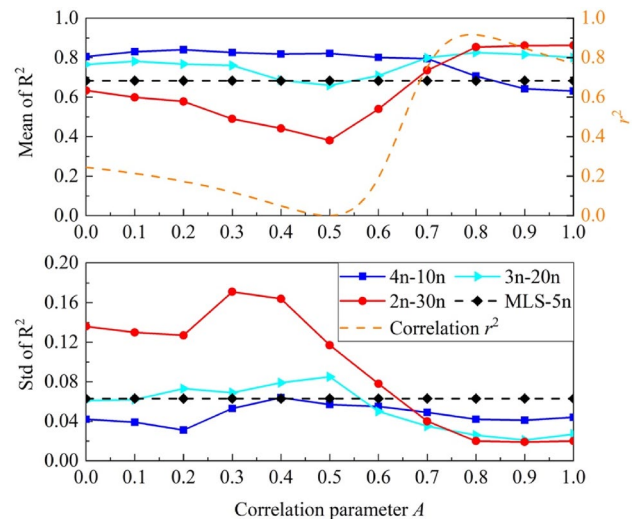


(b)

**Fig. 8** Effect of the cost ratio on the performance of **a** the MFS-MLS model and **b** the LF model

samples, 80% of the total budget is used to generate the HF samples, the remaining 20% budget is used to generate the LF samples for building MFS-MLS models. The cost ratios are set to 5, 10, 20, 40, respectively. The total budget and the number of the HF samples are fixed, the number of the LF samples is determined by the cost ratio. Then the number of training samples for the MFS-MLS model is shown in Table 4. The cost ratio determines the number of the LF samples and the LF model are constructed by the LF samples so that the cost ratio will affect the performance of the LF model.

Figure 8 illustrates the effect of different cost ratios on the performance of the MFS-MLS model. From Fig. 8a, it can be observed that the performance of the MFS-MLS model is getting better as the cost ratio gets increased when



**Fig. 9** Effect of the combination of HF and LF samples on the performance of MFS-MLS for test function 2

**Table 4** The number of training samples for the MFS-MLS model

Total budget	Cost ratio	No. HF	No. LF
50 HF	5	40	50
	10	40	100
	20	40	200
	40	40	400

$r^2 \geq 0.1$ . It is understandable because the LF model becomes more accurate with more LF samples are added, as shown in Fig. 8(b), which will provide a more accurate prediction trend for the MFS-MLS model. However, when the correlation  $r^2$  remains constant at extremely low values, for instance,  $r^2 \leq 0.1$ , the performance of the MFS-MLS model does not improve as the cost ratio increases. This is because when the correlation is extremely small, the similarity of the landscape of HF and LF function is weak, the LF model cannot provide a useful trend for the MFS-MLS model even if the LF model is accurate enough. It can be observed from Fig. 8b that when  $A=0.4$  or  $0.5$ , i.e.,  $r^2 = 0.05$  or  $0$ , the  $R^2$  of the MFS-MLS models are 0.827, 0.814, 0.854, 0.853 or 0.825, 0.804, 0.811, 0.830 when cost ratios are 5, 10, 20, and 40. when the correlation  $r^2$  takes the extremely small values, the performance of the MFS-MLS model does not improve as the cost ratio increases. Moreover, it can be seen that the performance of the MFS-MLS becomes more consistent with the tendency of the correlation as the cost ratio increases. When the cost ratio is relatively low, the low prediction accuracy of the LF model disturbs the effect of correlation on the performance of the MFS-MLS. Therefore,

both the cost ratio and correlation have an impact on the performance of MFS-MLS.

The effect of different combinations of HF and LF samples on the performance of the MFS-MLS model is investigated. From Fig. 9, it can be observed that when the correlation  $r^2 < 0.8$ , the case of “ $4n-10n$ ”, i.e.,  $4n$  HF samples and  $10n$  LF samples, performs best among the three cases (“ $4n-10n$ ”, “ $3n-20n$ ”, and “ $2n-30n$ ”); the case of “ $3n-20n$ ” performs better than the combination of “ $2n-30n$ ”. However, when the correlation  $r^2 \geq 0.8$ , the combination of “ $2n-30n$ ” performs best. The order of the performance of these three combinations reverses. This phenomenon can be explained by follows: Although the test function 2 is high-dimensional, the landscape of the HF function is not very bumpy, and the nonlinearity of the HF function is also lower than that of the LF function, as we can see from Eqs. (18) and (19). Therefore, when the correlation is large enough (here,  $r^2 \geq 0.8$ ), only a few HF samples are enough to calibrate or enhance the prediction accuracy of the MFS-MLS model as long as the LF model is accurate enough to reflect the real trend of HF function. When  $r^2 < 0.8$ , the performance of the case “ $2n-30n$ ” is unpromising because when the correlation is low, in other words, the similarity of the landscape of HF and LF function is weak,  $2n$  HF samples are not sufficient to fit the discrepancy between the HF and the scaled LF function even though the LF surrogate model is accurate enough.

In Fig. 9, only the MLS model of the three single-fidelity models is shown, because the MLS model performs best among the three single-fidelity models. The three single-fidelity models are the same as the ones in Fig. 6. Overall, the combination of “ $4n-10n$ ” and “ $3n-20n$ ” would be reasonable choices for constructing an MFS-MLS model for this test function, because, for most correlations, they both achieve better performance than the single-fidelity models.

### 3.5 Other benchmark functions

In this section, extra 16 test functions were employed to validate the performance of the MFS-MLS model further. The 16 test functions, comprising different dimensions and various degrees of nonlinearity, are selected from the website <https://www.sfu.ca/~ssurjano/index.html> and listed in “Appendix 1”. For each function,  $5n$  HF samples generated by LHS are employed to construct single-fidelity surrogate models. For MFS models, the number of HF and LF samples are  $4n$  and  $10n$ , respectively, and the cost ratio of HF samples to LF samples is set to 10, therefore, the total budget of constructing an MFS model is the cost of  $5n$  HF samples. To eliminate the effect of random sampling plan on the performance of surrogate models, all the results are averaged over 30 random DOEs. “Appendix 2” lists the comparison results of the MFS models and the three single-fidelity models in terms of the  $R^2$  of the 16 test functions. The best results of

each function are in bold italics. It is shown that the MFS-MLS model performs best among the four MFS and the three single-fidelity models except for the 3rd, 6th and 9th functions. It should be noted that, for the 3rd, 6th and 9th functions, the performance of the MFS-MLS model is just slightly weaker than those of the best.

Figure 10 compares the MFS-MLS model with the other three MFS models and the three single-fidelity models in terms of the prediction accuracy. The red columns represent the mean of prediction accuracy  $R^2$ , the mean is obtained by averaging the values of  $R^2$  over the 30 random DOE and then averaged the 16 test functions. The light blue columns represent the standard deviation (Std) of  $R^2$ . It can be found that all of the four MFS models outperform the three single-fidelity models appreciably in terms of the mean of the prediction accuracy except that the CoRBF model beats the single-fidelity RBF model by a narrow margin. The MFS-MLS has the largest mean  $R^2$  of 0.923, which is better than the other MFS and single-fidelity models with the same cost of  $5n$  HF samples. Moreover, the standard deviation of  $R^2$  of MFS-MLS is also smaller than the rest models except the MFS-RBF, which shows the strong robustness of the MFS-MLS model on these 16 test functions.

## 4 Engineering problem

In this section, a static analysis of the boom of a bucket wheel reclaimer (BWR) was used to validate the performance of the proposed model. BWRs, as shown in Fig. 11, are used for moving large amounts of bulk materials, such as coal and ores, in ports, power plants, and stockyards. This is how BWRs work: a heavy load is attached to the short end of the boom served as the balance weights. The bulk materials can then be reclaimed by the rotation of bucket wheel which

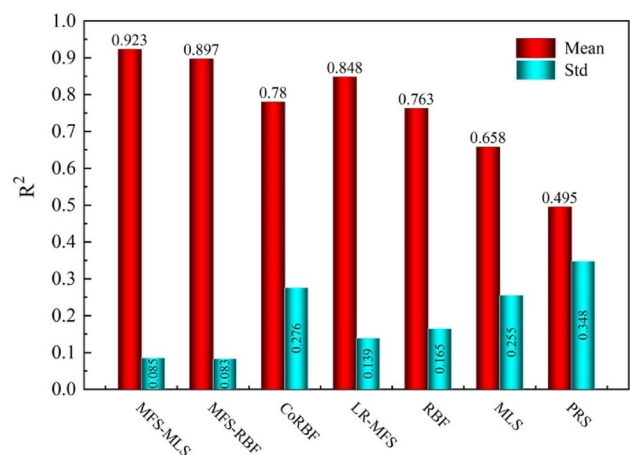


Fig. 10 Comparison of the mean and Std of the  $R^2$  for the 16 test functions



Fig. 11 Bucket wheel reclaimer

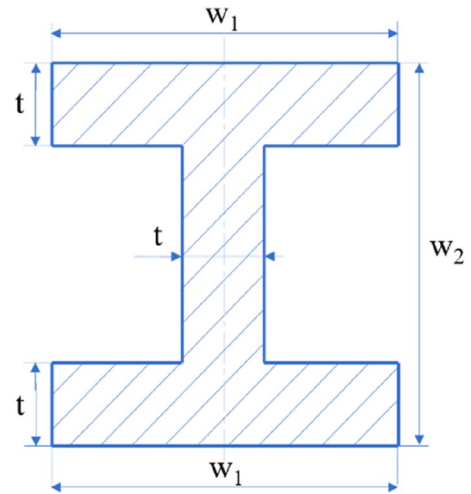


Fig. 13 The cross-section of an I-beam

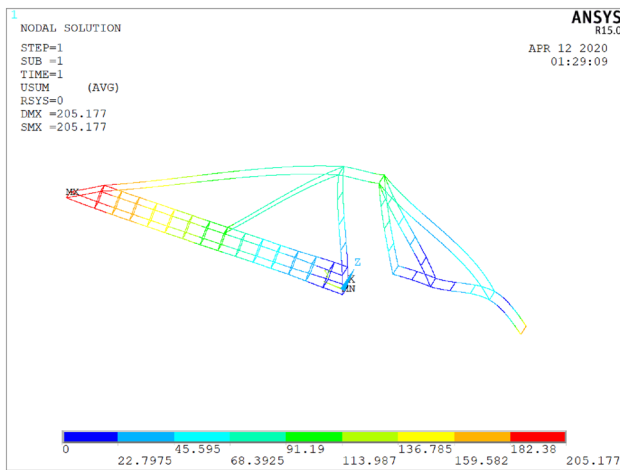


Fig. 12 The deformation cloud map of the boom

is mounted on the long end of the boom on the opposite side. Generally, the boom of a BWR consists of I-beams, and overload can cause deformation, vibration or even failure of the boom. Therefore, the relationship between the maximum deformation (see Fig. 12) and the cross-sectional area of I-beam under different balance weights was investigated. As shown in Fig. 13, the Flange width ( $W_1$ ), beam height

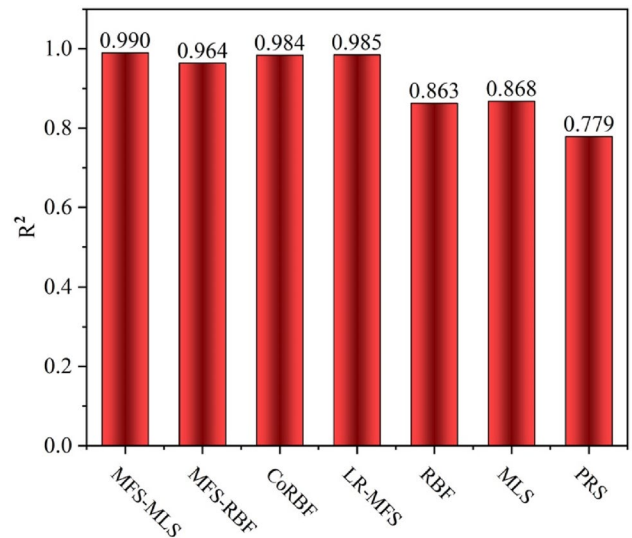


Fig. 14 Comparison of different MFS and single-fidelity models for the engineering problem

( $W_2$ ), the web thickness ( $t$ ), and the balance weight ( $P$ ) are selected as design variables with ranges of 60–70 mm, 100–120 mm, 4–5 mm, and 300–350 kN, respectively. The maximum deformation of the boom is the quantity of interest. The static analysis was conducted on a personal computer with an Intel Core i7 6700 CPU and 32G RAM, using the commercial software ANSYS 17.0. It is assumed that the cutting resistance is constant and the gravity load is considered. The model of the boom built by Timoshenko beam, consisting of 374,000 elements, was used as the HF simulation model, while the model of the boom built by Euler–Bernoulli beam, consisting of 46,750 elements, was used as the LF simulation model. It was found that running one HF simulation takes approximately 71 s, while running one LF simulation takes approximately 13 s; thus, the cost ratio of the HF model to the LF model is approximately 5. The MFS models are constructed by 12 HF samples and 20 LF samples. Three single-fidelity models are constructed by 16 HF samples. In addition, another 20 HF samples were used as testing data to validate the performance of the MFS and single-fidelity models. The comparison of MFS-MLS with the other three MFS models and the three single-fidelity surrogates is shown in Fig. 14. It is observed that the MFS-MLS model exhibits the best results among all the surrogates for this engineering problem.

## 5 Conclusions

A multi-fidelity surrogate model based on moving least squares, called MFS-MLS, was developed in this paper. In the proposed method, the MLS is used to combine the LF model and the discrepancy function to represent the HF responses. Unlike global MFS models, the coefficients of the MFS-MLS model at each prediction site are calculated using the weighted HF samples within the influence domain. Moreover, the size of the influence domain is determined adaptively by a new strategy. The MFS-MLS model was compared with three benchmark MFS models (i.e., MFS-RBF, CoRBF, and LR-MFS) and three popular

single-fidelity surrogate models (RBF, PRS, and MLS) in terms of the prediction accuracy through multiple numerical test functions and an engineering problem. The results show that the MFS-MLS model exhibited competitive performance in both the numerical cases and the practical case. In addition, the effects of key factors (i.e., the correlation of HF and LF samples, the cost ratio of HF to LF samples, and the combination of HF and LF samples) on the performance of the MFS-MLS were investigated using two test functions. The results show that the prediction accuracy of the MFS-MLS model is less sensitive to the correlation of HF and LF samples compared with the other MFS models, which is caused by the fact that, for the MFS-MLS model, the HF samples far away from the evaluation point have less or even no influence on the prediction. The performance of the MFS-MLS model, however, is still getting better as the correlation increases as long as the LF model is accurate. It is also found that under the same total computational budget, the performance of the MFS-MLS model will become better with the increase of the cost ratio. Moreover, it is suggested that 60–80% of the total budget should be allocated to HF samples and this percentage can be increased if the cost ratio is large.

It is worth noting that the MFS-MLS, like other MFS models, cannot be mathematically proved to be a universal approximator. Therefore, the prior information of the engineering problems, such as the dimensionality of the problem and the total computational budget, is encouraged to be considered before applying the MFS-MLS model. When the MFS-MLS model solves high-dimensional problems, the influence domain tends to be so large that all the HF samples are included in it if the quadratic monomials with cross terms are employed. This implies that it needs more and more samples for high-dimensional problems, otherwise the proposed strategy will turn the MFS-MLS model from local to global, which will incur a decline in the performance. In the future, we will focus on strategies to identify the influence domain to locally correct the surrogate with further less HF samples for high-dimensional problems.

## Appendix 1: 16 Test Functions

No.	HF/LF	Test functions	D	S	$r^2$
1	HF	$y_h = (6x - 2)^2 \sin(12x - 4)$	1	$(0,1]^D$	0.58
	LF	$y_l = 0.56y_h + 10(x - 0.5) - 5$			
2	HF	$y_h = \sin(2\pi(x - 0.1)) + x^2$	1	$[0,1]^D$	0.86
	LF	$y_l = \sin(2\pi(x - 0.1))$			
3	HF	$y_h = x \sin(x)/10$	1	$[0,10]^D$	0.73
	LF	$y_l = x \sin(x)/10 + x/10$			
4	HF	$y_h = \cos(3.5\pi x) \exp(-1.4x)$	1	$[0,1]^D$	0.75
	LF	$y_l = \cos(3.5\pi x) \exp(-1.4x) + 0.75x^2$			
5	HF	$y_h = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	2	$[-2,2]^D$	0.77
	LF	$y_l = 2x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + 0.5x_1x_2 - 4x_2^2 + 2x_2^4$			
6	HF	$y_h = \left[ x_2 - 1.275 \left( \frac{x_1}{\pi} \right)^2 + 5 \frac{x_1}{\pi} - 6 \right]^2 + 10 \left( 1 - \frac{1}{8\pi} \right) \cos(x_1)$	2	$x_1 \in [-5,10]$ $x_2 \in [0,15]$	0.98
	LF	$y_l = \frac{1}{2} \left[ x_2 - 1.275 \left( \frac{x_1}{\pi} \right)^2 + 5 \frac{x_1}{\pi} - 6 \right]^2 + 10 \left( 1 - \frac{1}{8\pi} \right) \cos(x_1)$			
7	HF	$y_h = [1 - 2x_1 + 0.05 \sin(4\pi x_2 - x_1)]^2 + [x_2 - 0.5 \sin(2\pi x_1)]^2$	2	$[0,1]^D$	0.85
	LF	$y_l = [1 - 2x_1 + 0.05 \sin(4\pi x_2 - x_1)]^2 + 4[x_2 - 0.5 \sin(2\pi x_1)]^2$			
8	HF	$y_h = \sum_{i=1}^2 x_i^4 - 16x_i^2 + 5x_i$	2	$[-3,4]^D$	0.83
	LF	$y_l = \sum_{i=1}^2 x_i^4 - 16x_i^2$			
9	HF	$y_h = \frac{1}{6} [(30 + 5x_1 \sin(5x_1)) (4 + \exp(-5x_2)) - 100]$	2	$[0,1]^D$	0.88
	LF	$y_l = \frac{1}{6} [(30 + 5x_1 \sin(5x_1)) (4 + \frac{2}{5} \exp(-5x_2)) - 100]$			
10	HF	$y_h = \cos(x_1 + x_2) \exp(x_1 x_2)$	2	$[0,1]^D$	0.86
	LF	$y_l = \cos[0.6(x_1 + x_2)] \exp(0.6x_1 x_2)$			
11	HF	$y_h = \sum_{i=1}^3 0.3 \sin \left( \frac{16}{15} x_i - 1 \right) + \left[ \sin \left( \frac{16}{15} x_i - 1 \right) \right]^2$	3	$[-1,1]^D$	0.40
	LF	$y_l = \sum_{i=1}^3 0.3 \sin \left( \frac{16}{15} x_i - 1 \right) + 0.2 \left[ \sin \left( \frac{16}{15} x_i - 1 \right) \right]^2$			
12	HF	$y_h = (x_1 - 1)^2 + (x_1 - x_2)^2 + x_2 x_3 + 0.5$	3	$[0,1]^D$	0.69
	LF	$y_l = 0.2y_h - 0.5x_1 - 0.2x_1 x_2 - 0.1$			
13	HF	$y_h = \sum_{i=1}^5 \left[ 100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2 \right]$	6	$[0,1]^D$	0.54
	LF	$y_l = \sum_{i=1}^5 \left[ 100(x_i^2 - 4x_{i+1})^2 + (x_i - 1)^2 \right]$			
14	HF	$y_h = \sum_{i=1}^8 x_i^4 - 16x_i^2 + 5x_i$	8	$[-3,3]^D$	0.74
	LF	$y_l = \sum_{i=1}^8 0.3x_i^4 - 16x_i^2 + 5x_i$			
15	HF	$y_h = \sum_{i=1}^2 \left[ (x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4 \right]$	8	$[0,1]^D$	0.66
	LF	$y_l = \sum_{i=1}^2 \left[ (x_{4i-3} + 10x_{4i-2})^2 + 125(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4 \right]$			
16	HF	$y_h = \sum_{i=1}^{10} \exp(x_i) \left[ A(i) + x_i - \ln \left( \sum_{k=1}^{10} \exp(x_k) \right) \right]$ $A = [-6.089, -17.164, -34.054, -5.914, -24.721, -14.986, -24.100, -10.708, -26.662, -22.662, -22.179]$	10	$[-2,3]^D$	0.94
	LF	$y_l = \sum_{i=1}^{10} \exp(x_i) \left[ B(i) + x_i - \ln \left( \sum_{k=1}^{10} \exp(x_k) \right) \right]$ $B = [-10, -10, -20, -10, -20, -20, -20, -10, -20, -20]$			

Notes: D stands for the dimension of functions, S is the design space, and  $r^2$  represents the correlation of HF and LF functions

## Appendix 2: Results of 16 test functions

Function	MFS-MLS	MFS-RBF	CoRBF	LR-MFS	RBF	MLS	PRS
1	<b>0.996 ± 0.000</b>	0.969 ± 0.052	0.987 ± 0.070	0.996 ± 0.000	0.756 ± 0.192	0.812 ± 0.157	0.250 ± 0.084
2	<b>0.999 ± 0.010</b>	<b>0.999 ± 0.000</b>	0.981 ± 0.012	0.975 ± 0.010	0.933 ± 0.048	0.956 ± 0.028	0.334 ± 0.093
3	0.994 ± 0.000	0.845 ± 0.372	<b>0.999 ± 0.000</b>	0.994 ± 0.000	0.572 ± 0.251	0.734 ± 0.230	0.064 ± 0.047
4	<b>0.999 ± 0.000</b>	0.982 ± 0.044	0.973 ± 0.024	0.966 ± 0.022	0.863 ± 0.168	0.784 ± 0.135	0.023 ± 0.015
5	<b>0.834 ± 0.057</b>	0.810 ± 0.045	0.677 ± 0.149	0.513 ± 0.079	0.515 ± 0.096	0.383 ± 0.206	0.666 ± 0.135
6	0.874 ± 0.047	<b>0.883 ± 0.046</b>	0.441 ± 0.406	0.876 ± 0.048	0.625 ± 0.152	0.523 ± 0.245	0.515 ± 0.165
7	<b>0.960 ± 0.015</b>	0.931 ± 0.038	0.942 ± 0.102	0.752 ± 0.065	0.815 ± 0.100	0.502 ± 0.253	0.574 ± 0.161
8	<b>0.817 ± 0.051</b>	0.811 ± 0.053	0.738 ± 0.332	0.817 ± 0.051	0.529 ± 0.109	0.259 ± 0.210	0.254 ± 0.120
9	0.963 ± 0.026	0.968 ± 0.012	<b>0.974 ± 0.031</b>	0.930 ± 0.017	0.899 ± 0.061	0.693 ± 0.270	0.916 ± 0.046
10	<b>0.990 ± 0.010</b>	0.970 ± 0.028	0.965 ± 0.055	0.892 ± 0.034	0.941 ± 0.058	0.935 ± 0.098	0.954 ± 0.148
11	<b>0.928 ± 0.018</b>	0.848 ± 0.063	0.771 ± 0.082	0.774 ± 0.063	0.911 ± 0.030	0.457 ± 0.070	0.759 ± 0.111
12	<b>0.985 ± 0.005</b>	0.962 ± 0.014	0.949 ± 0.070	0.858 ± 0.267	0.898 ± 0.045	0.984 ± 0.014	0.975 ± 0.046
13	<b>0.720 ± 0.098</b>	0.708 ± 0.050	0.427 ± 0.196	0.582 ± 0.064	0.505 ± 0.089	0.254 ± 0.197	0.483 ± 0.242
14	<b>0.835 ± 0.019</b>	0.820 ± 0.031	0.025 ± 0.018	0.824 ± 0.029	0.688 ± 0.069	0.452 ± 0.117	0.042 ± 0.000
15	<b>0.975 ± 0.005</b>	0.945 ± 0.011	0.920 ± 0.015	0.919 ± 0.013	0.932 ± 0.015	0.975 ± 0.006	0.958 ± 0.096
16	<b>0.896 ± 0.014</b>	0.893 ± 0.012	0.711 ± 0.013	0.892 ± 0.012	0.831 ± 0.021	0.832 ± 0.026	0.147 ± 0.071

The figure before “±” represents the mean of  $R^2$ ; the figure after “±” represents the standard deviation of  $R^2$

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00158-021-03044-5>.

**Funding** This research is financially supported by the National Key Research and Development Program of China (Grant No. 2018YFB1700704).

### Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Replication of results** The main codes and raw data are submitted as supplementary materials.

### References

- Absi GN, Mahadevan S (2019) Simulation and sensor optimization for multifidelity dynamics model calibration. *AIAA J* 58(2):879–888
- Alexandrov NM, Lewis RM, Gumbert CR, Green LL, Newman PA (2001) Approximation and model management in aerodynamic optimization with variable-fidelity models. *J Aircr* 38(6):1093–1101
- Bai J, Meng G, Zuo W (2019) Rollover crashworthiness analysis and optimization of bus frame for conceptual design. *J Mech Sci Technol* 33(7):3363–3373
- Basheer IA, Hajmeer M (2000) Artificial neural networks: fundamentals, computing, design, and application. *J Microbiol Methods* 43(1):3–31
- Batra R, Pilonia G, Uberuaga BP, Ramprasad R (2019) Multifidelity information fusion with machine learning: a case study of dopant formation energies in Hafnia. *ACS Appl Mater Interfaces* 11(28):24906–24918
- Berci M, Toropov VV, Hewson RW, Gaskell PH (2014) Multidisciplinary multifidelity optimisation of a flexible wing aerofoil with reference to a small UAV. *Struct Multidisc Optim* 50(4):683–699
- Breitkopf P, Naceur H, Rassineux A, Villon P (2005) Moving least squares response surface approximation: formulation and metal forming applications. *Comput Struct* 83(17–18):1411–1428
- Cai X, Qiu H, Gao L, Shao X (2017) Metamodeling for high dimensional design problems by multi-fidelity simulations. *Struct Multidisc Optim* 56(1):151–166
- Clarke SM, Griebisch JH, Simpson TW (2005) Analysis of support vector regression for approximation of complex engineering analyses. *J Mech Des* 127(6):1077–1087
- Durantin C, Rouxel J, Désidéri JA, Glière A (2017) Multifidelity surrogate modeling based on radial basis functions. *Struct Multidisc Optim* 56(5):1061–1075
- Eldred M, Giunta A, Collis S (2004) Second-order corrections for surrogate-based optimization with model hierarchies. In 10th AIAA/ISSMO multidisciplinary analysis and optimization conference.
- Eldred MS, Giunta AA, Collis SS (2004) Second-order corrections for surrogate-based optimization with model hierarchies. *Collect Tech Pap - 10th AIAA/ISSMO Multidiscip Anal Optim Conf* 3:1754–1768.
- Fang H, Horstemeyer MF (2006) Global response approximation with radial basis functions. *Eng Optim* 38(04):407–424
- Fernández-Godino MG, Park C, Kim NH, Haftka RT (2016) Review of multi-fidelity models. <http://arxiv.org/abs/1609.07196>
- Fernández-Godino MG, Park C, Kim NH, Haftka RT (2019) Issues in deciding whether to use multifidelity surrogates. *AIAA J* 57(5):2039–2054
- Forrester AI, Keane AJ (2009) Recent advances in surrogate-based optimization. *Prog Aersp Sci* 45(1–3):50–79
- Forrester AI, Sóbester A, Keane AJ (2007) Multi-fidelity optimization via surrogate modelling. *Proc R Soc A* 463(2088):3251–3269

- Forrester A, Sobester A, Keane A (2008) Engineering design via surrogate modelling: a practical guide. Wiley, Hoboken
- Goel T, Haftka RT, Shyy W, Queipo NV (2007) Ensemble of surrogates. *Struct Multidisc Optim* 33(3):199–216
- Haftka RT (1991) Combining global and local approximations. *AIAA J* 29(9):1523–1525
- Han ZH, Görtz S (2012) Hierarchical kriging model for variable-fidelity surrogate modeling. *AIAA J* 50(9):1885–1896
- Han ZH, Görtz S, Zimmermann R (2013) Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. *Aerosp Sci Technol* 25(1):177–189
- Hao P, Feng S, Zhang K, Li Z, Wang B, Li G (2018) Adaptive gradient-enhanced kriging model for variable-stiffness composite panels using Isogeometric analysis. *Struct Multidisc Optim* 58(1):1–16
- Hao P, Feng S, Li Y, Wang B, Chen H (2020) Adaptive infill sampling criterion for multi-fidelity gradient-enhanced kriging model. *Struct Multidisc Optim* 62:1–21
- Hutchison MG, Unger ER, Mason WH, Grossman B, Haftka RT (1994) Variable-complexity aerodynamic optimization of a high-speed civil transport wing. *J Aircr* 31(1):110–116
- Jin R, Chen W, Simpson TW (2001) Comparative studies of meta-modeling techniques under multiple modelling criteria. *Struct Multidisc Optim* 23(1):1–13
- Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. *J Global Optim* 13(4):455–492
- Kaplan AG, Kaplan YA (2020) Developing of the new models in solar radiation estimation with curve fitting based on moving least-squares approximation. *Renewable Energy* 146:2462–2471
- Kleijnen JP (2008) Response surface methodology for constrained simulation optimization: an overview. *Simul Model Pract Theory* 16(1):50–64
- Krishnamurthy T (2005) Comparison of response surface construction methods for derivative estimation using moving least squares, kriging and radial basis functions. In 46th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference
- Lancaster P, Salkauskas K (1981) Surfaces generated by moving least squares methods. *Math Comput* 37(155):141–158
- Lee J, Yong SC (2011) Role of conservative moving least squares methods in reliability based design optimization: a mathematical foundation. *J Mech Design* 133(12):121005
- Li J, Wang H, Kim NH (2012) Doubly weighted moving least squares and its application to structural reliability analysis. *Struct Multidisc Optim* 46(1):69–82
- Liu Y, Collette M (2014) Improving surrogate-assisted variable fidelity multi-objective optimization using a clustering algorithm. *Appl Soft Comput* 24:482–493
- Lü Q, Xiao ZP, Ji J, Zheng J, Shang YQ (2017) Moving least squares method for reliability assessment of rock tunnel excavation considering ground-support interaction. *Comput Geotech* 84:88–100
- Mainini L, Maggione P (2012) A Multifidelity Approach to Aerodynamic Analysis in an Integrated Design Environment. In 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA
- Majdisova Z, Skala V (2017) Radial basis function approximations: comparison and applications. *Appl Math Model* 51:728–743
- Myers RH, Montgomery DC, Anderson-Cook CM (2016) Response surface methodology: process and product optimization using designed experiments. Wiley, Hoboken
- Napier N, Sriraman SA, Tran HT, James KA (2020) An artificial neural network approach for generating high-resolution designs from low-resolution input in topology optimization. *J Mech Des* 142:1
- Park C, Haftka RT, Kim NH (2018) Low-fidelity scale factor improves Bayesian multi-fidelity prediction by reducing bumpiness of discrepancy function. *Struct Multidisc Optim* 58(2):399–414
- Sacks J, Welch WJ, Mitchell TJ, Wynn HP (1989) Design and analysis of computer experiments. *Stat Sci* 4:409–423
- Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14(3):199–222
- Song X, Lv L, Sun W, Zhang J (2019) A radial basis function-based multi-fidelity surrogate model: exploring correlation between high-fidelity and low-fidelity models. *Struct Multidisc Optim* 60(3):965–981
- Sun G, Li G, Zhou S, Xu W, Yang X, Li Q (2011) Multi-fidelity optimization for sheet metal forming process. *Struct Multidisc Optim* 44(1):111–124
- Tao J, Sun G (2019) Application of deep learning based multi-fidelity surrogate model to robust aerodynamic design optimization. *Aerosp Sci Technol* 92:722–737
- Toal DJ (2015) Some considerations regarding the use of multi-fidelity Kriging in the construction of surrogate models. *Struct Multidisc Optim* 51(6):1223–1245
- Viana FA, Simpson TW, Balabanov V, Toropov V (2014) Metamodeling in multidisciplinary design optimization: how far have we really come? *AIAA J* 52(4):670–690
- Viana FA (2010) SURROGATES Toolbox User's Guide, Version 2.1, <http://sites.google.com/site/felipeacviana/surrogatestoolbox>
- Wang B (2015) A local meshless method based on moving least squares and local radial basis functions. *Eng Anal Boundary Elem* 50:395–401
- Wang GG, Shan S (2007) Review of metamodeling techniques in support of engineering design optimization. *J Mech Des* 129(4):370–380
- Zhang Y, Kim NH, Park C, Haftka RT (2018) Multifidelity surrogate based on single linear regression. *AIAA J* 56(12):4944–4952
- Zhou Q, Shao X, Jiang P, Zhou H, Shu L (2015) An adaptive global variable fidelity metamodeling strategy using a support vector regression based scaling function. *Simul Model Pract Theory* 59:18–35
- Zhou Q, Wang Y, Choi SK, Jiang P, Shao X, Hu J (2017) A sequential multi-fidelity metamodeling approach for data regression. *Knowl-Based Syst* 134:199–212

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.