Computer Modeling in
Engineering & Sciences

Tech Science Press

Doi:10.32604/cmes.2025.069515

ARTICLE

A Novel Variable-Fidelity Kriging Surrogate Model Based on Global Optimization for Black-Box Problems

Yi Guan[1], **Pengpeng Zhi**[2,3,*] **and Zhonglai Wang**[1,4,*]

[1]School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China
[2]Yangtze Delta Region Institute (Huzhou), University of Electronic Science and Technology of China, Huzhou, 313000, China
[3]Institute of Electronic and Information Engineering of University of Electronic Science and Technology of China in Guangdong, Dongguan, 523808, China
[4]Institute of Electronics and Information Industry Technology of Kash, Kash, 844000, China

*Corresponding Authors: Pengpeng Zhi. Email: zhipeng17@yeah.net; Zhonglai Wang. Email: wzhonglai@uestc.edu.cn

Received: 25 June 2025; Accepted: 25 August 2025; Published: 30 September 2025

ABSTRACT: Variable-fidelity (VF) surrogate models have received increasing attention in engineering design optimization as they can approximate expensive high-fidelity (HF) simulations with reduced computational power. A key challenge to building a VF model is devising an adaptive model updating strategy that jointly selects additional low-fidelity (LF) and/or HF samples. The additional samples must enhance the model accuracy while maximizing the computational efficiency. We propose ISMA-VFEEI, a global optimization framework that integrates an Improved Slime-Mould Algorithm (ISMA) and a Variable-Fidelity Expected Extension Improvement (VFEEI) learning function to construct a VF surrogate model efficiently. First, A cost-aware VFEEI function guides the adaptive LF/HF sampling by explicitly incorporating evaluation cost and existing sample proximity. Second, ISMA is employed to solve the resulting non-convex optimization problem and identify global optimal infill points for model enhancement. The efficacy of ISMA-VFEEI is demonstrated through six numerical benchmarks and one real-world engineering case study. The engineering case study of a high-speed railway Electric Multiple Unit (EMU), the optimization objective of a sanding device attained a minimum value of 1.546 using only 20 HF evaluations, outperforming all the compared methods.

KEYWORDS: Global optimization; kriging; variable-fidelity model; slime mould algorithm; expected improvement

1 Introduction

Surrogate models, also known as metamodels, are widely used in the field of engineering design optimization. These models can be taken as an alternative to experimental processes or simulation models, which can save significant computational cost. The surrogate models are constructed from a finite amount of simulation data and are represented in a compact form consisting of elementary functions and are, therefore, faster to evaluate [1]. In system modelling, a variety of nonlinear and complex structures exist beyond the linear paradigm, including neural networks (NN), Hammerstein, Wiener [2], parallel cascade [3] and other architectures. These representations can capture nonlinearities and intricate dynamics far more effectively. Neural networks, for instance, offer powerful universal-approximation capabilities, accommodating highly nonlinear and ambiguous relationships. Hammerstein and Wiener models combine static nonlinear blocks with dynamic linear elements, yielding flexible yet tractable frameworks for nonlinear systems. Nonetheless, traditional surrogate models—such as linear or simplified nonlinear surrogates—retain decisive advantages:

low computational cost, parsimonious parameterization, and, in many applications, adequate fidelity. Consequently, surrogates are widely exploited to analyse and optimize simulation-based models whose direct evaluation is prohibitively expensive. Today, surrogate modelling spans applications from multidisciplinary design optimization to the reduction of analysis time and the enhancement of the tractability of complex analysis codes [4].

In order to reduce computational cost, variable-fidelity (VF, also known as multifidelity) surrogate methods leveraging both LF and HF samples are widely studied. Generally, VF surrogate methods are divided into three types. The first is the correction-based method, where a scaling function is used to characterize the ratio of HF and LF response values [5]. The scaling function can be multiplicative [6], additive [7], or hybrid [8]. The second is the spatial mapping-based method, where the LF samples can be transformed to the HF samples with the mapping function for a high-quality surrogate model [9]. The third is the VF Kriging method, where the typical method is the co-Kriging method. In the co-Kriging method, a mapping function is constructed by building a covariance matrix between the LF function and HF function [10]. Compared with Co-Kriging, the Hierarchical Kriging [11] structure can avoid calculating the covariance matrix but has higher accuracy in predicting the trend of the high-confidence function.

Surrogate models are essential in design optimization for significantly reducing the computational expense of evaluating objective functions, thereby playing a critical role in surrogate-based optimization frameworks. Recent research has seen considerable advancements in innovative methods aimed at enhancing the performance of these models. For instance, Huang et al. [12] developed a sequential Co-Kriging optimization method using an Augmented Expected Improvement (AEI) criterion to strategically select both sampling locations and fidelity levels. Similarly, Xiong et al. [13] proposed a model fusion technique based on Bayesian–Gaussian process modeling to improve prediction accuracy. To achieve balanced sampling efficiency across multiple fidelity levels, Zhang et al. [14] introduced a variable-fidelity expected improvement (VF-EI) method that adaptively selects new samples from both low- and high-fidelity sources. Further contributions include a sequential optimization framework by Liu et al. [15], which employs augmented collaborative Kriging to explore the impact of inter-model data correlations on hyperparameter estimation. Han et al. [16] addressed multi-fidelity integration through a variable-fidelity optimization approach that combines a multi-level hierarchical Kriging (MHK) model with the expected improvement criterion. In applied engineering contexts, Jiang et al. [17] proposed a variable-fidelity lower confidence bound (VF-LCB) method and demonstrated its efficacy in the design optimization of micro-aerial vehicle fuselages. For computationally intensive CFD simulations involving adaptive mesh refinement, Serani et al. [18] devised four adaptive sampling strategies founded on stochastic radial basis function (RBF) multi-fidelity metamodels. To tackle high-cost multimodal optimization problems, Yi et al. [19] developed a multi-fidelity RBF surrogate-based optimization framework (MRSO). Expanding the range of surrogate modeling techniques, Shi et al. [20] presented a novel multi-fidelity model based on support vector regression. In reliability analysis, He et al. [21] introduced a sequential optimization strategy incorporating variable-fidelity surrogates for failure assessment. With the aim of optimizing the distribution of high- and low-fidelity samples, Guo et al. [22] proposed a new infilling criterion named Filter-GEI to better balance computational cost and accuracy. Cheng et al. [23] formulated a variable-fidelity constrained lower confidence bound (VF-CLCB) criterion integrated with an adaptive mechanism for identifying elite sample points. For expensive black-box problems, Ruan et al. [24] put forward a variable-fidelity probability of improvement (VF-PI) approach. Liu et al. [25] constructed a multi-fidelity radial basis function model (MMFS) that adaptively determines scaling factors to leverage multi-fidelity information. Li and Dong [26] proposed a modified trust-region assisted variable-fidelity optimization (MTR-VFO) framework, substantially improving optimization efficiency for computationally demanding engineering designs. In the context of aerodynamic shape optimization, Tao et al. [27] established a data-driven framework based on a multi-fidelity convolutional neural network (MFCNN), which incorporates optimal solutions from prior cycles as new high-fidelity samples and uses a low-fidelity infilling strategy guided by maximum minimum Euclidean distance. Zhang et al. [28] introduced an information-theoretic acquisition function that balances information acquisition for the current optimization task and knowledge transfer for future tasks in multi-fidelity black-box optimization. Together, these studies contribute to the ongoing refinement of surrogate-based optimization, with a unified emphasis on improving adaptive updating strategies for surrogate models.

In optimization methods based on surrogate models, another key factor is the selection of the optimization method. During the optimization process, the algorithm is used to optimize the sampling criteria. Different optimization algorithms will ultimately result in different locations of the sampling points. This, in turn, directly affects the performance of global optimization. Jariego Pérez and Garrido Merchán [29] applied the grid search algorithm and the limited memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm to the optimization of the acquisition function, and also extended it to the hyperparameter optimization of surrogate models. Vincent and Jidesh [30] employed evolutionary algorithms such as differential evolution, genetic algorithms, and evolutionary strategies to maximize the expected improvement (EI) and compared the performance of Bayesian optimization. It can be concluded that intelligent optimization algorithms play an important role in maximizing the acquisition function. In this study, we adopt the slime mold algorithm (SMA) [31] as the main optimization method, primarily based on its excellent nature-inspired characteristics and global search ability. Compared with traditional optimization techniques, the slime mold algorithm has a strong global search capability, which can effectively avoid falling into local optima and is particularly suitable for complex, multi-modal, and multi-constrained optimization problems. Since its inception, SMA has been successfully applied to a wide range of engineering problems. Ajiboye et al. [32] optimized the operational strategy of hybrid renewable energy systems using a modified SMA, improving both cost and reliability. In the field of power systems, Zhu et al. [33] extended SMA to handle multi-objective structural optimization problems about a parallel hybrid power system. Singh [34] demonstrated the effectiveness of an enhanced SMA in solving the economic load dispatch (ELD) problem. In structural optimization, Wu et al. [35] applied an improved SMA with Lévy flight to optimize truss structures, achieving effective weight minimization under complex constraints. Devarajah et al. [36] proposed an enhanced slime mould algorithm (ESMA) for identifying the solar cells' parameters for five photovoltaic (PV) models, making two modifications to the original SMA.

Collectively, these studies highlight SMA's versatility and effectiveness in solving complex, real-world engineering problems and demonstrate its potential for further development and hybridization to meet emerging optimization challenges. Therefore, in this study, we employ the SMA as the method for optimizing the acquisition function.

Even though there are many studies on the VF surrogate models, the sample updating strategy and efficiency improvement are still challenges. In this paper, a new VF surrogate modelling method based on the global optimization method is proposed. The contributions of the paper can be summarized as

(1) An advanced variable-fidelity extension expected improvement (VFEEI) function is built to adaptively select new samples from LF and HF functions while considering the cost as well as the sample distance between the LF and HF samples.
(2) A new global optimization framework based on a high-fidelity surrogate model is constructed.
(3) An improved slime mould algorithm (ISMA) is proposed to search for the optimal solution of sample updating.

The rest of the paper is organized as follows. The hierarchical Kriging method is briefly introduced in Section 2. The proposed VF surrogate model based on the ISMA and an advanced variable-fidelity extension expectation improvement (VFEEI) will be elaborated in Section 3. The proposed method will be illustrated and validated by several numerical examples and engineering cases in Section 4. Followed by conclusions in Section 5.

2 A Brief Review of Hierarchical Kriging

The Hierarchical Kriging model [11,37] can be built with two or more layers. Taking the two-layer model as an example, the first Kriging model $\hat{y}_2$ is built on the LF sample dataset (S_2, y_2), which $\hat{y}_2$ is used to describe the global trend. Based on the model $\hat{y}_2$, the HF sample dataset (S_1, y_1) is employed to improve the modelling accuracy. The LF samples are usually obtained from simulations, while HF samples are from physical experiments or higher accurate simulations. But both LF and HF samples should have relatively high simulation accuracy from the aspect of engineering practices. Otherwise, the simulated samples are meaningless.

For the LF data, the primary function is to construct a low-fidelity surrogate model that can approximate the fitting trend of a high-fidelity model at a reduced computational cost. For the HF data, the main function is to build a high-fidelity model capable of fitting high-precision approximations by enhancing the accuracy of the model. Basically, a better balance among accuracy, efficiency, and computational cost can be achieved by combining the LF and HF data during the modelling process.

The corresponding low-fidelity Kriging model needs to be developed first, and the low-fidelity model can be expressed as

$$Y_{lf}(x) = \beta_{0,lf} + Z_{lf}(x) \tag{1}$$

where β is an unknown constant to be determined and $Z_{\mathrm{lf}}(x)$ is a smooth Gaussian stochastic process.

An LF approximation model is built based on LF samples$(S_{\mathrm{lf}}, y_{\mathrm{s,lf}})$. The prediction for an arbitrary untested point $\boldsymbol{x}$ can be written as

$$\hat{y}_{\mathrm{lf}}(x) = \beta_{0,\mathrm{lf}} + \mathbf{r}_{\mathrm{lf}}^{\mathrm{T}}(x)\mathbf{R}_{\mathrm{lf}}^{-1}(y_{\mathrm{s,lf}} - \beta_{0,\mathrm{lf}}\mathbf{E}) \tag{2}$$

where $\beta_{0,\mathrm{lf}} = \left(\mathrm{E}^{\mathrm{T}}\mathrm{R}_{\mathrm{lf}}^{-1}\mathrm{E}\right)^{-1}\mathrm{E}^{\mathrm{T}}\mathrm{R}_{\mathrm{lf}}^{-1}y_{\mathrm{s,lf}}$ after LF samples are implemented, $\mathbf{R}_{\mathrm{lf}}$ is a $n_{\mathrm{lf}} \times n_{\mathrm{lf}}$ dimensional matrix representing the correlation between observations, $n_{\mathrm{lf}}$is the number of low-fidelity samples, $\mathbf{E}$ is a n_{lf} dimensional vector with all elements being 1, and r_{lf} is the correlation vector between the training and prediction samples.

The Hierarchical Kriging model is expressed as

$$Y(x) = \beta_0\hat{y}_{\mathrm{lf}}(x) + Z(x) \tag{3}$$

where β_0 is used as the global trend function and $Z(x)$ is a Gaussian stochastic process with mean 0, whose covariance is expressed as

$$\mathrm{Cov}\left[Z(x), Z(x')\right] = \sigma^2 R(x, x') \tag{4}$$

where σ^2 is the process variance of $Z(x)$ and $R(x, x')$ is the spatial correlation function.

The predicted values of the Hierarchical Kriging model are expressed as [38]

$$\hat{y}(x) = \beta_0\hat{y}_{\mathrm{lf}}(x) + \mathbf{r}^{\mathrm{T}}\mathbf{R}^{-1}(y_{\mathrm{s}} - \beta_0\mathbf{F}) \tag{5}$$

where $\beta_0 = \left(\mathrm{F}^{\mathrm{T}}\mathrm{R}^{-1}\mathrm{F}\right)^{-1}\mathrm{F}^{\mathrm{T}}\mathrm{R}^{-1}y_{\mathrm{s}}$, n_{hf} is the number of HF samples, $\mathbf{F}$ is a n_{hf} dimensional vector with all elements being 1, and $\mathbf{R}$ is the correlation matrix between the HF training points.

3 The Proposed ISMA-VFEEI Method

3.1 The Proposed VFEEI Function

Currently, there are several methods to build the surrogate model from one kind of fidelity samples based on the efficient global optimization (EGO) [38,39] method. In multi-fidelity settings, a key challenge is to select the new samples, specifically which new samples to include and determining their respective fidelity sources. In order to adaptively allocate computational resources to the HF or LF model during iterative optimization based on the existing model information, the new index VFEEI is defined as

$$VFEEI\left(x,l\right)=EEI\left(x,l\right)\times CR\left(l\right)\times d\left(x,l\right) \tag{6}$$

where $EEI\left(x,l\right)$ represents the expanded EI function, $CR\left(l\right)$ represents the cost ratio and $d\left(x,l\right)$ represents the sample distance function.

In Hierarchical Kriging, β_0 is represented as the global trend coefficient, and the variable fidelity prediction error function can be calculated by

$$s\left(x,l\right)=\begin{cases}\beta_0 s_{\mathrm{lf}}\left(x\right), & l=1\\ s\left(x\right), & l=2\end{cases} \tag{7}$$

where l represents the fidelity level, $l=1$ represents the low-fidelity level, $l=2$ represents the high-fidelity level, $s_{\mathrm{lf}}\left(x\right)$represents the prediction standard deviation of the low-fidelity model, and $s\left(x\right)$ represents the prediction standard deviation of the high-fidelity model.

Correspondingly, the expressions of the variable fidelity minimum value $y_{\min}\left(x,l\right)$ and the variable fidelity predicted response value $\hat{y}\left(x,l\right)$ are provided by

$$y_{\min}\left(x,l\right)=\begin{cases}y_{\mathrm{minlf}}\left(x\right), & l=1\\ y_{\min}\left(x\right), & l=2\end{cases} \tag{8}$$

$$\hat{y}\left(x,l\right)=\begin{cases}\hat{y}_{\mathrm{lf}}\left(x\right), & l=1\\ \hat{y}\left(x\right), & l=2\end{cases} \tag{9}$$

where $y_{\mathrm{minlf}}\left(x\right)$ represents the minimum value of the low-fidelity model, $y_{\min}\left(x\right)$ represents the minimum value of the high-fidelity model, $\hat{y}_{\mathrm{lf}}\left(x\right)$ represents the predicted response value of the low-fidelity model, and $\hat{y}\left(x\right)$ represents the predicted response value of the high-fidelity model.

Then the expression of $EEI\left(x,l\right)$is

$$\mathrm{EEI}\left(x,l\right)=\begin{cases}\left(y_{\min}\left(x,l\right)-\hat{y}\left(x,l\right)\right)\Phi\left(\frac{y_{\min}\left(x,l\right)-\hat{y}\left(x,l\right)}{s\left(x,l\right)}\right) & \\ +s\left(x,l\right)\varphi\left(\frac{y_{\min}\left(x,l\right)-\hat{y}\left(x,l\right)}{s\left(x,l\right)}\right), & s\left(x,l\right)>0\\ 0, & s\left(x,l\right)<0\end{cases} \tag{10}$$

where $CR\left(l\right)$ is defined as the ratio of the computational cost of response evaluation between the HF model and the lth fidelity model. It indicates how many times a lth fidelity model can be evaluated for the same computational cost as evaluating an HF model. Its specific expression is

$$CR\left(l\right)=\begin{cases}C_{\mathrm{h}}, & \text{the HF model}\\ C_l, & \text{the } l\text{th fidelity model}\end{cases} \tag{11}$$

where C_h represents the computational cost of response assessment for the HF model and C_l represents the computational cost of response assessment for the lth fidelity model.

The sample distance function is developed to prevent the occurrence of overfitting due to overly dense samples. This function describes the relative distance between sample points and is expressed as

$$d(x,l) = \min\left|x - x_i^l\right|, i = 1, 2, \cdots, n_l \tag{12}$$

where n_l denotes the number of samples in the lth fidelity model and x_i^l denotes the i-th training sample of the lth fidelity model. As shown in Eq. (12), the sample distance function is defined as the minimum distance between the new sample and the observed samples. When the samples are densely distributed in an observation region, the sample distance function value will increase. Therefore, maximizing the VFEEI can effectively suppress the resampling of samples. When the sample distance is too small or when samples are repeated, the matrix constructed from these samples may become singular, which in turn degrades the prediction performance of the surrogate model.

The new sampling points and fidelity level can be achieved by maximizing the VFEEI function with the following expression

$$x, l = \mathrm{argmin} VFEEI(x,l) \tag{13}$$

3.2 The Improved Slime Mould Algorithm

In the flowchart of the proposed method, the optimization should be conducted for the acquisition of updated samples and the globally optimal solution of the black-box function. Despite the robust performance in various engineering applications, the original slime-mould algorithm (SMA) exhibits several limitations. First, while SMA has a strong global search ability, it may lack sufficient local search precision in the optimization of certain high-dimensional or multi-modal functions. This leads to a decrease in the algorithm's accuracy when approaching the optimal solution. Second, the search behavior of the original SMA is relatively monotonous and lacks an adaptive mechanism, which may result in an unsatisfactory balance between exploration and exploitation in specific optimization problems, thereby affecting the quality of the final solution. Therefore, although the original SMA has strong application potential, the probability of falling into local optimal solutions is relatively high. To address this, an opposition-based learning mechanism combined with Cauchy mutation is introduced to enhance the global search capability.

From the perspective of the slime mold position update mechanism, in the standard SMA, the position update rule of the slime mold is determined by the relationship between r and p. When $r < p$, the position update of the slime mold is determined by the positions of the current optimal individual and two random individuals. In this case, the slime mold exhibits random exploration around the current best position. However, such purposeless random exploration can also slow down the initial convergence speed of SMA. As the number of iterations increases, the slime mold population will converge towards the current best position, making SMA highly vulnerable to falling into local optima when solving functions with multiple local optima. When $r \geq p$, the position update of the slime mold is determined by the convergence factor v_c and the slime mold individual's own position. As the number of iterations increases, v_c linearly converges from 1 to 0, causing the slime mold population to converge towards the origin. This type of position update is not conducive to the optimization of functions whose optimal solutions are not at the origin. When optimizing such functions, the solution accuracy of standard SMA is relatively poor.

To address these issues, we have improved the position update behavior and the weight coefficients, enabling the algorithm to automatically update parameters and weight coefficients based on the number of iterations. This modification allows the algorithm to focus on global search in the early stages of optimization and to shift towards local exploitation in the later stages. Additionally, we have incorporated randomly selected individuals into the position update formula, which facilitates the escape of the best individual from local optima. The formula for updating the position of the slime moulds during the food discovery phase can be expressed by

$$X(t+1) = \begin{cases} X_b(t) + v_b \cdot ((W \cdot X_b(t) - X_A(t)) + (W \cdot X_b(t) - X_B(t))), & r < p \\ v_c \cdot X(t), & r \geq p \end{cases} \tag{14}$$

where $v_b = [-a, a]$, $v_c = e^{-\left(\frac{4t}{t_{\max}}\right)^2}$, $a = 2.5 - \frac{2.5}{1 + e^{-\frac{20t}{2(t_{\max}-8)}}}$, $p = \tanh|S(i) - F_D|$, $i \in 1, 2, \cdots, N$, r is the random value in the interval [0, 1], $X_b(t)$ is the current best position obtained, $X_A(t)$ and $X_B(t)$ are two randomly selected individuals, v_b is the control parameter, v_c is the adaptive adjustable feedback factor, t represents the current number of iterations, $t_{\max}$ represents the maximum number of iterations, r represents the current position, p is the selection switch, $S(i)$ is the population after sorting, F_D is the best value among all iterations. W is the weight

$$W(\text{SmellIndex}(i)) = \begin{cases} 1 + r \cdot \log\left(\frac{F_b - S(i)}{F_b - F_w} + 1\right), & i \leq \frac{N+1}{2} + \frac{N-1}{2}\cos\left(\frac{\pi t}{t_{\max}}\right) \\ 1 - r \cdot \log\left(\frac{F_b - S(i)}{F_b - F_w} + 1\right), & \text{else} \end{cases} \tag{15}$$

$$\text{SmellIndex} = \text{sort}(S) \tag{16}$$

where F_b denotes the best adaptation value during the current iteration, F_w denotes the worst adaptation value during the current iteration, and SmellIndex denotes the ranking of the adaptation values.

The slime moulds also split a portion of their search with random exploration. Then the location of slime moulds can be updated by

$$X(t+1) = \begin{cases} r \cdot (B_u - B_l) + B_l, & r < z & \text{(17a)} \\ X_b(t) + v_b \cdot ((W \cdot X_b(t) - X_A(t)) + (W \cdot X_b(t) - X_B(t))), & r \geq z, r < p & \text{(17b)} \\ v_c \cdot X(t), & r \geq z, r \geq p & \text{(17c)} \end{cases}$$

where B_u and B_l are the upper and lower limits of the search range, respectively, and z is a parameter that determines the proportion of randomly distributed slime individuals to the total slime population, which is taken as 0.3 in this paper.

Finally, the population is updated using opposition-based learning [40] and Cauchy mutation [41] to select individuals with smaller fitness for positional updating. The expressions of the opposition-based learning and the Cauchy mutation are respectively

$$X_1(t+1) = \min(X(t+1)) + \max(X(t+1)) - X(t+1) \tag{18}$$

$$X_2(t+1) = X_b(t) + X_b(t) \cdot \tan(\pi(\text{rand} - 0.5)) \tag{19}$$

where $X_1(t+1)$ is the population update position for reverse learning, $X_2(t+1)$ is the population updating position for the Cauthy mutation, and *rand* is a random number between [0, 1].

The pseudo code of the ISMA is shown in Algorithm 1.

Algorithm 1: Pseudo-code of ISMA

```
Initialize the parameter: Population size $N$, $B_u$ and $B_l$, $t_{\max}$
Initialize the position of the slime mould $X_i$ $(i = 1,2,\ldots,N)$
While $(t < t_{\max})$ do
    Calculate the fitness of all slime mould;
    Update $\boldsymbol{F_D}$, $\boldsymbol{F_b}$, $\boldsymbol{F_w}$, $X_b$
    Calculate the $W$ by Eq. (15)
    For $i = 1$ to $N$ do
        If rand $< z$
            Update position by Eq. (17a)
        Else
            Update $p$, $v_b$, $v_c$
            If $rand \geq z$ and rand $< p$
                Update position by Eq. (17b)
            Else
                Update position by Eq. (17c)
            End if
        End if
    End for
    For $i = 1$ to $N$ do
        Update position by Eq. (18)
        Update position by Eq. (19)
    End for
    $t = t + 1$;
End while
Return $\boldsymbol{F_D}$ $X_b$
```

Compared with the improvements in Reference [36], the difference between this paper and the reference lies in the position update behavior. The method in this paper employs adaptive adjustment of weight coefficients and random positions of individuals to escape from local optima. In contrast, Reference [36] developed an arbitrary average position to enable the algorithm to break free from local optima and explore a broader solution space. Moreover, the ISMA method in this paper takes into account the mutation behavior of the optimal individual to increase the probability of escaping from local optima.

3.3 Details of the Proposed ISMA-VFEEI Method

The aim of the proposed ISMA-VFEEI method is to construct a VF surrogate model by efficiently using the LF and HF samples, where the ISMA is provided to obtain sufficiently accurate optimal results when reducing the computational cost. The flowchart of the proposed ISMA-VFEEI method is shown in Fig. 1.

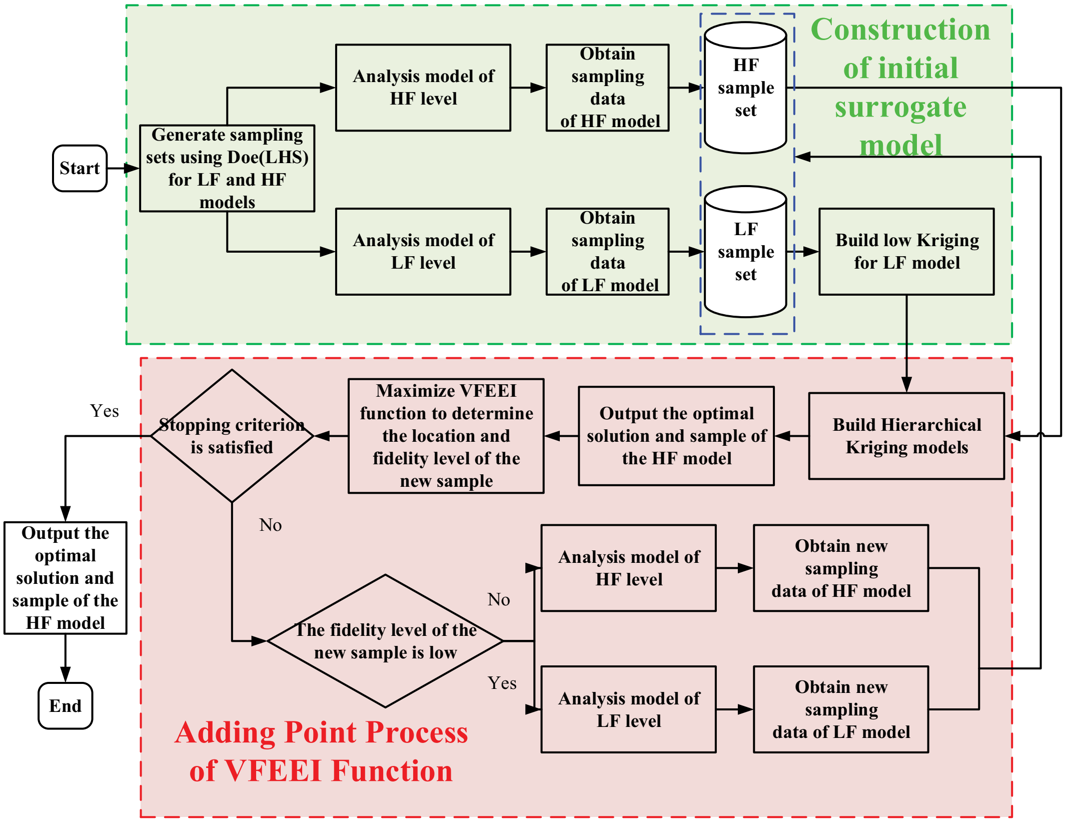

Construction of initial surrogate model
Start
Generate sampling sets using Doe(LHS) for LF and HF models
Analysis model of HF level
Obtain sampling data of HF model
HF sample set
Analysis model of LF level
Obtain sampling data of LF model
LF sample set
Build low Kriging for LF model
Build Hierarchical Kriging models
Output the optimal solution and sample of the HF model
Maximize VFEEI function to determine the location and fidelity level of the new sample
Stopping criterion is satisfied
Yes
No
Output the optimal solution and sample of the HF model
End
The fidelity level of the new sample is low
No
Yes
Analysis model of HF level
Obtain new sampling data of HF model
Analysis model of LF level
Obtain new sampling data of LF model
Adding Point Process of VFEEI Function

Figure 1: Flowchart of the proposed ISMA-VFEEI method

The specific steps of the proposed ISMA-VFEEI method are elaborated as follows.

Step 1: The Design of Experiment (DoE) method is used to generate the initial HF and LF sample points separately. For example, HF samples $x_h = \left\{x_1^h, x_2^h, \cdots, x_{n_h}^h\right\}$ and LF samples $x_l = \left\{x_1^l, x_2^l, \cdots, x_{n_l}^l\right\}$ are generated using Latin Hypercube Sampling (LHS).

Step 2: The design samples are imported into the corresponding models with different levels of fidelity and the response values $f_h = \left\{f_1^h, f_2^h, \cdots, f_{n_h}^h\right\}$ and $f_l = \left\{f_1^l, f_2^l, \cdots, f_{n_l}^l\right\}$ are output.

Step 3: A Hierarchical Kriging model is built based on the current set of samples.

Step 4: The proposed ISMA method is used to obtain the minimum of the high-fidelity model and output the location and response values of the minimum sample.

Step 5: The proposed ISMA method of maximizing the VFEEI function is used to determine the updated point and the corresponding level of fidelity.

Step 6: The termination condition is judged in each iteration. The total computational cost is considered, and the computational cost of the LF samples is converted to the computational cost of the HF samples and the expression for the total cost is

$$\begin{cases} \left|\hat{y}_{\min} - y_{\text{real}}\right| \le \varepsilon \\ C_{\text{total}} \ge C_{\max} \\ J \ge J_{\max} \end{cases} \tag{20}$$

where $\hat{y}_{\min}$ represents the predicted optimal solution in the optimization process and y_{real} is the known optimal solution, ε is the allowed minimum error, $C_{\max}$ is the allowed maximum computational cost and $J_{\max}$ is the maximum number of points allowed to be added. For engineering problems, the maximum computational cost and the maximum number of points to be added are set as termination conditions.

If this condition is met, the algorithm switches to step 9, otherwise, it returns to step 7. We will explicitly state that if the maximum limits for computational cost and the number of samples are reached and the error is still greater than the allowed range, the algorithm will proceed with the final assessment of performance. In this scenario, we recommend that after the calculation is terminated, the sample points of the initial model should be increased and the calculation should be carried out again until the requirements of the algorithm are met.

Step 7: The updated sample is imported into the corresponding models with different levels of fidelity and the response values of the updated samples are outputted.

Step 8: Add the updated sample to the HF or LF sample set and re-establish the Hierarchical Kriging model. Return to **Step 3**.

Step 9: If the termination condition is satisfied, the optimal solution and the corresponding sample position are outputted using the ISMA algorithm.

4 Examples and Discussions

4.1 Validation of ISMA

In order to verify the effectiveness of the ISMA algorithm, six testing functions are employed and compared to the original SMA algorithm. The testing functions from CEC2005 and reference [31]. The specific information of the testing functions is shown in Table 1.

Table 1: Test functions of ISMA

Testing functions	*Dim*	*Range*	*fmin*
$F1 = \sum_{i=1}^{n} -x_i \sin\left(\sqrt{\lvert x_i\rvert}\right)$	10	[−500, 500]	−4189.8
$F2 = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) + 10\right]$	30	[−5.12, 5.12]	0
$F3 = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[−600, 600]	0
$F4 = \sum_{i=1}^{11}\left[a_i - \frac{x_1\left(b_i^2 + b_i x_2\right)}{b_i^2 + b_i x_3 + x_4}\right]^2$	4	[5, 5]	0.000308
$F5 = -\sum_{i=1}^{4} \alpha_i \exp\left(-\sum_{j=1}^{6} A_{ij}\left(x_j - P_{ij}\right)^2\right)$, where $A = \begin{pmatrix} 10 & 3 & 17 & 3.50 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{pmatrix}$ $P = 10^{-4}\begin{pmatrix} 1312 & 1696 & 5569 & 124 & 8283 & 5886 \\ 2329 & 4135 & 8307 & 3736 & 1004 & 9991 \\ 2348 & 1451 & 3522 & 2883 & 3047 & 6650 \\ 4047 & 8828 & 8732 & 5743 & 1091 & 381 \end{pmatrix}$	6	[0, 1]	−3.322
$F6 = -\sum_{i=1}^{m}\left(\sum_{j=1}^{4}\left(x_j - C_{ji}\right)^2 + \beta_i\right)^{-1}$, where			

(Continued)

Table 1 (continued)

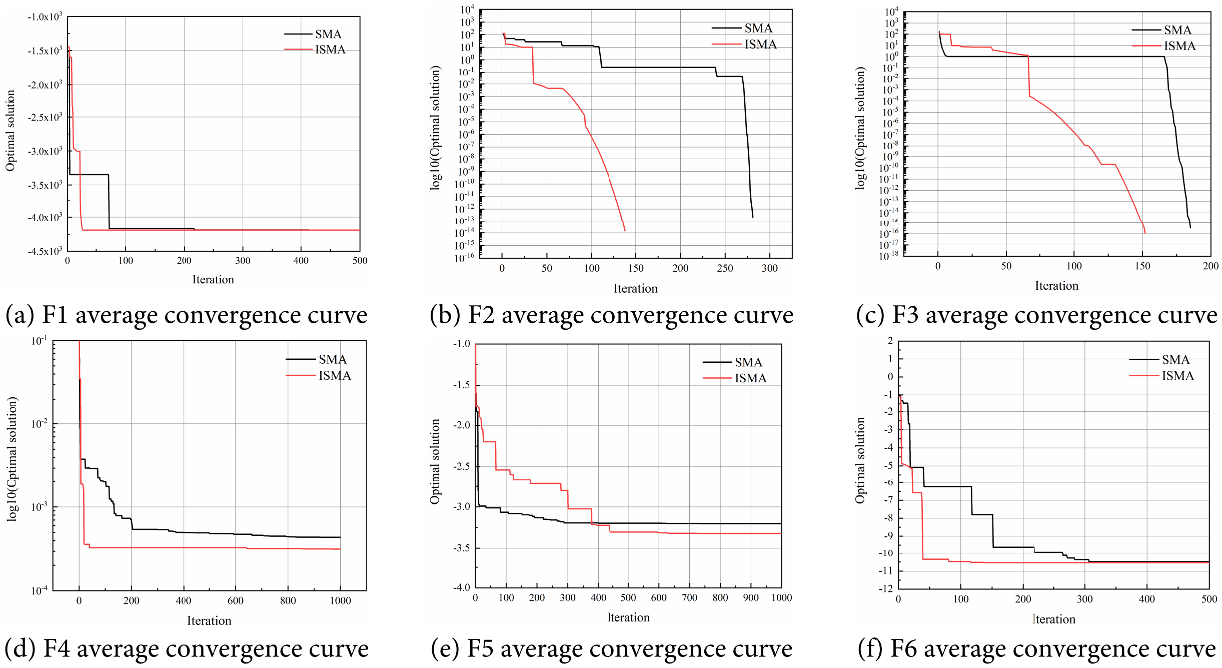
Testing functions	*Dim*	*Range*	*fmin*
$m = 10$ $\beta = 1/10\,(1, 2, 2, 4, 4, 6, 3, 7, 5, 5)^T$ $f_3 = \begin{pmatrix} 4.0 & 1.0 & 8.0 & 6.0 & 3.0 & 2.0 & 5.0 & 8.0 & 6.0 & 7.0 \\ 4.0 & 1.0 & 8.0 & 6.0 & 7.0 & 9.0 & 3.0 & 1.0 & 2.0 & 3.6 \\ 4.0 & 1.0 & 8.0 & 6.0 & 3.0 & 2.0 & 5.0 & 8.0 & 6.0 & 7.0 \\ 4.0 & 1.0 & 8.0 & 6.0 & 7.0 & 9.0 & 3.0 & 1.0 & 2.0 & 3.6 \end{pmatrix}$	4	[0, 10]	−10.5364

For a comparison with other methods, the population size $N = 30$, dimension $Dim = 10$ and maximum number of iterations $t_{\max} = 1000$ are set. 30 times are run independently on each testing function, and the average and standard deviation are provided in Table 2.

Table 2: Comparison of testing function results between the SMA and ISMA

Functions	Indicator	SMA	ISMA
F1	Minimum	−4189.8275	−4189.8236
	Maximum	−4189.7993	−4189.8215
	Average	−4189.8226	−4189.8272
	Standard deviation	3.61E−3	5.02E−4
F2	Minimum	0	0
	Maximum	0	0
	Average	0	0
	Standard deviation	0	0
F3	Minimum	0	0
	Maximum	0	0
	Average	0	0
	Standard deviation	0	0
F4	Minimum	3.31E−4	3.31E−4
	Maximum	4.71E−4	3.57E−4
	Average	3.42E−4	3.31E−4
	Standard deviation	1.71E−4	1.47E−4
F5	Minimum	−3.322	−3.322
	Maximum	−3.2388	−3.2723
	Average	−3.2824	−3.322
	Standard deviation	2.84E−2	2.62E−2
F6	Minimum	−10.5364	−10.5364
	Maximum	−9.0309	−9.1466
	Average	−10.5361	−10.5364
	Standard deviation	2.30E−4	1.98E−4

From Table 2, we can see that ISMA can obtain better results compared to the original SMA algorithm, both from the mean and standard deviation. In order to compare the convergence speed and accuracy of each algorithm more intuitively, the average convergence curves of the testing functions are presented in Fig. 2. From Fig. 2, we can see that the convergence speed of the ISMA is also better than that of the original SMA.



(a) F1 average convergence curve (b) F2 average convergence curve (c) F3 average convergence curve

(d) F4 average convergence curve (e) F5 average convergence curve (f) F6 average convergence curve

Figure 2: Average convergence curve of the test functions

To further evaluate the performance of ISMA, this paper employs the Wilcoxon rank-sum test to examine whether there is a significant difference between the performance of ISMA and SMA. The significance level is set at 0.05, and the optimization results from 30 runs of all algorithms are used as samples. When the p-value of the test is greater than 0.05, it indicates that there is no significant difference between the results of the two algorithms; otherwise, a significant difference exists. Table 3 presents the results of the Wilcoxon rank-sum test between ISMA and the comparison algorithms.

Table 3: Wilcoxon rank-sum test results

Functions	ISMA
F1	1.02E−11
F2	/
F3	/
F4	2.16E−2
F5	5.78E−8
F6	3.81E−5

As shown in Tables 2 and 3, ISMA outperforms ESMA and SMA on four test functions. Therefore, ISMA demonstrates statistically significant superiority in performance. The experimental results indicate that the proposed ISMA in this paper has significantly improved convergence speed, solution accuracy, and robustness.

4.2 An Illustrative Example of ISMA-VFEEI

A one-dimensional numerical case is applied to illustrate the detailed procedure of the method proposed in this subsection. The mathematical model is

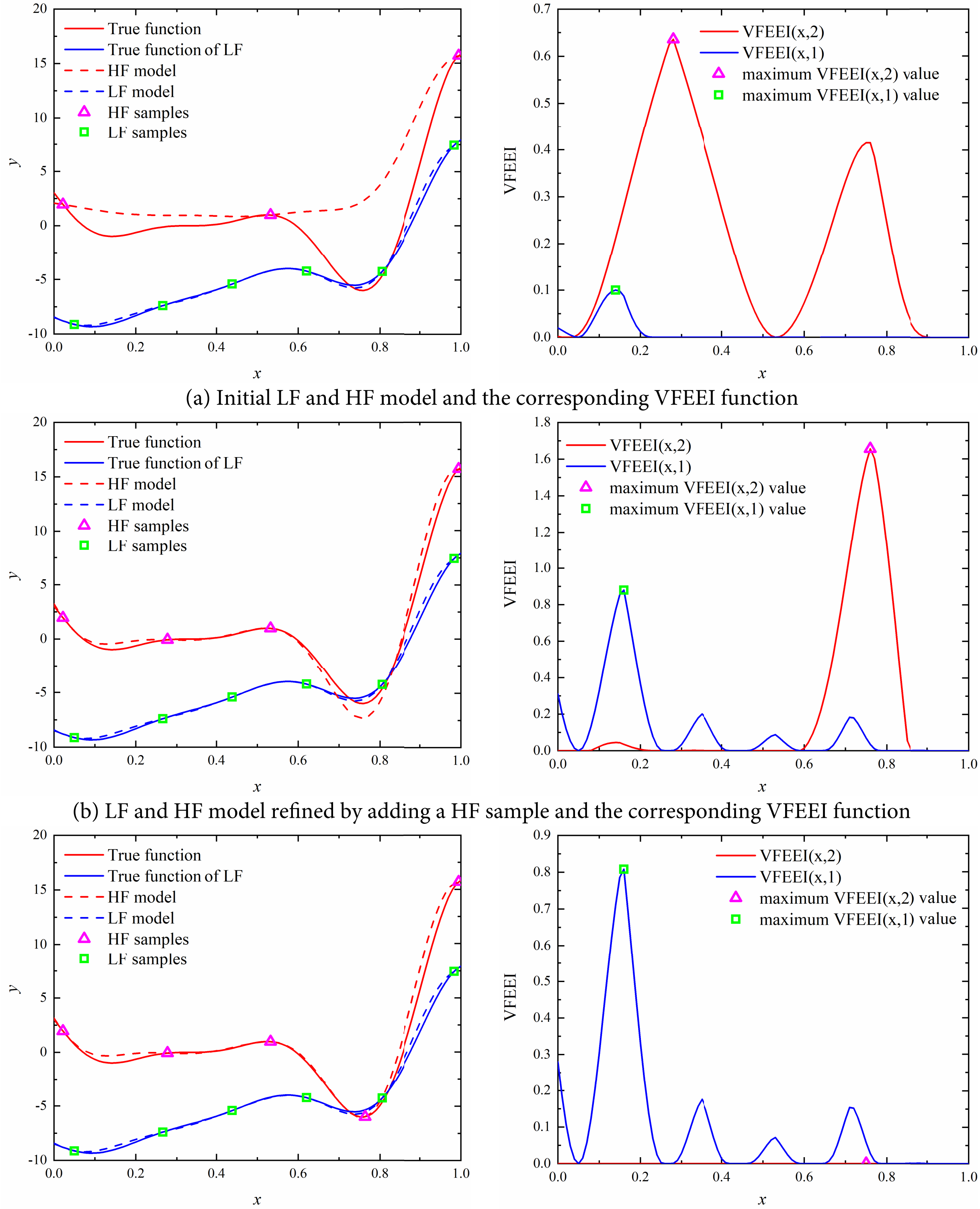
$$\begin{cases} \min f_2(x) = (6x-2)^2 \sin(12x-4) \\ f_1(x) = 0.5 f_2(x) + 10(x-0.5) - 5 \\ \text{s.t.} x \in [0,1] \end{cases} \tag{21}$$

where $f_2(x)$ denotes the HF analytical function and $f_1(x)$ denotes the LF analytical function. The theoretical optimal solution is $x^* = 0.75725$ with the optimal functional value. The computational cost ratio between the HF and LF models is assumed to be 4. Termination condition $\varepsilon = 0.01$.

First, two initial sample sets containing three HF sample points $\boldsymbol{x}_h = \{0.02198, 0.53247, 0.99455\}$ and six LF sample points $\boldsymbol{x}_l = \{0.04996, 0.26578, 0.43828, 0.62033, 0.80575, 0.98358\}$ are generated. Based on the initial sampling points, an initial VF metamodel is built. After the initial VF metamodel is built, the refinement process of the HF model using the proposed method is shown in Fig. 3. The left graph of Fig. 3 shows the iterative process of the surrogate model and the right side of Fig. 3 shows the distribution of the corresponding VFEEI functions in the design space. In the left panel of Fig. 3, the red solid curve denotes the true high-fidelity function, while the blue solid curve denotes the true low-fidelity function. The red dashed curve shows the surrogate's high-fidelity prediction, and the blue dashed curve shows its low-fidelity prediction. Pink hollow triangles mark the high-fidelity sample locations, and green hollow squares indicate the low-fidelity sample locations. In the right panel of Fig. 3, the red solid curve depicts the high-fidelity VFEEI acquisition function, and the blue solid curve depicts the low-fidelity VFEEI acquisition function. Pink hollow triangles denote the maxima of the high-fidelity VFEEI, while green hollow squares denote the maxima of the low-fidelity VFEEI. During the optimization process, the sizes of the VFEEI functions of these two samples are used to determine whether they should be added to the initial sample set.

From Fig. 3a, we can see that the LF surrogate model is not accurate but has a similar trend compared to the HF model, and the maximum value of the VFEEI function of the HF model is much larger than that of the HF model. When an HF sample in the second cycle is added, the accuracy can be improved greatly than the LF added. For this example, the iteration is terminated in the fifth cycle. The optimal solution is obtained by using the ISMA algorithm. For high-fidelity models, when the fitting samples cannot represent the overall trend of the function, global exploration will be conducted through the VFEEI function, as illustrated in Fig. 3b. Conversely, when the fitting samples can represent the overall trend of the function but the local prediction performance is poor, the VFEEI function will focus on the local development, as shown in Fig. 3c. After two iterations, the final high-fidelity function surrogate model is shown in Fig. 3e.

The optimization process is provided in Table 4. When one LF sample and three HF samples are added, the global optimal solutions can be obtained and the last iteration (the fifth cycle) of the surrogate model is provided in Fig. 3e. From the results, we can see that the proposed method can build a highly accurate VF surrogate model by combining the LF and HF samples.

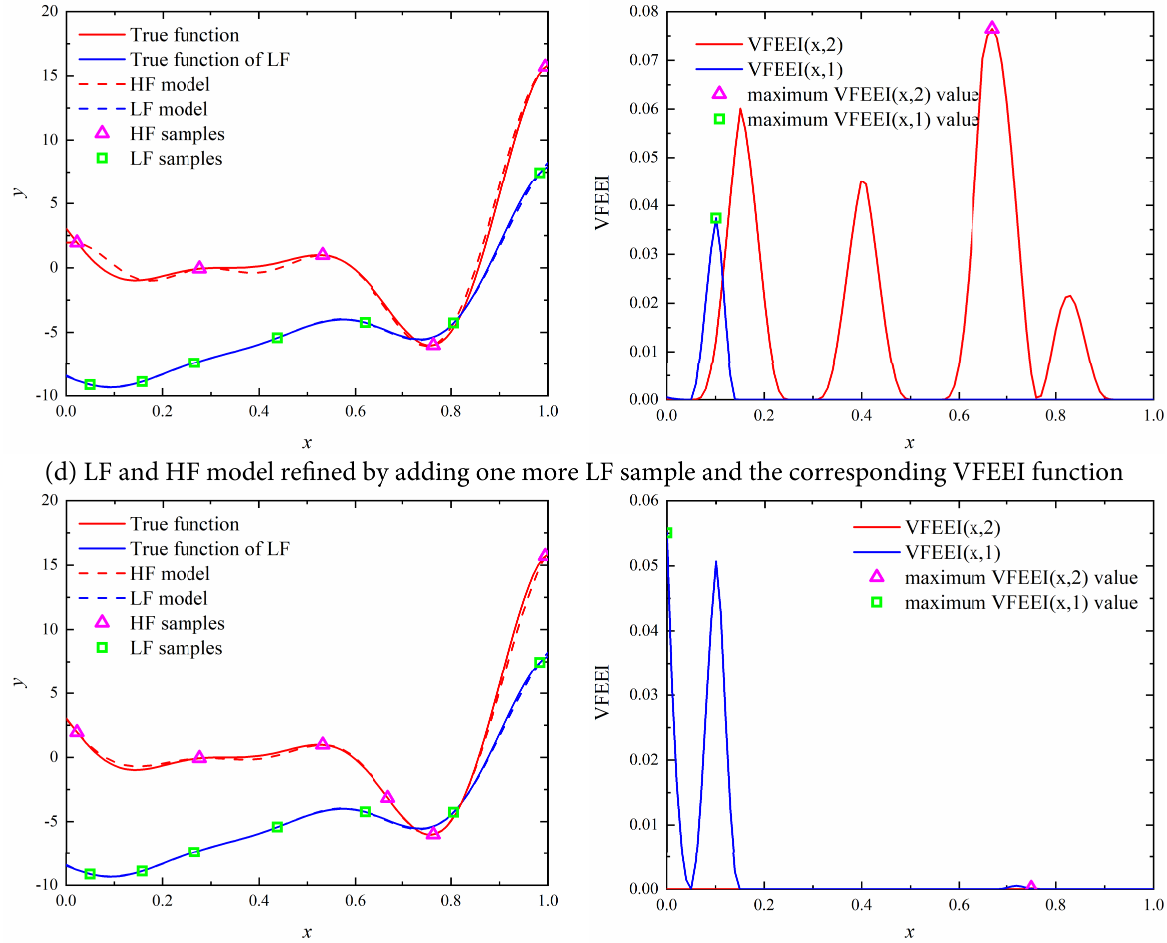


(a) Initial LF and HF model and the corresponding VFEEI function

(b) LF and HF model refined by adding a HF sample and the corresponding VFEEI function

(c) LF and HF model refined by adding a HF sample and the corresponding VFEEI function

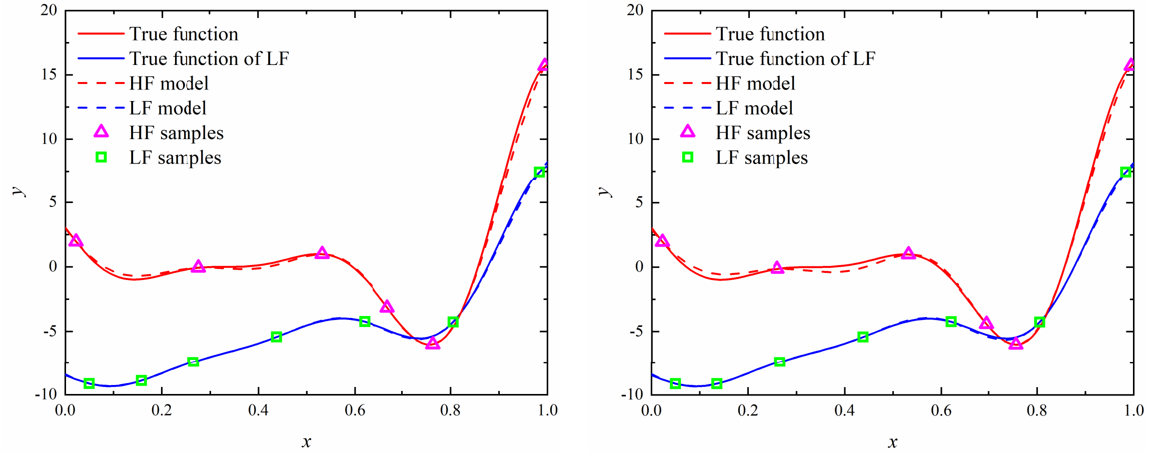
Figure 3: (Continued)



(d) LF and HF model refined by adding one more LF sample and the corresponding VFEEI function

(e) LF and HF model refined by adding one more HF sample and the corresponding VFEEI function

Figure 3: Refinement process of optimization based on the proposed VFEEI



(a) Sampling result of VFEEI function

(b) Sampling result of VFEEI-D function

Figure 4: Comparison chart of sampling results from different acquisition functions

Table 4: Refinement process of the example

Updating cycle	Added sample	Fidelity level	Optimal solution of the HF model
1	0.27723	HF	0.8439
2	0.76351	HF	−7.3530
3	0.15787	LF	6.0958
4	0.66942	HF	−6.0888
5	–	–	−6.0207

To further verify the effectiveness of the proposed VFEEI function, the impact of the sample distance function on global optimization is investigated next. In this case, the sample distance function is removed from the VFEEI function (referred to as VFEEI-D), and optimization is performed based on this modified function. The final comparative results are shown in Fig. 4.

As shown in Fig. 4, the sampling points obtained using the VFEEI-D function tend to cluster, especially near the minimum value. The high-fidelity sampling points obtained using the VFEEI function exhibit significantly better discreteness compared to those obtained using the VFEEI-D function. This conclusion also applies to the low-fidelity sampling process, particularly in the region near 0.1. Additionally, in areas not covered by high-fidelity samples, such as near 0.1 and 0.4, it is evident that the fitting effect of the high-fidelity function in Fig. 4b is slightly worse than that in Fig. 4a. These analyses fully demonstrate that the sample distance function is highly sensitive to the entire sampling process and plays a crucial role in the prediction accuracy of multi-fidelity models.

4.3 Testing Functions of ISMA-VFEEI

In this subsection, six commonly used numerical examples are used to illustrate the effectiveness and efficiency of the proposed ISMA-VFEEI method. Especially, the testing functions in cases 5 and 6 are highly nonlinear. These test functions were modified from references [14,42]. The expressions of the testing functions are shown in Table 5. The computational cost ratio between HF evaluation and LF evaluation is set to 4 for all testing examples.

Table 5: Testing functions

Name	Description of functions
Case 1	$f_2(x) = (6x-2)^2 \sin(12x-4)$ $f_1(x) = 0.5f_2(x) + 10(x-0.5) - 5$ $x \in [0,1], x^* = 0.75725, f^* = -6.020740$
Case 2	$f_2(x) = (6x-2)^2 \sin(12x-4)$ $f_1(x) = f_2(x) - 0.5$ $x \in [0,1], x^* = 0.75725, f^* = -6.020740$
Case 3	$f_2(x) = (6x-2)^2 \sin(12x-4)$ $f_1(x) = f_2(x+0.2)$ $x \in [0,1], x^* = 0.75725, f^* = -6.020740$

(Continued)

Table 5 (continued)

Name	Description of functions
Case 4	$f_2(\mathbf{x}) = \sum_{i=1}^{4} c_i \exp\left[\sum_{j=1}^{3} \alpha_{ij}\left(x_j - p_{ij}\right)^2\right]$ $\left[\alpha_{ij}\right] = \begin{bmatrix} 3 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3 & 10 & 30 \\ 0.1 & 10 & 35 \end{bmatrix}, \left[c_i\right] = \begin{bmatrix} 1 \\ 1.2 \\ 3 \\ 3.2 \end{bmatrix}, \left[p_{ij}\right] = \begin{bmatrix} 0.3689 & 0.1170 & 0.2673 \\ 0.4699 & 0.4387 & 0.7470 \\ 0.1091 & 0.8732 & 0.5547 \\ 0.0381 & 0.5743 & 0.8828 \end{bmatrix}$ $f_1(\mathbf{x}) = f_2(\mathbf{x}) + 7.6 \times \left(\begin{array}{l} 0.585 - 0.324x_1 - 0.379x_2 - 0.431x_3 - 0.208x_1x_2 + \\ 0.326x_1x_3 + 0.193x_2x_3 + 0.225x_1^2 + 0.263x_2^2 + 0.274x_3^2 \end{array}\right)$ $x_1, x_2, x_3 \in [0,1], \mathbf{x}^* = (0.114, 0.556, 0.852), f^* = -3.8627$
Case 5	$f_{\text{borehole}} = \dfrac{2\pi x_3\left(x_4 - x_6\right)}{\ln\left(\dfrac{x_2}{x_1}\right)\left[1 + \dfrac{2x_7x_4}{\ln\left(\dfrac{x_2}{x_1}\right)x_1^2x_8} + \dfrac{x_3}{x_5}\right]}$ $f_1(x) = 0.4f_{\text{borehole}}(x) + 0.07x_1^2x_8 + \dfrac{x_1x_7}{x_3} + \dfrac{x_1x_6}{x_2} + x_1^2x_4$ $f_2(x) = f_{\text{borehole}}(x)$ $x_1 \in [0.05, 0.15], \quad x_2 \in [100, 50000],$ $x_3 \in [63070, 115600], \quad x_4 \in [990, 1110],$ $x_5 \in [63.1, 116], \quad x_6 \in [700, 820],$ $x_7 \in [1120, 1680], \quad x_8 \in [9855, 12045].$ $\mathbf{x}^* = (0.05, 50000, 63070, 990, 63.1, 820, 1680, 9855)$ $f^* = 452.6289$
Case 6	$f_2(\mathbf{x}) = 10^{-4}\sum_{i=1}^{10} \exp\left(x_i\right)\left[A(i) + x_i - \ln\left(\sum_{j=1}^{10} \exp\left(x_j\right)\right)\right]$ $f_1(\mathbf{x}) = 10^{-4}\sum_{i=1}^{10} \exp\left(x_i\right)\left[B(i) + x_i - \ln\left(\sum_{j=1}^{10} \exp\left(x_j\right)\right)\right]$ $A = \left[\begin{array}{l} -6.089, -17.164, -34.054, -5.914, -24.721, \\ -14.986, -24.100, -10.708, -26.662, -22.662 \end{array}\right]$ $B = [-10, -10, -20, -10, -20, -20, -20, -10, -20, -20]$ $x_i \in [-5, 10], i = 1, 2, \ldots, 10, \mathbf{x}^* = (5, 5, 5, 5, 5, 5, 5, 5, 5, 5)$ $f^* = -3.118$

In the first five testing functions, the initial HF and LF samples are set to 3*d* and 6*d*, respectively, while in the last testing case, the initial HF and LF samples are set to 10*d* and 20*d*, respectively, where *d* denotes the dimension of the design variables. Considering the randomness of the LHS and ISMA methods, 10 runs are performed for each method. N_{HF} and N_{LF} denote the number of function calls for the HF and LF models, respectively. In order to measure the robustness of each method, the standard deviation of the total cost is used. The relative error is calculated to describe the computational accuracy. The compared methods used here are the AEI [12], the VF-EI [14], the VF-LCB [17] and the VF-PI [24]. The results are shown in Table 6. All acquisition functions are computed using ISMA.

Table 6: Summary of the comparison results for numerical cases

Testing case	Methods	N_{HF}	N_{LF}	C_{total}	Standard deviation of C_{total}	Difference with optima, %
Case 1	AEI	5.00	2.10	5.53	0.89	1.01×10^{-5}
	VF-EI	5.40	1.80	5.85	0.45	1.00×10^{-5}
	VF-PI	4.40	7.50	6.28	1.86	9.70×10^{-3}
	VF-LCB	6.10	5.70	7.53	1.01	8.11×10^{-4}
	VFEEI	3.20	2.70	3.88	1.10	1.03×10^{-5}
Case 2	AEI	4.09	3.45	4.96	1.50	2.50×10^{-4}
	VF-EI	3.27	3.27	4.09	0.83	5.10×10^{-4}
	VF-PI	4.10	7.36	5.93	1.77	7.47×10^{-3}
	VF-LCB	2.73	5.64	4.14	0.93	1.83×10^{-4}
	VFEEI	2.27	2.82	2.98	0.48	2.32×10^{-4}
Case 3	AEI	6.00	3.18	6.80	1.03	1.22×10^{-4}
	VF-EI	6.18	4.36	7.27	1.17	1.31×10^{-4}
	VF-PI	4.73	7.45	6.59	1.95	6.21×10^{-3}
	VF-LCB	4.82	0	4.82	1.34	8.13×10^{-4}
	VFEEI	4.36	2.45	4.98	0.92	1.21×10^{-4}
Case 4	AEI	16.40	12.00	19.40	9.04	7.64×10^{-3}
	VF-EI	13.40	14.90	17.13	6.21	8.45×10^{-3}
	VF-PI	10.82	16.45	14.93	4.27	9.50×10^{-3}
	VF-LCB	9.91	30.36	17.50	8.69	1.10×10^{-2}
	VFEEI	7.30	8.60	9.45	3.73	6.13×10^{-3}
Case 5	AEI	33.36	18.64	38.02	15.26	4.52×10^{-5}
	VF-EI	32.36	17.09	36.64	15.39	5.32×10^{-5}
	VF-PI	31.72	2.54	32.36	6.18	6.42×10^{-4}
	VF-LCB	36.82	114.27	65.39	27.06	1.29×10^{-3}
	VFEEI	15.55	14.27	19.11	2.76	5.66×10^{-5}
Case 6	AEI	8.18	11.45	8.47	5.58	1.59×10^{-5}
	VF-EI	19.09	11.36	19.38	6.83	1.58×10^{-5}
	VF-PI	18.81	16.36	19.23	4.58	1.59×10^{-5}
	VF-LCB	14.91	20.91	20.14	4.11	1.74×10^{-5}
	VFEEI	6.91	6.73	7.08	4.47	1.59×10^{-5}

Compared with the original EI, VFEEI extends EI to handle multi-fidelity data and additionally considers sample distance and computational cost. Compared with VFEEI, AEI further takes into account the correlation between data of different fidelities. Compared with VF-EI, VFEEI handles the original EI differently; in VF-EI, the same high-fidelity function prediction value is used for EI at different fidelities. All the aforementioned methods are based on EI. In contrast, VF-PI and VF-LCB are improved versions of the probability of improvement function and the low confidence bound function, respectively.

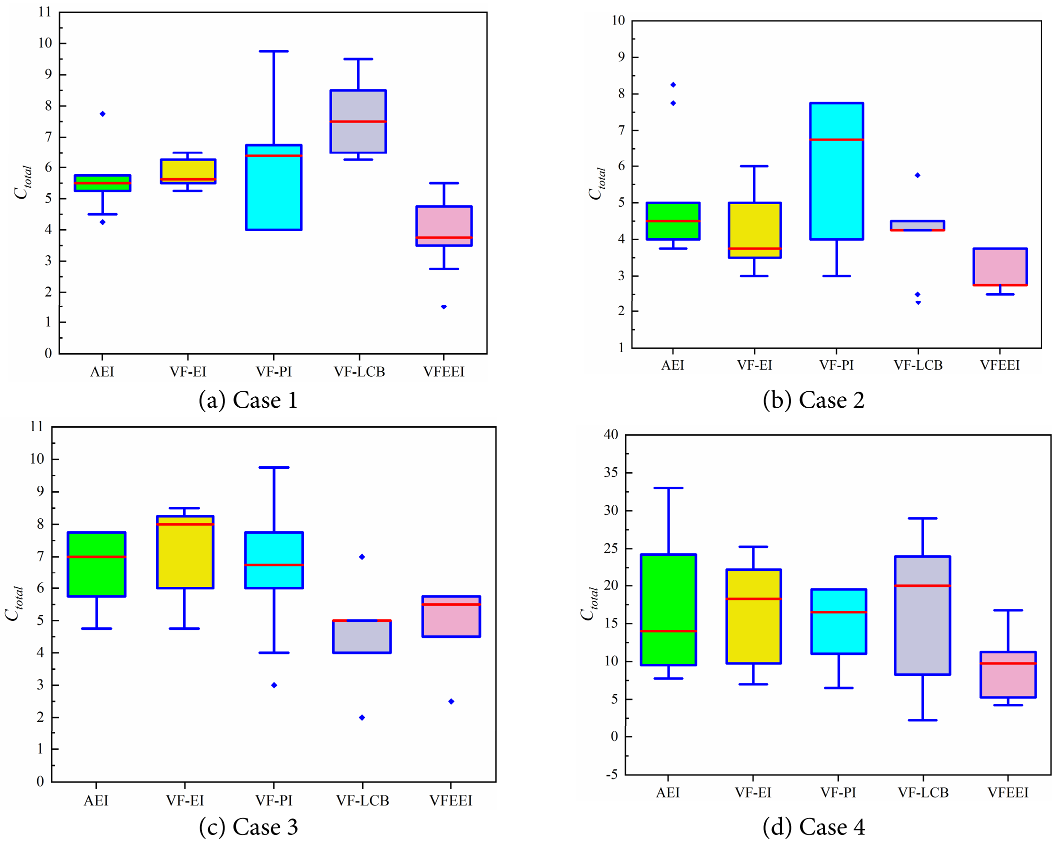
The comparison of different methods reveals significant reductions in both high-fidelity (HF) computation costs and total costs across the six examples. Specifically, compared to the AEI method, the HF calculation cost is reduced to 55% (improved by 36%), and the total cost to 51% (improved by 30%). These reductions indicate the effectiveness of the proposed method in improving computational efficiency.

In comparison with the VF-EI method, the HF calculation cost is reduced by up to 64%, with a corresponding decrease in total cost ranging from 27% to 63%. This shows a notable improvement in computational efficiency, particularly in the cases where the cost reduction is most pronounced.

The VF-PI method also shows a reduction in HF calculation costs (8% to 51%) and total costs (27% to 63%). However, the improvements are less consistent than those observed when compared to the VF-EI method, suggesting that the VF-EI method may offer a more stable advantage in terms of overall computational savings.

When compared to the VF-LCB method, the reductions in HF calculation costs are substantial, ranging from 10% to 58%, while the total cost reduction varies from −3% to 71%. The negative value for the total cost in one of the examples indicates that, in some cases, the VF-LCB method might perform better than the proposed method in terms of total cost, but the proposed method generally shows superior efficiency across the majority of examples.

In summary, while all methods demonstrate varying degrees of improvement, the proposed method consistently outperforms the AEI, VF-EI, VF-PI, and VF-LCB methods in reducing computational costs, particularly for high-fidelity calculations, making it a more efficient approach for the examples studied. To testify to the robustness of all the methods, the standard deviation of the total costs is also provided in Table 5. For an intuitive comparison, a box plot of the total costs for all methods in the six different cases is described in Fig. 5. We can see that all methods exhibit the desired robustness. The robustness of the proposed method is less than the other methods only in Case 1, but better than the other methods in other cases.



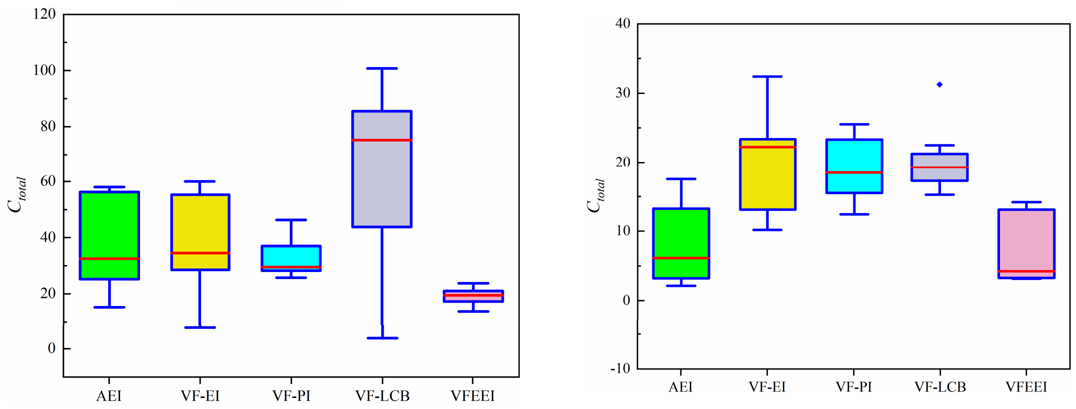
(a) Case 1

(b) Case 2

(c) Case 3

(d) Case 4

Figure 5: (Continued)



(e) Case 5

(f) Case 6

Figure 5: Comparison results of different cases for the proposed VFEEI method and other methods

The results of the Wilcoxon rank-sum test for different methods are shown in Table 7. Among the six test cases, all Wilcoxon rank-sum test p-values are greater than 0.05, indicating that there are no significant differences between the optimization results obtained using VFEEI and those obtained using the other four methods. This is because, in this subsection, the primary focus is on testing the computational effects of different acquisition functions, and the convergence criterion during the optimization process is the optimization precision error within a fixed cost range. Therefore, all methods are able to obtain the optimal solution. In this example, the Wilcoxon rank-sum test is used to compare the statistical results of the optimal solutions computed by different acquisition functions.

Table 7: Wilcoxon rank-sum test results

Testing case	AEI	VF-EI	VF-PI	VF-LCB
Case 1	0.471	0.489	0.884	0.475
Case 2	0.123	0.515	0.322	0.259
Case 3	0.305	0.132	0.246	0.686
Case 4	0.481	0.856	0.103	0.372
Case 5	0.650	0.581	0.364	0.235
Case 6	0.858	0.783	0.673	0.921

4.4 *Engineering Example*

With the quick development of the high-speed railway, the safety of the Electric Multiple Unit (EMU) becomes more and more important as the speed increases. As one of the important structural components of the EMU, the sanding device can improve the adhesion coefficient between the wheel and rail by sanding to ensure the smooth operation of the EMU in harsh weather such as rain and snow, which can prevent wheelset idling and vehicle taxiing, and effectively avoid the occurrence of wheel rub fault. This section mainly conducts static strength analysis on the sanding device.

The entire structure of the sanding device is made of 5-5083-H111 aluminium alloy. Its modulus of elasticity is 7 × 10^4 MPa, density is 2700 kg/m^3, Poisson's ratio is 0.33, yield strength is 139 MPa, and its main structural components include sand box, sand spreader, hoisting structure, sand inlet pipe, etc. The structural composition of the sanding device is shown in Fig. 6.

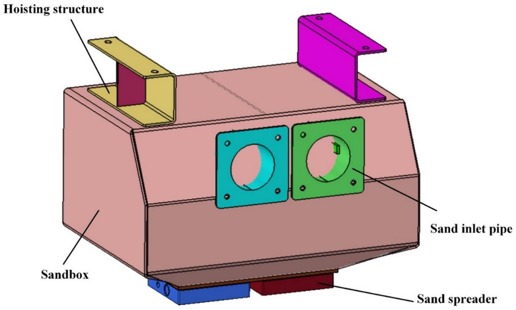


Figure 6: The structural composition of the sanding device

The sanding device is connected to the rest of the rolling stock by a hoisting structure. The constraints of the sanding device are established at the location of the bolt holes in the hoisting structure of the sand spreader, which constrains six directions of freedom, limiting in-plane movement and rotation in space. Also, an acceleration condition with a longitudinal acceleration of 5 g and a vertical acceleration of 1 g was determined as the working condition, where g is the acceleration of gravity. The finalised key parameter locations and constraint locations for the sanding device are shown in Fig. 7. The symbols and thicknesses of the key locations of the sanding device are summarised in Table 8.

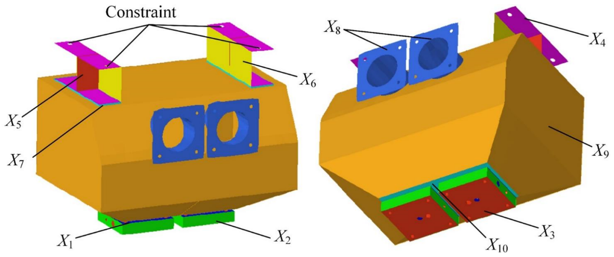


Figure 7: Location of key parameters and constraint positions of the sanding device

Table 8: Sanding device key position symbols and thicknesses

Symbol	Position	Thickness (mm)
X_1	Sander upper plate	4
X_2	Sander side plate	1.155
X_3	Sander base plate	5
X_4	Bottom and upper plates for hoisting	4
X_5	Hoisting rib plate	4.5
X_6	Hoisting side plate	5
X_7	Hoisting connection plate	5.5
X_8	Sand inlet pipe	5
X_9	Sandbox body	5
X_{10}	Sandbox bottom plate	10

In order to determine the sensitive parameters on the stress and mass of the sanding device from $X_1 \sim X_{10}$, the sensitivity analysis is performed and the results are shown in Fig. 8. From Fig. 8, we can see X_4, X_5, X_6, X_7, X_9 are sensitive parameters and the information on the sensitive parameters is provided in Table 9.

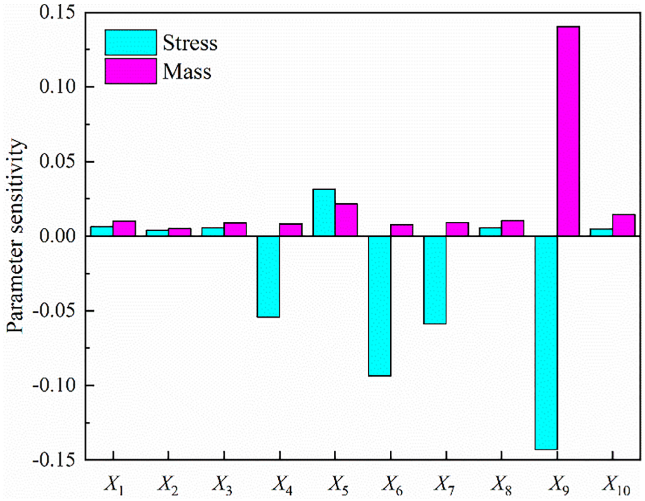


Figure 8: Sensitivity analysis on the stress and mass

Table 9: The information on the design variables of the supporting thin plate structure

Design variables	Ranges (mm)
X_4	[3, 5]
X_5	[3.5, 5.5]
X_6	[4, 6]
X_7	[4.5, 6.5]
X_9	[4, 6]

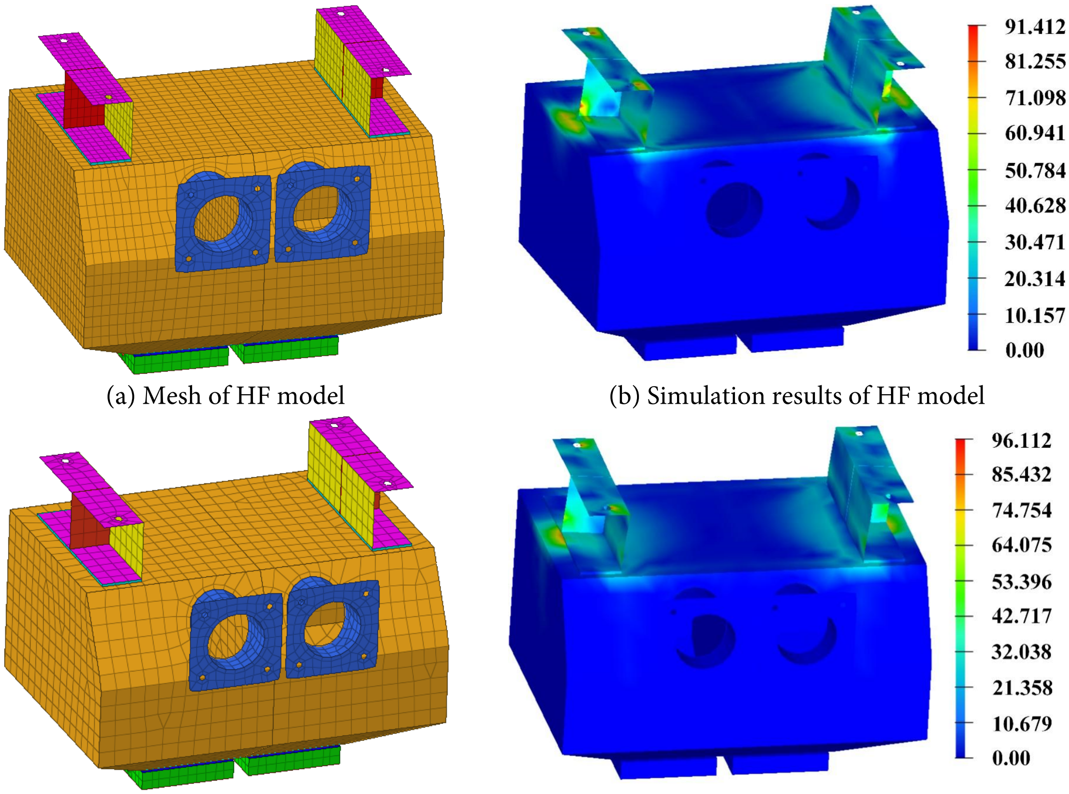
Then the optimization model can be expressed by

$$\begin{array}{l}\text{Find } \mathbf{x} = \{X_4, X_5, X_6, X_7, X_9\} \\ \min f = \dfrac{L_S\left(X_4, X_5, X_6, X_7, X_9\right)}{139} + \dfrac{L_M\left(X_4, X_5, X_6, X_7, X_9\right)}{0.02683} \\ \text{s.t.} 3 \le X_4 \le 5, 3.5 \le X_5 \le 5.5, 4 \le X_6, X_9 \le 6, 4.5 \le X_7 \le 6.5\end{array} \tag{22}$$

where L_S represents the performance function on stress and L_M represents the performance function on mass, which are both obtained from surrogate modelling; 139 MPa represents the yield stress and 0.02683 represents the quality of the original model.

A simulation model with approximately 4343 elements was chosen as the LF model, while the HF simulation model had approximately 7004 elements. The mesh and corresponding simulation results are shown in Fig. 9. The single computation time of the HF model is 62.5682 s, while the LF model is 25.2731 s. The computational time for the HF simulation model is approximately 2.5 times longer than that of the LF simulation model. Therefore, the cost ratio was set to 2.5.

The proposed method and other methods are applied to solve this engineering problem. The optimization starts with 5 HF initial sampling points and 10 LF initial sampling points. The termination condition for the optimization is set to the total computational cost, which is determined to be no more than the computational time of the 20 HF simulation models. The optimal results for the design variables for each method are given in Table 10. Table 11 summarizes the optimization target results and computational costs for all methods.



(a) Mesh of HF model

(b) Simulation results of HF model

(c) Mesh of LF model

(d) Simulation results of LF model

Figure 9: The mesh and corresponding simulation results of the sanding device

Table 10: The optima of different methods of the sanding device

Methods	X_4 (mm)	X_5 (mm)	X_6 (mm)	X_7 (mm)	X_9 (mm)
AEI	5.12	4.20	5.50	6.65	4.50
VF-EI	3.50	5.50	6.00	4.50	4.00
VF-PI	3.00	4.55	4.92	6.50	5.28
VFEEI	4.55	5.50	6.00	6.50	4.00

Table 11: The actual responses of different methods of the sanding device

Methods	f	N_{HF}	N_{LF}	C_{total}
AEI	1.602	10	25	20.0
VF-EI	1.553	13	18	20.2
VF-PI	1.564	12	20	20.0
VFEEI	1.546	14	15	20.0

As can be seen in Tables 10 and 11, all methods obtained the optimised design solution while satisfying the termination criterion. Among the four methods, the VFEEI obtained the globally optimal solution with the C_{total} of 20.0, which is reduced by 1.0% compared with the VF-EI method. From the optimization results, VFEEI achieved the minimum value of optimization, which was reduced by 3.50%, 0.45%, and 1.15%

compared with AEI, VF-EI, and VF-PI, respectively. The finite element simulation results of the optimal solution are provided in Fig. 10 from the proposed method.

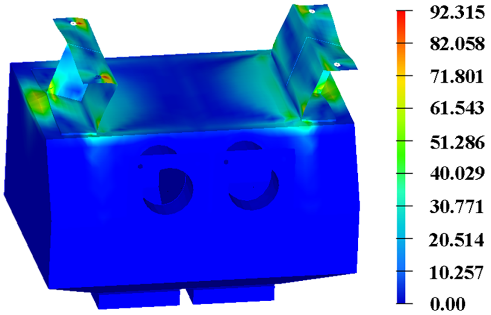


Figure 10: Simulation results of the optimum model by the proposed VFEEI method

5 Conclusion

In this paper, a novel VF Kriging surrogate model-based global optimization method for black-box function (ISMA-VFEEI) is proposed based on the Hierarchical Kriging model. In the proposed method, the VFEEI sampling function is developed to properly select new samples with low and high fidelity to update the surrogate model so that it obtains the global optimal solution. In addition, to improve the accuracy and efficiency of the optimization during the sample updating procedure, an improved slime mould algorithm (ISMA) is proposed to obtain the optimal value of the VFEEI function. Finally, a computational framework applicable to the proposed method is constructed.

The main conclusions can be summarized as follows: (1) ISMA-VFEEI allows for an objective balancing of global exploration and local exploitation to find the global optimal solution, and also achieves robust results by consuming fewer computational resources. (2) Six numerical examples illustrate the advantages of the proposed method in terms of computational cost and robustness. The quartile spacing of the proposed method is smaller than that of the other methods, which demonstrates the better robustness of the proposed method. (3) The optimization of the sanding device verifies the engineering application value of this method. Compared with the VF-EI and AEI methods, this method achieved better results with fewer calculations.

In our future work, the method proposed in this paper can be extended to conduct uncertainty optimization to be more applicable for complex engineering problems.

Acknowledgement: This work was funded by National Natural Science Foundation of China; Special Program of Huzhou; Projects of Huzhou Science and Technology Correspondent; Guangdong Basic and Applied Basic Research Foundation.

Funding Statement: This work was funded by National Natural Science Foundation of China (grant No. 52405255); Special Program of Huzhou (grant No. 2023GZ05); Projects of Huzhou Science and Technology Correspondent (grant No. 2023KT76); Guangdong Basic and Applied Basic Research Foundation (grant No. 2025A1515010487).

Author Contributions: Yi Guan: conceptualization, methodology, writing—original draft preparation; Pengpeng Zhi: supervision, project administration, funding acquisition; Zhonglai Wang: supervision, writing—review and editing. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Yang H, Hong SH, Wang Y. A sequential multi-fidelity surrogate-based optimization methodology based on expected improvement reduction. Struct Multidiscip Optim. 2022;65(5):153. doi:10.1007/s00158-022-03240-x.
2. Biagiola SI, Figueroa JL. Wiener and Hammerstein uncertain models identification. Math Comput Simul. 2009;79(11):3296–313. doi:10.1016/j.matcom.2009.05.004.
3. Harkin EF, Shen PR, Goel A, Richards BA, Naud R. Parallel and recurrent cascade models as a unifying force for understanding subcellular computation. Neuroscience. 2022;489:200–15. doi:10.1016/j.neuroscience.2021.07.026.
4. Zhang J, Chowdhury S, Messac A. An adaptive hybrid surrogate model. Struct Multidiscip Optim. 2012;46(2):223–38. doi:10.1007/s00158-012-0764-x.
5. Zhou Q, Shao X, Jiang P, Zhou H, Shu L. An adaptive global variable fidelity metamodeling strategy using a support vector regression based scaling function. Simul Model Pract Theory. 2015;59(3):18–35. doi:10.1016/j.simpat.2015.08.002.
6. Tyan M, Nguyen NV, Lee JW. Improving variable-fidelity modelling by exploring global design space and radial basis function networks for aerofoil design. Eng Optim. 2015;47(7):885–908. doi:10.1080/0305215X.2014.941290.
7. Liu Y, Collette M. Improving surrogate-assisted variable fidelity multi-objective optimization using a clustering algorithm. Appl Soft Comput. 2014;24:482–93. doi:10.1016/j.asoc.2014.07.022.
8. Li X, Qiu H, Jiang Z, Gao L, Shao X. A VF-SLP framework using least squares hybrid scaling for RBDO. Struct Multidiscip Optim. 2017;55(5):1629–40. doi:10.1007/s00158-016-1588-x.
9. Viana FAC, Simpson TW, Balabanov V, Toropov V. Special section on multidisciplinary design optimization: metamodeling in multidisciplinary design optimization: how far have we really come? AIAA J. 2014;52(4):670–90. doi:10.2514/1.j052375.
10. Xiao M, Zhang G, Breitkopf P, Villon P, Zhang W. Extended Co-Kriging interpolation method based on mul-ti-fidelity data. Appl Math Comput. 2018;323(4):120–31. doi:10.1016/j.amc.2017.10.055.
11. Han ZH, Görtz S. Hierarchical Kriging model for variable-fidelity surrogate modeling. AIAA J. 2012;50(9):1885–96. doi:10.2514/1.j051354.
12. Huang D, Allen TT, Notz WI, Miller RA. Sequential Kriging optimization using multiple-fidelity evaluations. Struct Multidiscip Optim. 2006;32(5):369–82. doi:10.1007/s00158-005-0587-0.
13. Xiong Y, Chen W, Tsui KL. A new variable-fidelity optimization framework based on model Fusi on and objective-oriented sequential sampling. J Mech Des. 2008;130(11):111401. doi:10.1115/1.2976449.
14. Zhang Y, Han ZH, Zhang KS. Variable-fidelity expected improvement method for efficient global optimization of expensive functions. Struct Multidiscip Optim. 2018;58(4):1431–51. doi:10.1007/s00158-018-1971-x.
15. Liu Y, Chen S, Wang F, Xiong F. Sequential optimization using multi-level cokriging and extended expected improvement criterion. Struct Multidiscip Optim. 2018;58(3):1155–73. doi:10.1007/s00158-018-1959-6.
16. Han Z, Xu C, Zhang L, Zhang Y, Zhang K, Song W. Efficient aerodynamic shape optimization using varia-ble-fidelity surrogate models and multilevel computational grids. Chin J Aeronaut. 2020;33(1):31–47. doi:10.1016/j.cja.2019.05.001.
17. Jiang P, Cheng J, Zhou Q, Shu L, Hu J. Variable-fidelity lower confidence bounding approach for engineering optimization problems with expensive simulations. AIAA J. 2019;57(12):5416–30. doi:10.2514/1.j058283.
18. Serani A, Pellegrini R, Wackers J, Jeanson CE, Queutey P, Visonneau M, et al. Adaptive multi-fidelity sampling for CFD-based optimisation via radial basis function metamodels. Int J Comput Fluid Dyn. 2019;33(6–7):237–55. doi:10.1080/10618562.2019.1683164.
19. Yi J, Shen Y, Shoemaker CA. A multi-fidelity RBF surrogate-based optimization framework for computationally expensive multi-modal problems with application to capacity planning of manufacturing systems. Struct Multidiscip Optim. 2020;62(4):1787–807. doi:10.1007/s00158-020-02575-7.
20. Shi M, Lv L, Sun W, Song X. A multi-fidelity surrogate model based on support vector regression. Struct Multidiscip Optim. 2020;61(6):2363–75. doi:10.1007/s00158-020-02522-6.

21. He Y, Sun J, Song P, Wang X. Variable-fidelity hypervolume-based expected improvement criteria for mul-ti-objective efficient global optimization of expensive functions. Eng Comput. 2022;38(4):3663–89. doi:10.1007/s00366-021-01404-9.
22. Guo Z, Wang Q, Song L, Li J. Parallel multi-fidelity expected improvement method for efficient global optimization. Struct Multidiscip Optim. 2021;64(3):1457–68. doi:10.1007/s00158-021-02931-1.
23. Cheng J, Lin Q, Yi J. An enhanced variable-fidelity optimization approach for constrained optimization problems and its parallelization. Struct Multidiscip Optim. 2022;65(7):188. doi:10.1007/s00158-022-03283-0.
24. Ruan X, Jiang P, Zhou Q, Hu J, Shu L. Variable-fidelity probability of improvement method for efficient global optimization of expensive black-box problems. Struct Multidiscip Optim. 2020;62(6):3021–52. doi:10.1007/s00158-020-02646-9.
25. Liu Y, Wang S, Zhou Q, Lv L, Sun W, Song X. Modified multifidelity surrogate model based on radial basis func-tion with adaptive scale factor. Chin J Mech Eng. 2022;35(1):77. doi:10.1186/s10033-022-00742-z.
26. Li C, Dong H. A modified trust-region assisted variable-fidelity optimization framework for computationally expensive problems. Eng Comput. 2022;39(7):2733–54. doi:10.1108/ec-08-2021-0456.
27. Tao G, Fan C, Wang W, Guo W, Cui J. Multi-fidelity deep learning for aerodynamic shape optimization using convolutional neural network. Phys Fluids. 2024;36(5):056116. doi:10.1063/5.0205780.
28. Zhang Y, Park S, Simeone O. Multi-fidelity Bayesian optimization with across-task transferable max-value en-tropy search. IEEE Trans Signal Process. 2025;73(1):418–32. doi:10.1109/TSP.2025.3528252.
29. Jariego Pérez LC, Garrido Merchán EC. Towards automatic Bayesian optimization: a first step involving acquisition functions. In: Advances in artificial intelligence. Cham: Springer International Publishing; 2021. p. 160–9. doi:10.1007/978-3-030-85713-4_16.
30. Vincent AM, Jidesh P. An improved hyperparameter optimization framework for AutoML systems using evolutionary algorithms. Sci Rep. 2023;13(1):4737. doi:10.1038/s41598-023-32027-3.
31. Li S, Chen H, Wang M, Heidari AA, Mirjalili S. Slime mould algorithm: a new method for stochastic optimiza-tion. Future Gener Comput Syst. 2020;111:300–23. doi:10.1016/j.future.2020.03.055.
32. Ajiboye OK, Ofosu EA, Gyamfi S, Oki O. Hybrid renewable energy system optimization via slime mould algorithm. Int J Eng Trends Technol. 2023;7(6):83–95. doi:10.14445/22315381/ijett-v71i6p210.
33. Zhu T, Wan H, Ouyang Z, Wu T, Liang J, Li W, et al. Multi-objective slime mold algorithm: a slime mold approach using multi-objective optimization for parallel hybrid power system. Sens Mater. 2022;34(10):3837. doi:10.18494/sam4020.
34. Singh T. Chaotic slime mould algorithm for economic load dispatch problems. Appl Intell. 2022;52(13):15325–44. doi:10.1007/s10489-022-03179-y.
35. Wu S, Heidari AA, Zhang S, Kuang F, Chen H. Gaussian bare-bone slime mould algorithm: performance optimization and case studies on truss structures. Artif Intell Rev. 2023;56(9):1–37. doi:10.1007/s10462-022-10370-7.
36. Devarajah LAP, Ahmad MA, Jui JJ. Identifying and estimating solar cell parameters using an enhanced slime mould algorithm. Optik. 2024;311:171890. doi:10.1016/j.ijleo.2024.171890.
37. Zhang L, Jin G, Liu T, Zhang R. Generalized hierarchical expected improvement method based on black-box functions of adaptive search strategy. Appl Math Model. 2022;106(4):30–44. doi:10.1016/j.apm.2021.12.041.
38. Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. J Glob Optim. 1998;13(4):455–92. doi:10.1023/A:1008306431147.
39. Huang H, Liu Z, Zheng H, Xu X, Duan Y. A proportional expected improvement criterion-based multi-fidelity sequential optimization method. Struct Multidiscip Optim. 2023;66(2):30. doi:10.1007/s00158-022-03484-7.
40. Long W, Jiao J, Liang X, Cai S, Xu M. A random opposition-based learning grey wolf optimizer. IEEE Access. 2019;7:113810–25. doi:10.1109/access.2019.2934994.
41. Cao L, Yue Y, Chen Y, Chen C, Chen B. Sailfish optimization algorithm integrated with the osprey optimization algorithm and cauchy mutation and its engineering applications. Symmetry. 2025;17(6):938. doi:10.3390/sym17060938.
42. Hu J, Peng Y, Lin Q, Liu H, Zhou Q. An ensemble weighted average conservative multi-fidelity surrogate modeling method for engineering optimization. Eng Comput. 2022;38(3):2221–44. doi:10.1007/s00366-020-01203-8.