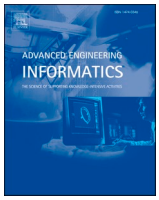
Advanced Engineering Informatics 69 (2026) 103973

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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



A novel adaptive quality-based multi-fidelity surrogate framework for multiple low-fidelity data sources

Mingyu Lee [a], Juyoung Lee [b], Jae-Hoon Choi [c], Nam H. Kim [d], Ikjin Lee [a,b,*]

[a] KAIST InnoCORE PRISM-AI Center, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea
[b] Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea
[c] Division of Mechanical Design Engineering, Jeonbuk National University, Jeonbuk-do 54896, Republic of Korea
[d] Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Keywords:
Adaptive quality-based multi-fidelity (AQBMF) surrogate
Multiple low-fidelity data sources
Multi-fidelity combination method
Data quality
Basis function

ABSTRACT

In this paper, a novel adaptive quality-based multi-fidelity (AQBMF) surrogate framework is introduced to maximize the utilization of low-fidelity (LF) data from various domains. The main goal of the proposed method is to adaptively select and combine LF data, by assessing its quality, to create the most accurate surrogate. The core idea lies in interpreting the quality levels of LF data sources as the relative importance of LF surrogates that serve as basis functions in a multi-fidelity (MF) surrogate. Based on this approach, the proposed AQBMF surrogate framework comprises four main stages. In the first stage, a newly defined augmented MF formulation is constructed, initially assuming equal importance for all LF data sources. In the second stage, LF surrogates are ranked by importance through the proposed MF basis screening method. In the third stage, promising candidate surrogates are systematically constructed based on the importance ranking of the LF surrogates. During this stage, both the selection and filtering of LF data, as well as the hierarchical and ensemble combination-based MF methods are considered. In the last stage, the best surrogate is selected from the candidates using the proposed algorithm. Various benchmark test results demonstrate the superior performance of the proposed framework. Finally, engineering application results show that the proposed AQBMF surrogate achieves higher accuracy than existing ones within the same computational budget.

1. Introduction

A surrogate model is a data-driven mathematical model that approximates the response of a real-world physical system [1]. Once constructed, the surrogate models can quickly predict responses, making them highly effective for processes such as optimization [2–6], uncertainty propagation [7,8], and sensitivity analysis [9], which are often resource- and time-intensive. Despite these advantages, obtaining sufficient data remains a significant challenge when the behavior of the physical system is highly complex or its responses are costly to obtain.

To alleviate this problem, the multi-fidelity (MF) surrogate method emerges as a promising strategy, aggregating auxiliary data from diverse sources of varying accuracy levels to make accurate predictions [10,11]. The MF surrogate model is a data fusion technique that uses a large number of inexpensive, less accurate low-fidelity (LF) samples to quickly capture overall trends and corrects them with a small number of costly, high-accuracy high-fidelity (HF) samples [10,11]. Due to their characteristic of complementarily leveraging data with different fidelity levels to create synergy, MF strategies successfully deliver rapid and accurate solutions to complex real-world problems [12–18].

In real-world scenarios, it is common to encounter MF datasets containing LF data from multiple sources. Depending on whether the fidelity levels of multiple LF data sources are known in advance as prior information, MF surrogates can be broadly classified into (1) hierarchical and (2) ensemble combination approaches. The former method, especially in early studies, assumes that the fidelity levels are known in advance, so all LF resources are combined in a hierarchical manner and then HF data is utilized for calibration. Representative surrogates are scaling function-based methods [19], Co-Kriging [11], hierarchical Kriging (HK) [20], improved hierarchical Kriging (IHK) [21], generalized hierarchical Co-Kriging (GCK) [22], MF surrogate based on design variable correlations (MFS-DVC) [23], and ensemble learning based MF

* Corresponding author at: KAIST InnoCORE PRISM-AI Center & Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea.
E-mail address: ikjin.lee@kaist.ac.kr (I. Lee).

https://doi.org/10.1016/j.aei.2025.103973
Received 21 June 2025; Received in revised form 9 September 2025; Accepted 10 October 2025
Available online 17 October 2025

surrogate (EL-MFS) [24]. However, when the fidelity level of LF data is not known in advance, applying hierarchical combination methods becomes impossible, prompting more recent studies to address this issue. In such cases, as in the latter approach, all LF resources are combined in an ensemble manner, while HF data is utilized for correction. Methods like linear regression MF surrogate (LR-MFS) [25], extended Co-Kriging (ECK) [26], variance-weighted sum MF surrogate (VWS-MFS) [27], non-hierarchical Co-Kriging (NHLF-Co-Kriging) [28], extended hierarchical Kriging (EHK) [29], local correlation-weighted fusion-based MF surrogate (LCWF-MFS) [30], and weighted-sum of multi-HK (WSMHK) [31] fall under this approach. However, in practice, differences in fidelity levels may exist among LF sources, even if they are not explicitly labeled. This suggests that in certain situations, hierarchical MF models may offer advantages over ensemble-based ones. In particular, the definition of non-hierarchical relationships is often ambiguous in industrial settings. For instance, suppose there are two LF sources, and the first has a slightly higher correlation with the HF data than the second. In such a case, it is unclear whether the two LF sources should be treated in a hierarchical or ensemble manner. This ambiguity makes it difficult to determine which type of MF modeling—hierarchical or ensemble—is more appropriate. Moreover, there are cases where both LF sources have weak correlations with the HF data individually. Yet, when combined in a hierarchical manner, their joint information may result in a much stronger correlation. Although such scenarios are realistic, existing ensemble-based MF methods have not considered them. In other words, when fidelity information is unavailable, existing methods do not provide guidance on which approach—hierarchical or ensemble—is more reasonable.

Besides, existing MF methods have another critical limitation that they unconditionally use all LF data sources without scrutiny. Contrary to the common assumption in many MF studies that LF data is always informative, in practice, low-quality LF data can negatively impact the MF surrogate model. For example, if the quality of LF data is poor, it can be detrimental to identifying the overall trend when constructing an MF surrogate. Moreover, even if high-quality LF data is available, a lack of sufficient LF data can hinder the construction of an accurate MF surrogate. For example, LF data may be of good quality but generated long ago, making it impossible to reproduce, thereby hindering trend identification due to insufficient LF data when constructing an MF surrogate. These situations are frequently encountered in industrial environments. In the worst-case scenario, the accuracy of the MF surrogate may fall below that of a single-fidelity (SF) surrogate constructed solely with HF samples. Therefore, careful consideration is required when applying the MF method. Existing studies have utilized discriminant metrics such as normalized cross validation error (NCVE) [32], maximum likelihood estimation (MLE) [32], Pearson correlation coefficient (PCC) [19,33]. In addition, attempts have been made to address this issue by excluding only highly biased LF data sources when building the MF surrogates [34]. However, these studies still have the limitation that they only consider the effect of a single LF data source among multiple LF data sources to determine whether to use that LF source. In particular, even if individual LF sources do not contribute much to the MF surrogate, using them together can sometimes produce a synergistic effect, which has not been fully considered in existing research. In other words, existing studies have not considered the potential coupling effects of multiple LF sources while determining whether to utilize them.

From the review of existing literature surveys, two ongoing challenges stand out when dealing with multiple LF data sources whose fidelity levels are not pre-defined. Firstly, there is no research on choosing the appropriate MF combination method. Specifically, there is a need for studies focusing on choosing between hierarchical and ensemble combination-based MF methods. Secondly, there is a lack of research on MF methods that consider the quality of the LF data sources. This involves identifying combinations of LF data sources that produce synergistic effects and pinpointing LF data sources to disregard or overlook. To the authors' best knowledge, existing approaches have not effectively addressed both of these two challenges related to the flexible utilization of LF data.

In this paper, a novel adaptive quality-based multi-fidelity (AQBMF) surrogate framework is proposed. This framework introduces a higher-level decision-making paradigm for the flexible and effective utilization of multiple LF data sources. The key innovation of this study lies in proposing a unified and adaptive framework that constructs the best surrogate model for a given scenario. Unlike conventional approaches that focus on improving specific surrogate formulations, the proposed method guides the selection and application of surrogate strategies—such as hierarchical-based, quality-aware ensemble-based, or even SF models—based on the characteristics and quality of the available data. To systematically orchestrate multiple data sources, the core idea involves treating the quality levels of LF data sources as the relative significance of the LF surrogates, which serve as basis functions, also known as trend functions, for the MF surrogate. Based on this idea, the proposed AQBMF method consists of four main stages: (1) construction of an augmented MF formulation, (2) MF basis screening, (3) candidate surrogate generation, and (4) best surrogate selection. The effectiveness of the proposed method is demonstrated through various numerical and engineering examples.

The main contributions of the proposed method can be summarized as follows:

- A novel AQBMF surrogate method is introduced, providing a higher-level paradigm for adaptive selection and combination of LF sources, particularly when prior fidelity information is unavailable.
- The proposed framework systematically incorporates multiple LF sources through a four-stage procedure, by treating LF data quality as the relative importance (ranking) of basis functions in the MF surrogate.
- The proposed method achieves superior performance compared to existing approaches under limited computational budgets through systematic surrogate model generation.

The remainder of this paper is structured as follows. Section 2 provides a comprehensive review of conventional MF surrogates and basis screening methods. Section 3 provides a step-by-step explanation of the newly proposed AQBMF surrogate framework. In Section 4, the superiority of the proposed method is validated through various numerical examples. Finally, Section 5 concludes the paper and outlines potential avenues for future research.

2. Review of conventional methods

This section provides a concise overview of the fundamental information underpinning the proposed research. Section 2.1 introduces the basic concept of Gaussian process surrogates and their related formulations. Additionally, Section 2.2 explains the basis selection method along with its accompanying algorithm.

2.1. Gaussian process surrogates

Assuming that a set of input and output data is given, the Kriging model [35–37], also known as a Gaussian process model, can be constructed. In the Kriging method, the random function $Y(\boldsymbol{x})$ consists of a summation of deterministic and stochastic parts, which can be expressed as

$$Y(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\beta} + Z(\boldsymbol{x}) \tag{1}$$

where $\boldsymbol{f}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\beta}$ reflects the trend of the true function, $\boldsymbol{f}(\boldsymbol{x})$ is a design matrix based on pre-selected basis functions (e.g., polynomial basis functions), $\boldsymbol{\beta}$ is a coefficient vector of the corresponding basis functions, and $Z(\boldsymbol{x})$ is the zero-mean Gaussian process with a process variance of $\widehat{\sigma}^2$. Then, the mean predictor $\widehat{y}(\boldsymbol{x})$ and its mean-squared error (MSE)

$s^2(\boldsymbol{x})$ can be obtained by applying the Kriging theory [35]. Here, if $\boldsymbol{f}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\beta}$ becomes $1 \times \beta$, it can be called ordinary Kriging [37].

When dealing with datasets with multiple fidelity levels defined, it becomes feasible to build an MF surrogate model, such as hierarchical Kriging (HK) [20]. If there are MF datasets with $N+1$ levels, LF information replaces the traditional polynomial basis functions in the trend function of the MF surrogate model. Therefore, the stationary random process corresponding to the true function can be written as

$$Y_i(\boldsymbol{x}) = \begin{cases} \boldsymbol{f}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\beta} + Z_i(\boldsymbol{x}), & i = 1, \\ s_{i-1}\widehat{y}_{\mathrm{LF},i-1}(\boldsymbol{x}) + Z_i(\boldsymbol{x}), & i = 2, \ldots, N+1, \end{cases} \tag{2}$$

where s and $\widehat{y}_{\mathrm{LF}}(\boldsymbol{x})$ are the scaling factor and mean prediction of LF surrogate, respectively, and subscript i means the corresponding value belongs to the i^{th} level, where $i = 1$ and $i = N+1$ indicate the lowest and highest fidelity levels, respectively.

More recently, to cope with the problems where there is no prior information about fidelity levels among multiple LF models (i.e., $i = 1, \ldots, N$), MF surrogate models combining ensemble LF models with HF data have been introduced [26–30]. One of the representative methods is the extended hierarchical Kriging (EHK) [29], which can be represented as

$$Y_{N+1}(\boldsymbol{x}) = \sum_{i=1}^{N} w_i \widehat{y}_{\mathrm{LF},i}(\boldsymbol{x}) + Z_{N+1}(\boldsymbol{x}) \tag{3}$$

where w_i is the weight factor.

2.2. Basis selection method

This section explains the method for selecting basis functions when constructing surrogate models. In general, selecting appropriate basis functions has a significant impact on improving the performance of the surrogate model. As mentioned earlier in Section 2.1, in the Kriging method, a response typically comprises two main components, the mean structure and the stochastic parts. In ordinary Kriging, the mean structure is assumed to be constant, while in universal Kriging, it is typically constructed as a first- or second-order polynomial function. Therefore, the choice of basis functions in the mean structure part plays a critical role in determining the performance of Kriging. Several research studies have explored different approaches for selecting these basis functions, such as blind Kriging [38], dynamic Kriging [39], and basis screening Kriging [40]. Among them, the basis screening Kriging method [40] is known for being more accurate and efficient than other approaches. The algorithm for this technique can be summarized as follows:

Step 1: Set a maximum order for polynomial basis functions.
Step 2: Estimate the significance of individual basis functions.
Step 3: Gradually add the basis functions based on their estimated importance, constructing the candidate basis function set.
Step 4: Identify the optimal basis function set from the candidate set using cross-validation error.

3. Proposed AQBMF surrogate framework

In this section, a novel AQBMF surrogate framework that fully leverages LF data is proposed. Firstly, the approach to reinterpreting the MF surrogate model from the perspective of basis functions is introduced. Secondly, the proposed MF framework is illustrated step-by-step. Sequentially, a validation metric and the overall procedure of the proposed method are described.

3.1. Thoughts on the MF surrogate

This section elaborates on our insights into the MF surrogate model. In real-world industrial settings, it is common to encounter a lack of HF data required to develop predictive models in the current domain, as shown in Fig. 1. In such cases, leveraging low-cost LF data from various domains to construct data fusion models, such as MF surrogates, can be a promising solution. Many previous studies have successfully applied MF surrogates with the support of high-quality LF data that effectively reflect the underlying physics. However, it cannot be guaranteed that such LF information will always be beneficial for HF data. In the worst scenario, poor-quality LF data can even degrade the overall surrogate performance. Moreover, it can be challenging even for physics domain experts to distinguish the quality levels of LF data sources. Therefore, it is essential to develop data-driven methods that automatically determine how to effectively utilize multiple LF data sources.

To address these challenges, in this study, we focus on the formulation of Kriging and HK in terms of basis functions in surrogate models. As mentioned earlier in Section 2.1, both data-driven models are divided into a mean structure part and a stochastic part, where the mean structure takes the form of a multiplication of the basis function set and the coefficient vector. Afterwards, the estimated hyper-parameters through MLE depend on how the basis function set is configured, naturally influencing the values of the coefficient vector and other components. Here, when the polynomial function set is employed as basis functions, it becomes a Kriging model. On the other hand, utilizing an LF data-driven surrogate that efficiently captures the system's trend as basis functions results in an HK model. Therefore, MF surrogate modeling can be redefined as an efficient surrogate modeling approach that utilizes high-quality basis vectors. In other words, Kriging is a special case of HK where the basis function is a polynomial function.

Based on these insights, the use of multiple LF data sources can be addressed by translating from a physical perspective to a surrogate model perspective. Generally, LF data contribute to the construction of an MF surrogate in the form of LF surrogates. Moreover, the performance of these LF surrogates directly impacts the performance of the MF surrogate as basis functions. Therefore, this study interprets the quality levels of LF data sources as the relative importance of basis functions within the MF surrogate, which can be summarized as follows:

- High-quality LF sources can be considered important basis functions in the surrogate model.
- Low-quality LF sources can be regarded as less important basis functions in the surrogate model.

Here, when MF surrogates such as scaling function-based methods (e.g., the comprehensive approach [19]) are used instead of the HK model mainly employed in this study, the importance of the basis function can be interpreted as the importance of the trend function.

To better understand the proposed insights, Fig. 2 illustrates surrogate models using various LF data sources. Fig. 2 clearly shows that the performance of surrogate models can vary, even with the same set of HF samples, depending on the quality of LF data sources (i.e., basis vectors).

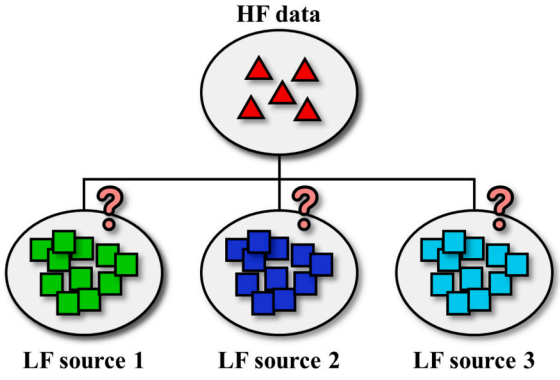


Fig. 1. Data collected from various source domains.

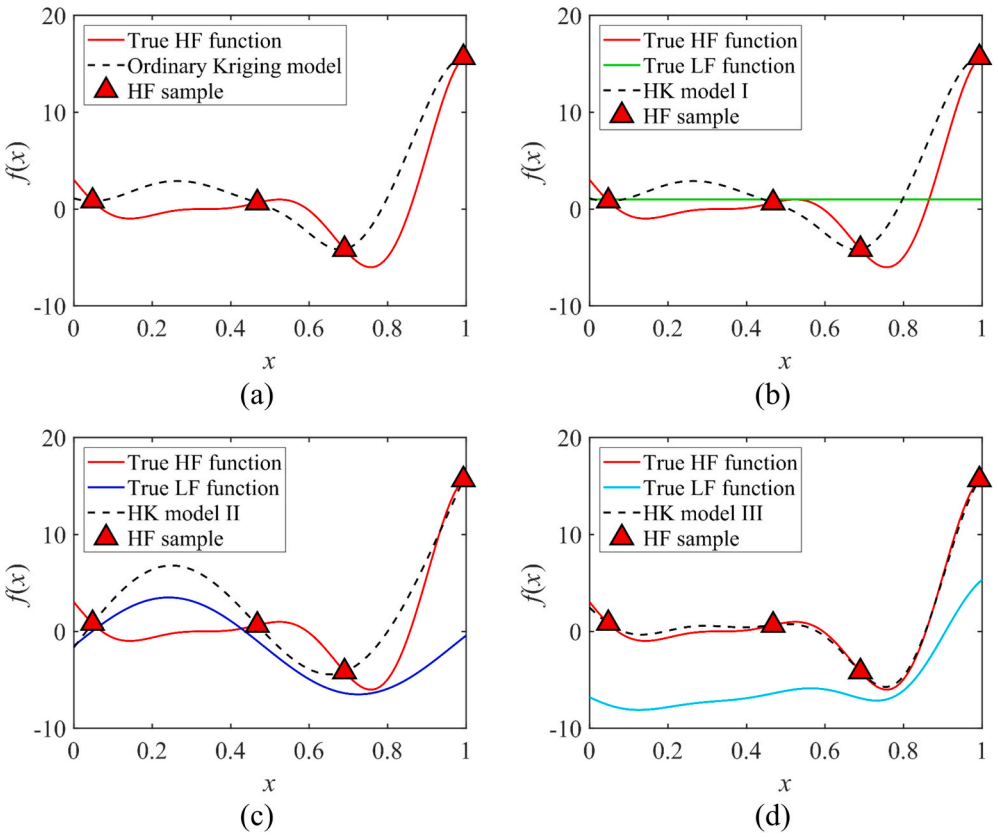


Fig. 2. Visualization of a 1D analytical function using (a) ordinary Kriging model, (b) HK model I with $y_{LF} = 1$, (c) HK model II with poor quality basis, and (d) HK model III with good quality basis.

Table 1
Performance results of illustrative examples.

	LF model	ϵ_{NRMSE}
Fig. 2(a)	–	1.381×10^{-1}
Fig. 2(b)*	$y_{LF,1} = 1$	1.381×10^{-1}
Fig. 2(c)*	$y_{LF,2} = 5\sin(6.5x) - 1.5$	1.859×10^{-1}
Fig. 2(d)*	$y_{LF,3} = 0.4y_{HF} + 10(0.4x^4 + 0.1x^2 + 0.2x + 0.2) - 10$	1.856×10^{-2}

* The HF model used in Fig. 2 is $y_{HF} = (6x-2)^2\sin(12x-4)$.

Table 1 lists the related LF surrogates and values of normalized root mean-squared error (NRMSE). When comparing Figs. 2(a) & (b), the two models appear to have similar shapes, and Table 1 confirms that their accuracy is identical. This indicates that in the HK model I, when the LF model has $y_{LF,1} = 1$, it is consistent with the ordinary Kriging model. In other words, it is worth noting once again that ordinary Kriging is a special case of HK with a basis where $y_{LF,1} = 1$. Fig. 2(c) and Table 1 demonstrate a decrease in the accuracy of the HK model II when a poor quality LF model ($y_{LF,2}$) is used as the basis. Conversely, when a relatively good quality LF model ($y_{LF,3}$) serves as the basis, the accuracy of the HK model III appears to increase, as shown in Fig. 2(d) and Table 1. Therefore, the basis functions can be considered important in the order of $y_{LF,3}$, $y_{LF,1}$, and $y_{LF,2}$, indicating that the LF data sources are of high quality in the same order. Furthermore, when the comprehensive approach is applied instead of the HK method, the NRMSE values for the four cases in Fig. 2 are 0.1381, 0.1381, 0.15509, and 0.09906, respectively, and the importance ranking of trend functions remains consistent with the HK model results. In summary, the accuracy of the MF surrogate can fluctuate based on the quality of the LF surrogate used as the basis function, highlighting the importance of high-quality basis vectors in surrogate modeling. In light of these insights, this study approaches the quality of LF data from the perspective of the importance of basis functions in MF surrogates, developing strategies for their utilization and combination. The detailed strategies will be explained in Section 3.2.

3.2. Main concept of the proposed method

The main concept of the proposed AQBMF method is to maximize the performance of the surrogate by appropriately utilizing multiple LF data sources, whose fidelity levels are not known in advance. This study is based on the insights from Section 3.1, which interprets the quality levels of LF data sources as the relative importance of basis functions in the MF surrogate model. The proposed AQBMF surrogate framework consists of four main stages. In the first stage, the proposed augmented MF surrogate formulation is constructed by integrating all available LF information. The second stage involves estimating the importance of LF surrogates using the proposed MF basis screening method. In the third stage, based on the relative importance of LF surrogates, promising candidate surrogates are systematically generated. Finally, in the fourth stage, the proposed surrogate selection algorithm chooses the best surrogate from the candidates. The detailed strategies for each stage will be explained in the following sections. Additionally, it should be noted that the term ensemble method in this study refers to weight-based ensemble combination in Eq. (3), rather than ensemble learning, which involves combining different types of surrogate models.

3.2.1. Augmented MF formulation

The first stage constructs an augmented MF formulation based on HK theory. The core idea of the proposed formulation is to initially assume that all basis functions are equally important, enabling the representa-

tion of both SF and MF surrogates. To incorporate the SF surrogate, the proposed formulation adopts the classic polynomial basis functions. Furthermore, to handle multiple LF data sources, an ensemble MF combination technique is employed. Reflecting these ideas, the proposed augmented MF formulation is defined as

$$\begin{aligned} Y(\boldsymbol{x}) &= \sum_{k=1}^{p+N} \beta_{\text{aug},k} f_{\text{aug},k}(\boldsymbol{x}) + Z(\boldsymbol{x}) \\ &= \boldsymbol{f}_{\text{aug}}(\boldsymbol{x})^{\text{T}} \boldsymbol{\beta}_{\text{aug}} + Z(\boldsymbol{x}) \end{aligned} \tag{4}$$

where p is the number of used polynomial basis functions, $\boldsymbol{f}_{\text{aug}}(\boldsymbol{x})$ is an expanded design matrix, and $\boldsymbol{\beta}_{\text{aug}}$ is a coefficient vector of the corresponding basis functions. Note that $\boldsymbol{f}_{\text{aug}}(\boldsymbol{x}) = \left[f_1(\boldsymbol{x}), \ldots, f_p(\boldsymbol{x}), \widehat{y}_{\text{LF},1}(\boldsymbol{x}), \ldots, \widehat{y}_{\text{LF},N}(\boldsymbol{x})\right]^{\text{T}}$, which contains the polynomial and LF functions for SF and MF surrogates, respectively.

Suppose that the inputs of LF and HF data are denoted as $\boldsymbol{S}_{\text{LF}} \in \mathbb{R}^{n_{\text{LF}} \times dm}$ and $\boldsymbol{S}_{\text{HF}} \in \mathbb{R}^{n_{\text{HF}} \times dm}$, respectively, and corresponding outputs are $\boldsymbol{Y}_{\text{LF}} \in \mathbb{R}^{n_{\text{LF}}}$ and $\boldsymbol{Y}_{\text{HF}} \in \mathbb{R}^{n_{\text{HF}}}$, respectively, where n_{LF}, n_{HF}, and dm are the number of LF data, HF data, and dimensions, respectively. Then, the vector form representing random response $Y(\boldsymbol{S}_{\text{HF},i})$ at n_{HF} sampled inputs in Eq. (4) can be expressed as

$$\boldsymbol{Y} = \boldsymbol{F}_{\text{aug}} \boldsymbol{\beta}_{\text{aug}} + \boldsymbol{Z} \tag{5}$$

where

$$\boldsymbol{Y} = \begin{bmatrix} Y(\boldsymbol{S}_{\text{HF},1}) \\ Y(\boldsymbol{S}_{\text{HF},2}) \\ \vdots \\ Y(\boldsymbol{S}_{\text{HF},n_{\text{HF}}}) \end{bmatrix}, \boldsymbol{F}_{\text{aug}} = \begin{bmatrix} f_{\text{aug}}(\boldsymbol{S}_{\text{HF},1})^{\text{T}} \\ f_{\text{aug}}(\boldsymbol{S}_{\text{HF},2})^{\text{T}} \\ \vdots \\ f_{\text{aug}}(\boldsymbol{S}_{\text{HF},n_{\text{HF}}})^{\text{T}} \end{bmatrix}, \boldsymbol{Z} = \begin{bmatrix} Z(\boldsymbol{S}_{\text{HF},1}) \\ Z(\boldsymbol{S}_{\text{HF},2}) \\ \vdots \\ Z(\boldsymbol{S}_{\text{HF},n_{\text{HF}}}) \end{bmatrix}, \tag{6}$$

and $\boldsymbol{F}_{\text{aug}}$ is defined as $(n_{\text{HF}} \times (p+N))$ expanded design matrix.

From Eq. (4), the predictor and its MSE are derived by minimizing the MSE of the stochastic process while satisfying the unbiasedness constraint. It is worth noting that the proposed formulation has the same structure as the standard Kriging or HK. The only difference lies in the augmentation of the existing expanded design matrix. Therefore, it is possible to directly apply the same procedure as in Kriging or HK theory to Eq. (4). The predictor and its MSE for the proposed formulation at $\boldsymbol{x}$ can be formulated as

$$\widehat{y}(\boldsymbol{x}) = \boldsymbol{f}_{\text{aug}}(\boldsymbol{x})^{\text{T}} \widehat{\boldsymbol{\beta}}_{\text{aug}} + \boldsymbol{r}(\boldsymbol{x})^{\text{T}} \boldsymbol{R}^{-1} \left(\boldsymbol{Y}_{\text{HF}} - \boldsymbol{F}_{\text{aug}} \widehat{\boldsymbol{\beta}}_{\text{aug}}\right) \tag{7}$$

and

$$\begin{aligned} s^2(\boldsymbol{x}) = \widehat{\sigma}^2 \Bigg[& 1 - \boldsymbol{r}(\boldsymbol{x})^{\text{T}} \boldsymbol{R}^{-1} \boldsymbol{r}(\boldsymbol{x}) \\ & + \left(\boldsymbol{F}_{\text{aug}}^{\text{T}} \boldsymbol{R}^{-1} \boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{f}_{\text{aug}}(\boldsymbol{x})\right)^{\text{T}} \left(\boldsymbol{F}_{\text{aug}}^{\text{T}} \boldsymbol{R}^{-1} \boldsymbol{F}_{\text{aug}}\right)^{-1} \left(\boldsymbol{F}_{\text{aug}}^{\text{T}} \boldsymbol{R}^{-1} \boldsymbol{r}(\boldsymbol{x}) \right. \\ & \left. - \boldsymbol{f}_{\text{aug}}(\boldsymbol{x})\right) \Bigg] \end{aligned} \tag{8}$$

where $\widehat{\boldsymbol{\beta}}_{\text{aug}} = \left(\boldsymbol{F}_{\text{aug}}^{\text{T}} \boldsymbol{R}^{-1} \boldsymbol{F}_{\text{aug}}\right)^{-1} \boldsymbol{F}_{\text{aug}}^{\text{T}} \boldsymbol{R}^{-1} \boldsymbol{Y}_{\text{HF}}$, $\widehat{\sigma}^2 = \left\{\left(\boldsymbol{Y}_{\text{HF}} - \boldsymbol{F}_{\text{aug}} \widehat{\boldsymbol{\beta}}_{\text{aug}}\right)^{\text{T}} \boldsymbol{R}^{-1} \left(\boldsymbol{Y}_{\text{HF}} - \boldsymbol{F}_{\text{aug}} \widehat{\boldsymbol{\beta}}_{\text{aug}}\right)\right\} / n_{\text{HF}}$, $\boldsymbol{R}$ is the $(n_{\text{HF}} \times n_{\text{HF}})$ correlation matrix between observed HF input points, and $\boldsymbol{r}(\boldsymbol{x})$ is the $(n_{\text{HF}} \times 1)$ correlation vector between untried and sampled HF input points. Here, when $N = 0$, the proposed formulation reduces to the standard SF Kriging with only polynomial basis functions. On the other hand, when $p = 0$, it corresponds to EHK, which combines all LF data in an ensemble manner. Therefore, the proposed augmented formulation makes it possible to flexibly handle both SF and MF surrogates depending on the selection of basis functions.

3.2.2. MF basis screening

In the second stage, the proposed MF basis screening method is performed to determine the quality (i.e., importance) of each LF data source. The proposed method assesses the individual influence of each basis function of the MF surrogate. Specifically, it integrates the polynomial and LF basis functions collected in Section 3.2.1 to generate the augmented vector of basis functions (i.e., $\boldsymbol{f}_{\text{aug}}(\boldsymbol{x})$), followed by estimating the importance of each individual basis function. The impact of each basis function is estimated using a new metric called the ranking measure (RM). In the proposed measure, NCVE is used to estimate the accuracy of the surrogate constructed with the selected basis functions, drawing inspiration from the basis selection method described in Section 2.2. In other words, a lower NCVE indicates a more accurate surrogate, whereas a higher NCVE implies a less accurate one. Additionally, the PCC is employed to evaluate how effectively the selected basis functions contribute meaningful LF correlation for building the MF surrogate. As discussed in Sections 2.1 and 3.1, the selected basis functions serve to represent the general trend of the MF surrogate. Therefore, as reported in previous studies [19,33], a higher PCC value indicates a more desirable basis function for constructing an MF surrogate, while a lower PCC suggests that the basis function is less suitable for use. Therefore, the proposed RM consists of NCVE and PCC indicators, which can be defined as

$$I_\alpha = \frac{\varepsilon_{\text{NCVE},\alpha}}{\left|\widetilde{r}_\alpha\right|} \tag{9}$$

where I_α, $\varepsilon_{\text{NCVE},\alpha}$, and $\widetilde{r}_\alpha$ are the RM, NCVE, and PCC values of the MF surrogate model built exclusively with each α^{th} basis function, respectively, and related formulations are provided in Appendix A. In general, the smaller $\varepsilon_{\text{NCVE},\alpha}$ and the larger $\left|\widetilde{r}_\alpha\right|$, the greater is the impact of the corresponding basis function. Therefore, a smaller RM indicates higher importance for the corresponding basis function, while a larger RM suggests lower importance. For example, consider $N = 3$ and $p = 1$ with $f_p(\boldsymbol{x}) = 1$, forming $\boldsymbol{f}_{\text{aug}}(\boldsymbol{x}) = \left[1, \widehat{y}_{\text{LF},1}(\boldsymbol{x}), \widehat{y}_{\text{LF},2}(\boldsymbol{x}), \widehat{y}_{\text{LF},3}(\boldsymbol{x})\right]^{\text{T}}$. If the RM value is highest for $\widehat{y}_{\text{LF},3}(\boldsymbol{x})$ and lowest for $\widehat{y}_{\text{LF},2}(\boldsymbol{x})$, then the importance of the basis functions is ranked in the order of $\widehat{y}_{\text{LF},2}(\boldsymbol{x})$ as most important, followed by $\widehat{y}_{\text{LF},1}(\boldsymbol{x})$, and finally $\widehat{y}_{\text{LF},3}(\boldsymbol{x})$.

3.2.3. Establishment of candidate MF surrogates

In the third stage, multiple candidate surrogates are constructed based on the estimated ranking of LF surrogates obtained in Section 3.2.2. The candidate MF surrogates include SF, hierarchical-based MF, and ensemble-based MF surrogates. To identify the best surrogate, metrics such as NCVE, PCC, and RM are used for each candidate, and their formulas are provided in Appendix A.

Firstly, the candidate SF surrogate is created using only HF data with a polynomial basis function of 1. In other words, an ordinary Kriging model with $f_{\text{aug}}(\boldsymbol{x}) = 1$, as described in Eq. (1), is employed. This surrogate does not utilize any of the collected LF data sources. If all LF data sources are ultimately deemed unnecessary, this surrogate will be selected as the final model.

Secondly, the candidate hierarchical combination-based MF surrogate is constructed using HF data, LF data, and Eq. (2). In this process, fidelity levels are treated as the importance of the basis functions, as estimated in Section 3.2.2. Therefore, LF surrogate is created using data with the lowest fidelity level, and each subsequent surrogate is refined using the next higher fidelity data. This process continues until the highest fidelity data (i.e., HF data) is incorporated, resulting in the candidate hierarchical MF surrogate. For example, assume that LF source 2 ($\widehat{y}_{\text{LF},2}(\boldsymbol{x})$), LF source 1 ($\widehat{y}_{\text{LF},1}(\boldsymbol{x})$), and LF source 3 ($\widehat{y}_{\text{LF},3}(\boldsymbol{x})$) are ranked by importance, with LF source 2 as the most important and LF source 3 as the least important. Then, an SF surrogate is first created using LF source 3, and it is subsequently updated hierarchically in the

order of LF source 1, LF source 2, and finally HF data, resulting in the construction of the candidate hierarchical MF surrogate.

Thirdly, the candidate ensemble combination-based MF surrogates are built using HF data, LF data, and Eq. (4). In this process, LF data sources, which are treated as basis functions, are added one by one according to the basis ranking, resulting in the creation of candidate MF surrogates. Specifically, the polynomial basis function 1 is fixed, and the remaining basis functions are added in order of their importance, resulting in the creation of multiple candidate surrogates. Therefore, excluding the polynomial basis function 1, if there are n basis functions, n candidate ensemble combination-based MF surrogates are constructed. Naturally, the last candidate ensemble combination-based MF surrogate corresponds to the model using all the basis functions. For example, if the basis functions are ranked in importance as LF source 2 ($\widehat{y}_{\mathrm{LF},2}(\boldsymbol{x})$), LF source 1 ($\widehat{y}_{\mathrm{LF},1}(\boldsymbol{x})$), and LF source 3 ($\widehat{y}_{\mathrm{LF},3}(\boldsymbol{x})$), then three candidate ensemble combination-based MF surrogates can be created. In this case, the expanded design matrices $\boldsymbol{f}_{\mathrm{aug}}(\boldsymbol{x}) = \left[1, \widehat{y}_{\mathrm{LF},2}(\boldsymbol{x})\right]^{\mathrm{T}}$, $\boldsymbol{f}_{\mathrm{aug}}(\boldsymbol{x}) = \left[1, \widehat{y}_{\mathrm{LF},2}(\boldsymbol{x}), \widehat{y}_{\mathrm{LF},1}(\boldsymbol{x})\right]^{\mathrm{T}}$, and $\boldsymbol{f}_{\mathrm{aug}}(\boldsymbol{x}) = \left[1, \widehat{y}_{\mathrm{LF},2}(\boldsymbol{x}), \widehat{y}_{\mathrm{LF},1}(\boldsymbol{x}), \widehat{y}_{\mathrm{LF},3}(\boldsymbol{x})\right]^{\mathrm{T}}$ in Eq. (4) are constructed and corresponding three candidate ensemble combination-based MF surrogates are generated.

Fig. 3 visually compares the conventional SF, ensemble combination-based MF, and the proposed MF surrogate frameworks through illustrative diagrams involving three LF data sources with unknown fidelity levels. Fig. 3(a) shows the conventional SF surrogate uses only polynomials and HF data. In addition, Fig. 3(b) describes the conventional ensemble combination-based MF surrogate, which combines all three LF data sources. However, as shown in Fig. 3(c), the proposed method first estimates the fidelity levels of the LF data sources and generates candidate SF, hierarchical combination-based MF, and ensemble combination-based MF surrogates. Afterwards, in Stage 4, the proposed algorithm is applied to select the best model from the candidate surrogate models.

3.2.4. Best surrogate selection

The final stage involves the selection of the best surrogate from multiple candidate surrogates built in Section 3.2.3. The best surrogate is determined based on the PCC, NCVE, and RM values. Specifically, the strategy for selecting surrogates is categorized according to the PCC, which represents the correlation between LF and HF systems. If at least one surrogate's PCC exceeds a pre-defined threshold, it indicates favorable conditions for MF surrogate creation and prompts an aggressive strategy for adaptively selecting LF basis functions. In this case, the best surrogate is determined by comparing the RM values of the selected candidate MF surrogates that meet the PCC criterion. However, if all surrogates' PCC values are lower than a pre-defined threshold, a conservative strategy is adopted, selecting between models that use either all or none of the LF basis functions. In this scenario, the best surrogate is evaluated by comparing the NCVE value of the candidate SF surrogate with the RM values of the two candidate MF surrogates using all LF basis functions. This environment is not conducive to MF surrogate creation,

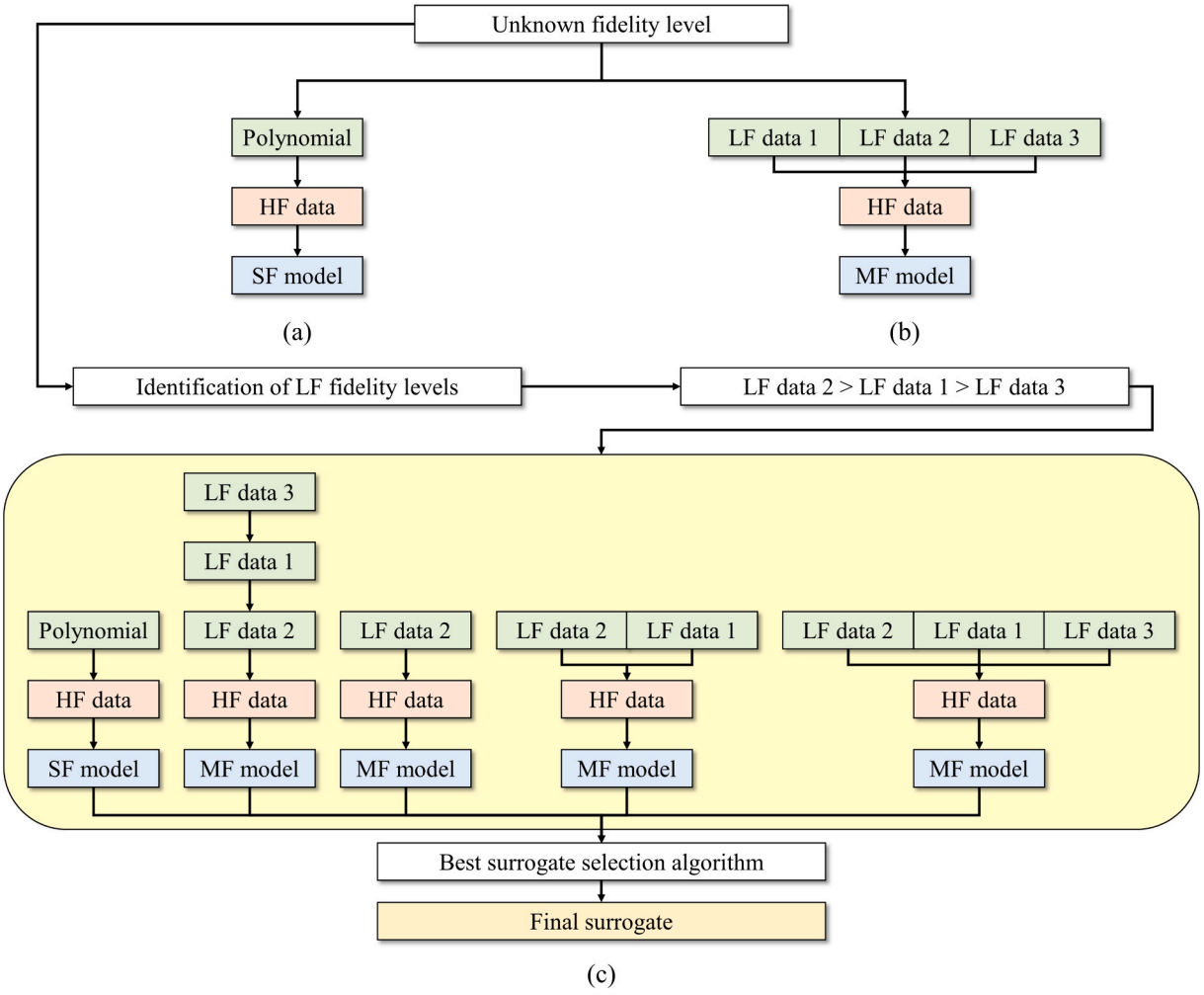


Fig. 3. Illustrative diagrams of three LF data sources with unknown fidelity levels: (a) conventional SF, (b) conventional MF, and (c) proposed MF surrogate frameworks.

therefore the MF surrogate employs RM, which applies a penalty for PCC in addition to NCVE. Since the absolute value of PCC is always less than or equal to 1, the MF surrogate using RM is less likely to be selected compared to the SF surrogate using NCVE. In particular, as the PCC decreases, the RM increases, thereby lowering the likelihood of selecting the MF surrogate. Conversely, if the RM value of the MF surrogate is lower than the NCVE of the SF surrogate despite unfavorable conditions for MF surrogate creation, it suggests that choosing the MF surrogate would be a reasonable decision. Motivated by this concept, the proposed method is developed, the algorithm of which can be summarized as follows:

Step 1: Select the MF surrogate models from multiple candidates whose PCC is equal to or greater than the pre-defined PCC threshold $r_{\min}$.
Step 2: If there are selected MF surrogates, choose the one with the lowest RM value among them as the best surrogate model.
Step 3: Otherwise, set the comparison metrics as the NCVE for the SF surrogate and the RM for the hierarchical and ensemble combination-based MF surrogates. In this step, both MF surrogates are constructed using all basis functions. Then, select the model with the lowest value among these comparison indicators as the best surrogate model.

3.3. Validation metric

To assess the global accuracy of the surrogates, the NRMSE is adopted in this study. If there is a dataset of n_{test} pairs of input and output test data, denoted as $\boldsymbol{x}_{\text{test}}$ and $\boldsymbol{y}_{\text{test}}$, respectively, then the NRMSE, $\varepsilon_{\text{NRMSE}}$, can be expressed as

$$\varepsilon_{\text{NRMSE}} = \frac{\sqrt{\frac{1}{n_{\text{test}}}\sum_{i=1}^{n_{\text{test}}}\left(y\left(\boldsymbol{x}_{\text{test}}^{(i)}\right)-\widehat{y}\left(\boldsymbol{x}_{\text{test}}^{(i)}\right)\right)^2}}{\max\left(\boldsymbol{y}_{\text{test}}\right)-\min\left(\boldsymbol{y}_{\text{test}}\right)} \tag{10}$$

where $\boldsymbol{x}_{\text{test}}^{(i)}$ is the i^{th} $\boldsymbol{x}_{\text{test}}$; $y\left(\boldsymbol{x}_{\text{test}}^{(i)}\right)$ and $\widehat{y}\left(\boldsymbol{x}_{\text{test}}^{(i)}\right)$ are the true and predicted output values of $\boldsymbol{x}_{\text{test}}^{(i)}$, respectively.

3.4. Overall procedure

In this section, the workflow of the proposed AQBMF surrogate framework is illustrated in Fig. 4, and its algorithm is summarized as follows:

- Step 1: Generate the inputs for the HF system and each LF system via design of experiments (DoE). In this study, a well-known Latin hypercube sampling (LHS) is used.
- Step 2: Obtain the HF and LF outputs based on DoE results.
- Step 3: Build LF surrogate models for each data source.
- Step 4: Construct the augmented MF formulation to create an integrated basis function set that includes both the polynomial and LF basis functions.
- Step 5: Estimate the importance of individual basis functions using the MF basis screening method. In this step, the RM, which is composed of PCC and NCVE, is used.
- Step 6: Build all candidate surrogates based on the ranking of the basis functions. In this step, SF, hierarchical-based MF, and ensemble combination-based MF surrogates are constructed and the corresponding metrics—PCC, NCVE, and RM—are also calculated.
- Step 7: Decide on the final surrogate among the candidate surrogates using the best surrogate selection algorithm.
- Step 8: Terminate the overall process.

3.5. Discussion about the pros and cons of the proposed method

In this section, the advantages and disadvantages of the proposed method compared to existing methods are analyzed. The discussion can be summarized as follows:

(1) The proposed AQBMF surrogate method is recommended for situations where LF information from various sources is available, but it is uncertain which LF data will be beneficial or which combination method will be most effective. If prior knowledge about the use of LF data sources is available, the proposed method may be somewhat inefficient. However, in real engineering problems, it is generally impossible to know in advance which data sources and MF combination methods will be most suitable,

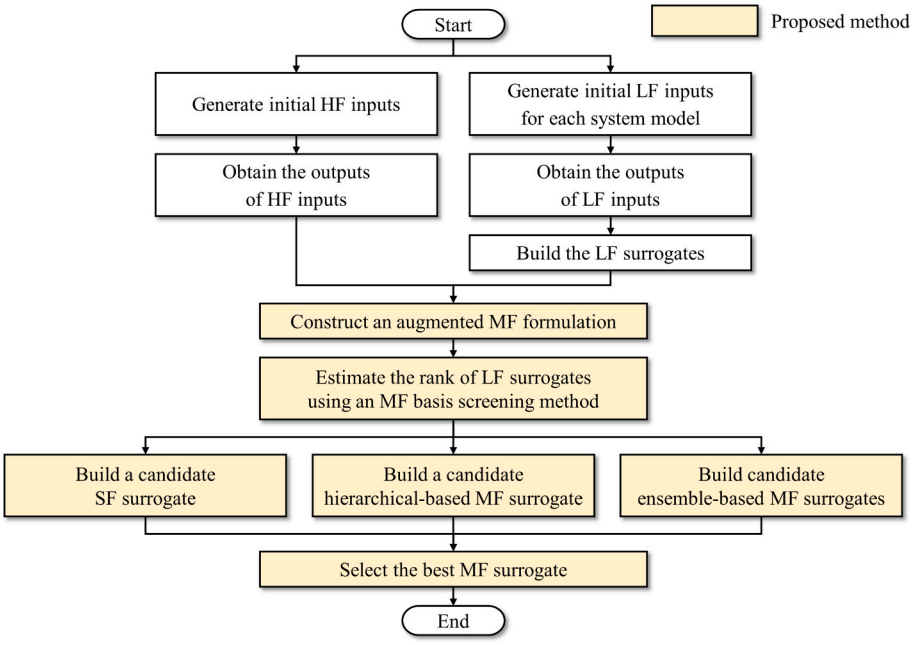


Fig. 4. Flowchart of the proposed AQBMF surrogate framework.

making the proposed method highly useful for addressing such issues.

(2) The proposed strategy becomes more effective as the number of LF data sources increases, compared to the approach of generating all possible candidate surrogates, as demonstrated in Appendix B. When all possible combinations are considered, the number of candidate surrogates that need to be created increases exponentially with respect to the number of LF sources, leading to computational times that become nearly prohibitive. In contrast, the proposed method increases linearly with respect to the number of LF sources, making it significantly more efficient.

(3) The proposed method is a flexible method that considers cases ranging from not using any LF data sources to using all of them. In particular, existing studies on creating MF surrogates using ensemble combination-based LF surrogates [27,30,31] typically assume that the sum of the weights for the LF surrogates equals one, implying that at least one LF data source must always be beneficial. However, sometimes LF data from all sources may not be useful, and the proposed method can handle such situations as well.

(4) The proposed framework has the advantage of being highly extensible, as it can be easily integrated with existing techniques or advanced methods yet to be developed. For example, it can be extended to various data-driven modeling techniques beyond the Gaussian process surrogates used in this study. If the surrogate employed is a regression model (e.g., neural networks; see [41,42]) rather than an interpolation model (e.g., Kriging), metrics such as the coefficient of determination (R^2) or NRMSE can be used instead of NCVE for ranking LF surrogates. In such cases, the concept of basis importance could be interpreted as the importance of trend functions. Moreover, in scenarios involving image data from multiple sources, the proposed framework can be combined with generative artificial intelligence (AI) models to construct more flexible and high-performing systems [43,44]. Within the proposed framework, fidelity levels can be interpreted as labels, and the AI model architecture can be adapted accordingly to a hierarchical or ensemble form. Data quality can be accounted for by adjusting the weights of the loss functions, allowing lower-quality data to have less influence during training.

(5) The performance of the proposed method has the drawback of being indirectly influenced by the performance of the adopted MF methodology. This is because generating an inaccurate MF surrogate can result in an unreliable estimate of the importance of LF data sources. Therefore, if a state-of-the-art MF surrogate is developed, the performance of the proposed method is expected to improve further.

4. Results

In this section, the effectiveness of the proposed AQBMF surrogate method is evaluated using numerical tests, comparing its performance to existing approaches. In the first numerical example, a demonstrative example is used to verify that the proposed method works effectively. In the second numerical example, the performance of the proposed algorithm is assessed using benchmark test functions with various characteristics. The third numerical example illustrates the practical utility of the proposed method by applying it to a real-world engineering problem. Particularly, it is important to note that all problems assume a common scenario in industrial environments where the fidelity levels among LF data sources are initially unknown. For all numerical examples, the proposed method is compared to three existing approaches:

- **Method 1**: Use only HF data without utilizing any LF data sources (e.g., Ordinary Kriging).
- **Method 2**: Combine all available LF data in an ensemble manner (e.g., EHK [29], NHLF-Co-Kriging [28], and WSMHK-OWSdiag [31]). More specifically, these three methods are implemented as comparative ensemble-based MF approaches, referred to as Method 2–1 (EHK), Method 2–2 (NHLF-Co-Kriging), and Method 2–3 (WSMHK-OWSdiag), respectively.
- **Method 3**: Determine whether to use the single LF data source based on the conventional MF dataset selection algorithm provided in Appendix C (e.g., Ordinary Kriging or HK).
- **Proposed**: Utilize LF information by adaptively selecting LF data sources and MF combination methods (e.g., Ordinary Kriging, HK, or EHK).

Here, it is worth noting once again that the term "ensemble method" in this study specifically means the weight-based ensemble combination in Eq. (3), rather than ensemble learning, which integrates different types of surrogates. During the surrogate creation, initial sample points are generated using LHS with a maximin criterion through 10^3 iterations [4,12]. To validate the accuracy of the surrogate models, test samples generated by LHS are used, equal to 200 times the number of design variables. Details on the implementation of the surrogate construction are provided in Appendix A. In low-dimensional cases, a surrogate-based modified PCC [16] is employed. In high-dimensional cases (e.g., when the dimensionality is greater than 10), due to the curse of dimensionality, a nested DoE [11,16] is generated and the conventional sampling-based PCC [19,33] is applied. Furthermore, the threshold parameter $r_{\min}$ is set to 0.85, referring to previous studies [16,33,45]. The entire computational process is run on a PC equipped with 12th Gen Intel® Core™ i7-12700 K and 32 GB RAM.

4.1. Demonstrative example

The aim of this section is to illustrate the process of the proposed AQBMF framework through a demonstrative example. The formulation of an adopted 1D function can be expressed as

$$\begin{aligned} y_{\mathrm{HF}} &= \sin(x) + 0.2x + (x-5)^2 \big/ 16 + 0.5, \\ y_{\mathrm{LF},1} &= (x-0.5)(x-4)(x-9) \big/ 20 + 2, \\ y_{\mathrm{LF},2} &= \sin(x) + 0.2x + 0.5, \\ & 0 \leqslant x \leqslant 10 \end{aligned} \tag{11}$$

which includes one HF function (y_{HF}) and two LF functions ($y_{\mathrm{LF},1}$ and $y_{\mathrm{LF},2}$). The observed HF and LF sample points are located at $[0.0808, 3.101, 6.037, 8.983]^{\mathrm{T}}$ and $[0.0, 0.5, 1.0, \ldots, 9.5, 10.0]^{\mathrm{T}}$, respectively, as shown in Fig. 5. All LF models share the same sample points. Therefore, the datasets D_{HF}, $D_{\mathrm{LF},1}$, and $D_{\mathrm{LF},2}$ are constructed, with the subscript indicating the source of each dataset. In this problem setting, the proposed method is compared with Methods 1, 2–1, and 3.

Fig. 5 presents visualizations of the results for the conventional and the proposed methods. The surrogates generated by Methods 1, 2–1, and 3 show noticeable discrepancies from the true function. In contrast, the surrogate created using the proposed method closely approximates the true HF function. In addition, detailed quantitative results comparing the existing and proposed methods are presented in Table 2. As previously mentioned, Method 1 uses only HF data, while Method 2–1 combines LF data in an ensemble manner and then calibrates it with HF data. Method 3 is equivalent to Method 1 because the traditional MF dataset selection algorithm does not select any LF data. This happens because the number of HF data points is limited, making it difficult to accurately calculate the PCC in the conventional MF dataset selection algorithm. However, unlike existing methods, the proposed approach begins by treating the two LF data sources as equally important in Stage 1. In Stage 2, it evaluates their relative significance and identifies $y_{\mathrm{LF},2}$ as a more influential LF data source – serving as a basis function – than $y_{\mathrm{LF},1}$. Based on this ranking, Stage 3 generates a set of candidate

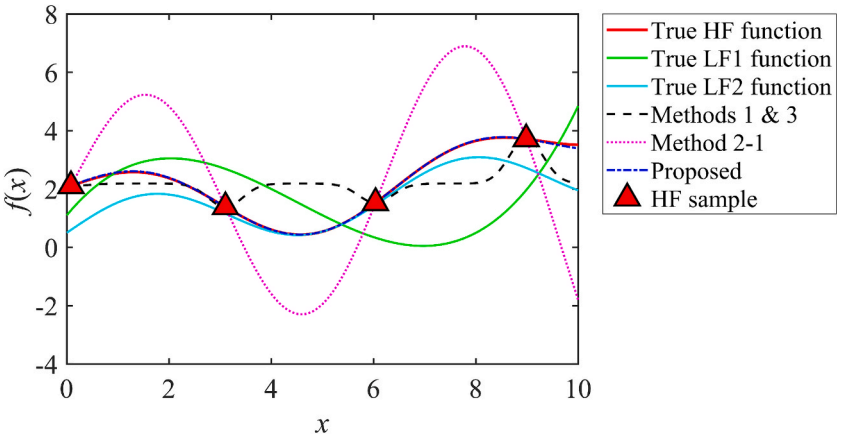


Fig. 5. Comparison of the plot results for surrogates created with four different approaches.

surrogates, which also includes ensemble combination-based MF surrogates similar to those used in Method 2–1. Finally, Stage 4 selects the best-performing surrogate, which in this case corresponds to the one constructed using the hierarchical combination strategy. Based on the explicit analytical expressions in Eq. (11), it can be intuitively inferred that $y_{LF,2}$ is more advantageous for y_{HF} than $y_{LF,1}$, which is consistent with the results of the proposed method. Therefore, unlike Method 2–1, which uniformly combines all LF sources, the AQBMF framework adaptively prioritizes and combines LF data based on estimated quality, thereby preventing performance degradation. However, such information is typically unknown beforehand in black-box systems, and the proposed method demonstrates robust performance even in the absence of prior knowledge. In other words, these findings imply that the proposed algorithm can provide practical guidelines for effectively utilizing multiple LF data sources with unknown fidelity levels. In the next section, the superiority of the proposed AQBMF method will be evaluated through various numerical examples.

4.2. Benchmark test examples

In this section, the performance of the proposed AQBMF framework is evaluated using benchmark test functions under various conditions, such as system correlations, dimensionality, the number and locations of HF samples, the number and locations of LF samples, and the number of LF data sources. Details of the adopted benchmark test functions are provided in Table 3. The entire process is repeated 50 times with different initial sample sets to account for the random effects of the DoE. It is also important to note that these numerical tests assume a common scenario where it is not known in advance which LF data will be beneficial and which will not be. The numerical results for the various scenarios are presented in Fig. 6 and Tables 4–9. In addition, ρ represents the proportion of utilized LF data sources relative to the total available LF data sources, and $\mathbf{M}_{sel}$ denotes the selection count vector for each candidate surrogate model. The upper bars represent the average values of the corresponding variables (e.g., $\overline{\varepsilon}_{NRMSE}$, $\overline{\rho}$).

Table 3
Information on the adopted problems.

Label	Name*	dm	n_{HF}	n_{LF} of each source
P1	Test function 1	1	4/5/6	(20, 20)
P2	Test function 2	1	4/5/6	(20, 20)
P3	Test function 3	1	4/5/6	(20,20)
P4	Test function 4	2	10	(40, 40)/(40, 20)/ (20, 40)
P5	Test function 5	4	16	(60, 60, 120, 120)
P6	Test function 6	6	24	(120, 120, 120)
P7	Test function 7	8	32	(160, 160)
P8	Test function 8	12	48	(120, 240)

* Detailed expressions of the benchmark test functions can be found in Appendix D.

Firstly, the results for P1 are presented in Fig. 6 and Table 4. In these problems, it is assumed that n_{LF} for each source is sufficient, making the LF models fairly accurate. The objective of this test is to compare the accuracy of the six approaches as n_{HF} increases and the HF sample locations change. As shown in Fig. 6(a), when n_{HF} is 4, Method 1 outperforms Methods 2–1 and 2–2, but as n_{HF} increases to 5 and 6, as shown in Figs. 6(b) & (c), Methods 2–1, 2–2, and 2–3 becomes more accurate than Method 1. Method 3 utilizes very little LF information across all three cases, resulting in outcomes that are nearly identical to those of Method 1. However, the proposed method consistently provides more accurate results than both Method 1 and all three Method 2 approaches across all cases. A detailed analysis indicates that the proposed method identifies $y_{LF,2}$ as more informative than $y_{LF,1}$ in constructing MF surrogates across all 50 repetitions. Notably, the calculated value of $\overline{\rho}$ after applying the proposed algorithm shows that LF sources are utilized in more than 75 % of the cases. In particular, the number of times each candidate surrogate model ($\mathbf{M}_{sel}$) was selected is summarized as follows. The first column of $\mathbf{M}_{sel}$ represents models using only HF data. The second column corresponds to models using only $y_{LF,2}$. The third column represents ensemble combination-based MF models using both $y_{LF,2}$ and $y_{LF,1}$, and the fourth column corresponds to hierarchical combination-

Table 2
Performance comparison results for the demonstrative example.

	n_{HF}*	Importance ranking**	Final MF combination***	ε_{NRMSE}
Method 1	4	–	D_{HF}	2.759×10^{-1}
Method 2–1	4	–	$(D_{LF,1} + D_{LF,2}) \rightarrow D_{HF}$	6.616×10^{-1}
Method 3	4	–	D_{HF}	2.759×10^{-1}
Proposed	4	$(D_{LF,2}, D_{LF,1})$	$D_{LF,1} \rightarrow D_{LF,2} \rightarrow D_{HF}$	6.835×10^{-3}

* Since all LF information is assumed to be given, all n_{HF} values are the same.
** LF data sources are listed in order of importance from the most important (left) to the least important (right).
*** Final combination process is represented, where the plus (+) and arrow (→) indicate the ensemble and hierarchical combination approaches, respectively.

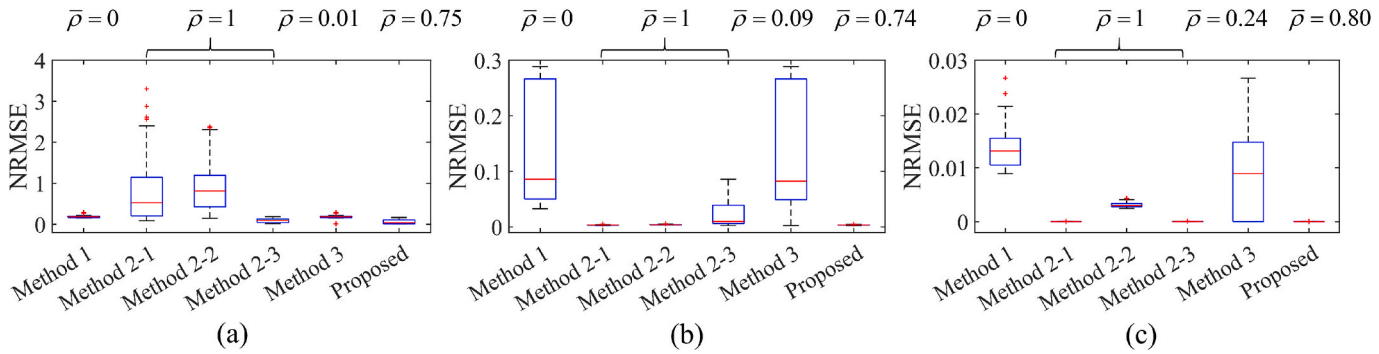


Fig. 6. Performance comparison results for P1 when (a) $n_{HF} = 4$, (b) $n_{HF} = 5$, and (c) $n_{HF} = 6$.

Table 4
Performance comparison results for P1.

	$n_{HF} = 4$		$n_{HF} = 5$		$n_{HF} = 6$	
	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$
Method 1	0.1916	0.0	0.1448	0.0	0.0137	0.0
Method 2–1	0.8439	1.0	0.0033	1.0	1.517×10^{-5}	1.0
Method 2–2	0.9233	1.0	0.0038	1.0	0.0031	1.0
Method 2–3	0.0864	1.0	0.0229	1.0	2.555×10^{-5}	1.0
Method 3	0.1864	0.01	0.1331	0.09	0.0074	0.24
Proposed	0.0591	**0.75 (0,25,2,23)***	0.0033	**0.74 (0,26,0,24)***	1.442×10^{-5}	**0.80 (0,20,2,28)***

* The number in parentheses represents the number of surrogates selected from 50 repetitions, in the following order: SF surrogate, the set of ensemble combination-based MF surrogates added one by one based on the basis ranking, and the hierarchical-based MF surrogate.

Table 5
Performance comparison results for P2.

	$n_{HF} = 4$		$n_{HF} = 5$		$n_{HF} = 6$	
	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$
Method 1	0.1776	0.0	0.1208	0.0	0.1011	0.0
Method 2–1	1.0011	1.0	0.2279	1.0	0.1519	1.0
Method 2–2	0.7502	1.0	0.3235	1.0	0.1281	1.0
Method 2–3	0.2451	1.0	0.1450	1.0	0.1163	1.0
Method 3	0.1776	0.0	0.1208	0.0	0.1011	0.0
Proposed	0.2204	**0.09 (44,3,3,0)***	0.1208	**0.0 (50,0,0,0)***	0.1107	**0.04 (47,2,1,0)***

* The number in parentheses represents the number of surrogates selected from 50 repetitions, in the following order: SF surrogate, the set of ensemble combination-based MF surrogates added one by one based on the basis ranking, and the hierarchical-based MF surrogate.

Table 6
Performance comparison results for P3.

	$n_{HF} = 4$		$n_{HF} = 5$		$n_{HF} = 6$	
	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$	$\overline{\varepsilon}_{NRMSE}$	$\overline{\rho}(\mathbf{M}_{sel})$
Method 1	0.1579	0.0	0.1135	0.0	0.0478	0.0
Method 2–1	8.886×10^{-5}	1.0	4.37×10^{-6}	1.0	4.05×10^{-6}	1.0
Method 2–2	8.881×10^{-5}	1.0	8.57×10^{-6}	1.0	1.34×10^{-5}	1.0
Method 2–3	0.759	1.0	0.0517	1.0	0.0068	1.0
Method 3	0.1554	0.47	0.1121	0.07	0.0478	0.0
Proposed	8.886×10^{-5}	**1.0 (0,0,50,0)***	4.37×10^{-6}	**1.0 (0,0,50,0)***	4.05×10^{-6}	**1.0 (0,0,50,0)***

* The number in parentheses represents the number of surrogates selected from 50 repetitions, in the following order: SF surrogate, the set of ensemble combination-based MF surrogates added one by one based on the basis ranking, and the hierarchical-based MF surrogate.

based MF models using $y_{LF,1}$, $y_{LF,2}$, and HF data sequentially. According to $\mathbf{M}_{sel}$ results, the majority of cases involve constructing MF surrogates using only $y_{LF,2}$ as the LF data source (25, 26, and 20 out of 50 cases in each respective scenario), or building MF surrogates by hierarchically combining $y_{LF,1}$, $y_{LF,2}$, and HF data in sequence (23, 24, and 28 out of 50 cases in each respective scenario). In other words, the proposed method adaptively utilizes LF models, effectively distinguishing useful LF information and achieving superior performance.

Secondly, the performance results for P2 are shown in Table 5. In this case, it is also assumed that n_{LF} is sufficient, and the six methods are evaluated as n_{HF} increases. The main feature of this problem is that neither $y_{LF,1}$ nor $y_{LF,2}$ (i.e., any LF data source) are beneficial to y_{HF}. This information is not known in advance, and the numerical results in all cases indicate that Method 1 outperforms Methods 2–1, 2–2, and 2–3. This aligns with the problem's characteristics, where it is more advantageous not to use all LF sources. Method 3 effectively differentiates the use of LF data sources and produces results nearly identical to Method 1. The proposed method, with a $\overline{\rho}$ value close to 0 in all cases, does not use any LF sources, yielding results similar to Methods 1 and 3. This shows that the proposed method selectively avoids using LF sources when they are not informative.

Thirdly, the performance results for P3 are listed in Table 6. In this case, n_{LF} is also assumed to be sufficient, and the six methods are evaluated as n_{HF} increases. A key feature of this problem is that while the

Table 7
Performance comparison results for P4.

	$dm = 2$, $n_{HF} = 10$ $n_{LF} = (40, 40)$		$dm = 2$, $n_{HF} = 10$ $n_{LF} = (40, 20)$		$dm = 2$, $n_{HF} = 10$ $n_{LF} = (20, 40)$	
	$\bar{\varepsilon}_{NRMSE}$	$\bar{\rho}(\mathbf{M}_{sel})$	$\bar{\varepsilon}_{NRMSE}$	$\bar{\rho}(\mathbf{M}_{sel})$	$\bar{\varepsilon}_{NRMSE}$	$\bar{\rho}(\mathbf{M}_{sel})$
Method 1	0.0806	0.0	0.0806	0.0	0.0806	0.0
Method 2–1	0.0161	1.0	0.0211	1.0	0.0877	1.0
Method 2–2	0.0495	1.0	0.0484	1.0	0.0877	1.0
Method 2–3	0.0339	1.0	0.0319	1.0	0.0709	1.0
Method 3	0.0244	0.47	0.0244	0.47	0.0721	0.47
Proposed	0.0163	**0.84 (0,16,25,9)***	0.0194	**0.85 (0,15,21,14)***	0.0322	**0.93 (0,7,6,37)***

* The number in parentheses represents the number of surrogates selected from 50 repetitions, in the following order: SF surrogate, the set of ensemble combination-based MF surrogates added one by one based on the basis ranking, and the hierarchical-based MF surrogate.

Table 8
Performance comparison results for P5 and P6.

	$dm = 4$, $n_{HF} = 16$ $n_{LF} = (60, 60, 120, 120)$		$dm = 6$, $n_{HF} = 24$ $n_{LF} = (120, 120, 120)$	
	$\bar{\varepsilon}_{NRMSE}$	$\bar{\rho}(\mathbf{M}_{sel})$	$\bar{\varepsilon}_{NRMSE}$	$\bar{\rho}(\mathbf{M}_{sel})$
Method 1	0.1086	0.0	0.1215	0.0
Method 2–1	0.0917	1.0	0.0732	1.0
Method 2–2	0.1113	1.0	5.0407	1.0
Method 2–3	0.0809	1.0	0.0830	1.0
Method 3	0.0817	0.22	0.0909	0.23
Proposed	0.0808	**0.53 (0,18,14,12,6,0)***	0.0731	**0.83 (0,6,13,31,0)***

* The number in parentheses represents the number of surrogates selected from 50 repetitions, in the following order: SF surrogate, the set of ensemble combination-based MF surrogates added one by one based on the basis ranking, and the hierarchical-based MF surrogate.

Table 9
Performance comparison results for P7 and P8.

	$dm = 8$, $n_{HF} = 32$ $n_{LF} = (160, 160)$		$dm = 12$, $n_{HF} = 48$ $n_{LF} = (120, 240)$	
	$\bar{\varepsilon}_{NRMSE}$	$\bar{\rho}(\mathbf{M}_{sel})$	$\bar{\varepsilon}_{NRMSE}$	$\bar{\rho}(\mathbf{M}_{sel})$
Method 1	0.0153	0.0	0.1502	0.0
Method 2–1	0.0068	1.0	0.1229	1.0
Method 2–2	0.0260	1.0	0.1233	1.0
Method 2–3	0.0089	1.0	0.0960	1.0
Method 3	0.0105	0.46	0.0958	0.50
Proposed	0.0055	**0.82 (0,18,3,29)***	0.0904	**0.74 (0,26, 3,21)***

* The number in parentheses represents the number of surrogates selected from 50 repetitions, in the following order: SF surrogate, the set of ensemble combination-based MF surrogates added one by one based on the basis ranking, and the hierarchical-based MF surrogate.

correlations between y_{HF} and $y_{LF,1}$, as well as between y_{HF} and $y_{LF,2}$, are relatively weak, combining $y_{LF,1}$ and $y_{LF,2}$ results in a strong correlation with y_{HF}. Generally, this information is not known in advance. Under these circumstances, the numerical results indicate that the three Method 2 approaches outperforms Method 1 in nearly all cases. This finding aligns with the characteristics and trends associated with problems where utilizing all LF sources is advantageous. Method 3 exhibits performance similar to that of Method 1, as it fails to effectively distinguish the use of LF data sources and, even if it could, fundamentally does not consider the combination of LF data sources. The proposed method incorporates all LF sources, thereby yielding results identical to those of Method 2–1. Furthermore, the $\bar{\rho}$ value obtained from the proposed method is 1 in all cases, indicating that the final model selected utilizes all LF sources. Notably, the proposed algorithm identifies the ensemble combination-based MF model – corresponding to the third column of $\mathbf{M}_{sel}$ – as the optimal surrogate for this problem, rather than the hierarchical combination-based model in the fourth column. This result demonstrates the effective applicability of the proposed method.

Fourthly, the performance results for P4 are presented in Table 7. In these numerical examples, n_{HF} is fixed at 10, while the combinations of n_{LF} are varied to compare the performances of the six methods. A key characteristic of both problems is the strong correlation between y_{HF} and $y_{LF,1}$, while the correlation between y_{HF} and $y_{LF,2}$ is relatively weak. Additionally, there is a slight correlation between $y_{LF,1}$ and $y_{LF,2}$. This implies that the fidelity levels can be distinguished in the increasing order of $y_{LF,2}$, $y_{LF,1}$, and y_{HF}, though this prior knowledge is assumed to be unknown beforehand. As listed in Table 7, when n_{LF} is 40 for both LF sources, the three Method 2 approaches and the proposed method yield more accurate results than Method 1, indicating that LF sources provide useful information. Moreover, when n_{LF} is (40, 20), the accuracy decreases compared to the previous case, but the three Method 2 approaches and the proposed method still perform similarly and better than Method 1. However, when n_{LF} is (20, 40), Methods 2–1 and 2–2 perform worse than Method 1, while the proposed method continues to deliver the best results among the six approaches. This implies that $y_{LF,2}$ provides detrimental information to y_{HF}, and the superior performance of Methods 2–1 and 2–2 over Method 1 in the earlier cases is due to $y_{LF,1}$. Thus, it can be confirmed that $y_{LF,1}$ has a more positive influence on y_{HF} than $y_{LF,2}$. Specifically, when n_{LF} is (40, 20), n_{LF} for $y_{LF,1}$ is sufficient to mitigate the negative impact of $y_{LF,2}$. In addition, the performance of Method 3 is superior to Method 1 since it selects a single LF data source; however, overall, its superiority compared to the three Method 2 approaches appears unclear. This is because, as previously explained, Method 3 does not take into account decisions regarding the selective usage of LF data sources or the methods for combining them. However, the proposed method identifies y_{HF} as the highest-fidelity source, followed by $y_{LF,1}$ and $y_{LF,2}$, based on the inferred fidelity ranking. It then considers the use and combination of all sources to generate multiple candidate surrogates and selects the optimal surrogate that outperforms traditional approaches. In the proposed method, $y_{LF,1}$ is used in all cases as a mandatory source, as indicated by the first column of $\mathbf{M}_{sel}$ being 0 across all scenarios. Notably, when n_{LF} is (20, 40), the hierarchical-based MF model using $y_{LF,2}$ as the lowest LF source is predominantly selected as the final surrogate (37 times out of 50), which corresponds to the fourth column of $\mathbf{M}_{sel}$. This can also be easily inferred from the fact that $\bar{\rho}$ of the proposed method, like $\bar{\rho}$ of the three Method 2 approaches when using all LF data sources in an ensemble manner, approaches nearly 1, but the proposed method still demonstrates superior performance over the three Method 2 approaches. These results highlight the flexibility of the proposed method in handling varying LF sample sizes and effectively utilizing LF sources.

Lastly, additional numerical experiments were conducted on various test problems. The results for P5 and P6 are presented in Table 8, and the results for P7 and P8 are presented in Table 9. As indicated, the NRMSE values for the proposed method are consistently lower than those of existing approaches, highlighting the enhanced accuracy of the proposed surrogates. This enhanced performance is due to the adaptive utilization of LF sources and the effective combination strategies employed. Furthermore, these results imply that the proposed approach can at least serve as a useful guideline when it is unclear how to utilize all available LF data.

In summary, the proposed AQBMF surrogate framework is demonstrated to be a decision-making algorithm that adaptively and effectively selects LF sources and combination strategies to create a flexible surrogate. Notably, when there is no prior information about which LF models are helpful to the HF model or the accuracy levels among the LF models, the proposed method shows a stronger impact. Therefore, it can be concluded that the proposed AQBMF method is superior to existing methods based on the numerical results.

4.3. Real-world engineering example

This section validates the performance of the proposed method using a real-world engineering problem [46–48]. This engineering example reflects a common scenario encountered in real-world industrial settings, where the objective is to maximize the use of existing databases when designing and developing new products under a limited budget. The selected application is a vehicle model, which is typically optimized with respect to various performance criteria, including mass, bending stiffness, torsional stiffness, and natural frequency [47]. Among these, this study focuses on torsional stiffness, which is widely regarded as one of the most critical performance metrics in the optimization process [47], as it plays a key role in both vehicle handling and structural integrity. Fig. 7 presents three distinct finite element models of a vehicle, along with their corresponding boundary and loading conditions, and their specifications are summarized in Table 10. Given the limited computational budget of the HF model, the objective of this numerical test is to construct the most accurate surrogate model by maximizing the utilization of available LF data sources.

In this test, there are one HF model and two LF models. The HF model, as shown in Figs. 7(a) and (b), is a shell element-based finite element model of a car with model type A. As shown in Fig. 7(c), the LF model I is created by simplifying the HF model, which is composed of shells, into a beam-based model [49], significantly reducing the degrees of freedom. The beam cross-section is assumed to be rectangular, with the width reflecting the shell's geometry and the height approximated by the shell thickness. In other words, the model type remains the same, but the mesh types differ significantly, as listed in Table 10. LF model II, on the other hand, is a reference model that also uses shells but has a different model type from the HF model, as shown in Fig. 7(d). This reference model was previously developed for model type B, and has degrees of freedom comparable to those of the HF model, as listed in Table 10. LF model I has only 570 degrees of freedom and is extremely fast to analyze, with all 280 samples generated in less than a few seconds. LF model II consists of previously computed results stored in a corporate database, requiring no additional computation. Accordingly, both LF datasets are either cost-free or computationally negligible, and thus their cost was not considered in the analysis.

All finite element models are primarily made of steel, with a Young's modulus of 210 GPa, a Poisson ratio of 0.3, and a density of 7.89×10^{-6} kg/mm^3, respectively. The inputs for the surrogate models are the shell thicknesses of seven structural components, as illustrated in Fig. 7: the A-pillar, B-pillar, C-pillar, roof bar, floor bar, forward bar, and rear bar. The output is the torsional stiffness of the vehicle, defined as the applied torque divided by the resulting relative twist angle. It can be expressed as

$$y_{\mathrm{TS}} = \frac{T}{\theta_{\mathrm{TS}}} = \frac{(F \times L)}{\left(\tan^{-1}\left(\frac{\delta_{\mathrm{left}}(\boldsymbol{t}) + \delta_{\mathrm{right}}(\boldsymbol{t})}{L}\right)\right)} \tag{12}$$

where y_{TS} is the torsional stiffness, T denotes the torque applied to the front suspensions, θ_{TS} is the resulting twist angle, $\boldsymbol{t}$ represents the shell thicknesses of the seven components, F is the vertical force on the frontal suspension supports, L is the lateral distance between the suspension supports, $\delta_{\mathrm{left}}(\boldsymbol{t})$ and $\delta_{\mathrm{right}}(\boldsymbol{t})$ are the vertical displacements at the left and right loading points, respectively, obtained via finite element analysis. The lower and upper bounds of $\boldsymbol{t}$ are 0.5 mm and 5.0 mm. The applied torque T is 2000 N•m. The values of L for the HF model, LF model I, and LF model II are 1124 mm, 997 mm, and 1138 mm, respectively. Naturally, the corresponding forces F are 1778.8 N, 2005.9 N, and 1770 N, respectively, based on previous studies [46,47]. Furthermore, the number of samples obtained from LF model I and LF model II is set to 280 and 70, respectively, while three cases are considered for the number of samples from the HF model with values of 7, 14, and 28, as represented in Table 10. Therefore, the datasets D_{HF}, $D_{\mathrm{LF,I}}$, and $D_{\mathrm{LF,II}}$ are constructed, with the subscripts representing their respective sources. In this situation, the proposed algorithm determines which information from the two LF models is beneficial for supporting the HF model. In this study, the HF model is analyzed using the commercial finite element analysis (FEA) software OptiStruct [50], while the two LF models are analyzed using our in-house MATLAB codes.

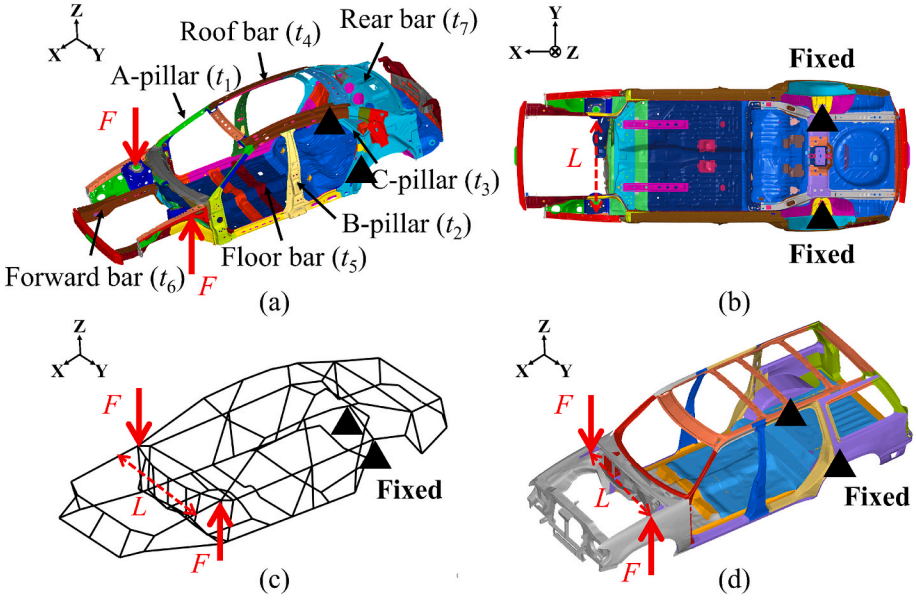


Fig. 7. Finite element models of the vehicle: (a) HF model (isometric view), (b) HF model (bottom view), (c) LF model I (isometric view), and (d) LF model II (isometric view).

Table 10
Information on finite element models.

	Model type	Mesh type	Analysis type	The number of samples	Degrees of freedom	Comments
HF model	A	Shell	Torsion	7/14/28	2,190,874	Original model
LF model I	A	Beam	Torsion	280	570	Simplified model
LF model II	B	Shell	Torsion	70	648,478	Reference model

Table 11
Performance comparison results for the engineering example.

	n_{HF}*	Importance ranking**	Final MF combination***	ϵ_{NRMSE}
Method 1	7	–	D_{HF}	0.2181
	14			0.08392
	28			**0.04249**
Method 2–1	7	–	$(D_{LF,I} + D_{LF,II}) \rightarrow D_{HF}$	**0.1360**
	14			0.07947
	28			0.04916
Method 2–2	7	–	$(D_{LF,I} + D_{LF,II}) \rightarrow D_{HF}$	1.7773
	14			0.06304
	28			0.05378
Method 2–3	7	–	$(D_{LF,I} + D_{LF,II}) \rightarrow D_{HF}$	0.22801
	14			0.06301
	28			0.05225
Method 3	–	–	–	–
Proposed	7	$(D_{LF,II}, D_{LF,I})$	$(D_{LF,I} + D_{LF,II}) \rightarrow D_{HF}$	**0.1360**
	14	$(D_{LF,II}, D_{LF,I})$	$D_{LF,I} \rightarrow D_{LF,II} \rightarrow D_{HF}$	**0.05608**
	28	$(D_{LF,II}, D_{LF,I})$	D_{HF}	**0.04249**

* Since all LF information is assumed to be given, all n_{HF} values are the same.
** LF data sources are listed in order of importance from the most important (left) to the least important (right).
*** Final combination process is represented, where the plus (+) and arrow (→) indicate the ensemble and hierarchical combination approaches, respectively.

Table 11 presents the results of surrogate models created using both existing and proposed methods for the engineering problem. As listed in Table 11, when n_{HF} is 7, the performance varies across the methods. When n_{HF} is 14, Methods 2–1, 2–2, and 2–3 show higher accuracy than Method 1, whereas when n_{HF} is 28, Method 1 outperforms Methods 2–1, 2–2, and 2–3. This indicates that LF information is helpful when n_{HF} is small, but the contribution of LF information diminishes as n_{HF} increases. In addition, Method 3 cannot be applied because there are no common sample points between the original LF data and the HF data, which means that Method 3 is impractical. In contrast, the proposed method produces results that are as accurate as or more accurate than the existing methods in all cases, as indicated by its lower NRMSE values. Specifically, the proposed method identifies $D_{LF,II}$ as more significant information than $D_{LF,I}$ for all cases, indicating that LF model II provides more useful information for the HF model than LF model I. This inference is also supported by the global sensitivity analysis shown in Fig. 8, where the overall trend of the Sobol' indices [9] for the HF model is more closely aligned with that of LF model II than with LF model I. In other words, the type of elements has a greater impact than the type of models in this problem. Based on the estimated quality levels of the LF data sources, the proposed method adaptively selects the best surrogate, whether it is the ensemble-based MF, hierarchical-based MF, or SF surrogate, as the number of HF samples increases to 7, 14, and 28, achieving consistently accurate results. In the future, if a more experienced expert refines LF model I, its relative importance may increase, allowing for more varied results through comparison with LF model II. In conclusion, the proposed AQBMF method demonstrates its capability as a flexible data-driven decision-making algorithm that maximizes the

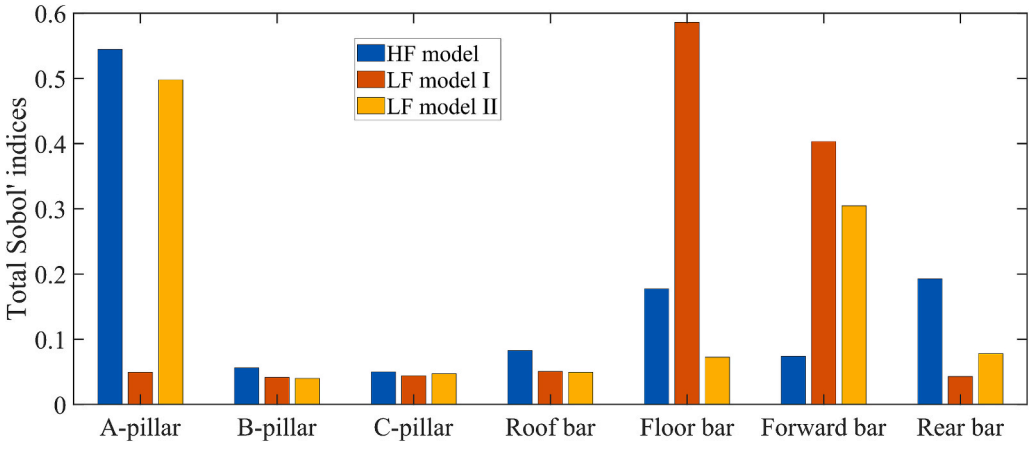


Fig. 8. Global sensitivity analysis results for the three engineering models.

utilization of multiple LF data sources to enhance surrogate model performance compared to existing methods.

5. Conclusions

This study presents a new paradigm, the AQBMF surrogate framework, designed to effectively utilize multiple LF data sources with unknown fidelity levels. The proposed method is particularly recommended for situations where multiple LF data sources with undefined fidelity levels are available, but there is uncertainty about which sources to select and how to combine them. The core idea of the proposed method is to interpret the quality level of LF data sources as the relative significance of basis functions in the surrogate model. Based on this idea, the proposed approach systematically builds the best surrogate in four stages, adaptively selecting LF data sources and MF combination methods. The first stage is to construct the augmented MF formulation, in which all basis functions are initially assigned equal importance. The second stage is to assess the importance of LF surrogates using the proposed MF basis screening strategy. The third stage is to build promising candidate surrogates based on the estimated ranking of LF surrogates. The last stage is to select the best surrogate among the candidate surrogates. The numerical results confirm that the proposed method not only adaptively utilizes the LF data sources but also appropriately employs the MF combination method. The first numerical results demonstrate the flexible application process of the proposed approach step-by-step through a 1D example. The second numerical results verify that the performance of the proposed method outperforms that of conventional methods through various benchmark test functions. The final numerical results show that the proposed data-driven decision algorithm generates a more accurate surrogate than other approaches in the same computational efficiency setting using a real-world engineering example.

In the future, each sub-stage of the proposed framework could be replaced or integrated with more advanced techniques, further enhancing overall performance. For instance, a new metric could be developed to more precisely assess the utility of LF sources even with a limited number of HF data points. Moreover, the proposed AQBMF method is expected to be extended to achieve more robust performance in problems where strong nonlinear correlations exist between LF and HF models [51,52]. Finally, the proposed method will be incorporated into sequential sampling processes (e.g., Bayesian optimization) to further broaden its applicability [53].

CRediT authorship contribution statement

Mingyu Lee: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Juyoung Lee:** Writing – original draft, Methodology, Investigation. **Jae-Hoon Choi:** Software, Data curation. **Nam H. Kim:** Writing – original draft, Methodology. **Ikjin Lee:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the InnoCORE program of the Ministry of Science and ICT (N10250154).

Appendix A. Implementation details

In this section, the two implementation details necessary for generating fair and reliable numerical results are provided. The methods employed in constructing the surrogates and the procedures for calculating the PCC metrics are explained sequentially.

Firstly, scaling techniques are used to mitigate potential performance degradation in the surrogate model caused by variations in the ranges of design variables. In this study, standardization techniques [4] are applied to both inputs and outputs. The hyperparameter range for the Gaussian spatial correlation function is set to $[10^{-6}, 10^{2}]$ [4,12], and the pattern search algorithm is employed with various initial points to estimate the hyperparameters [4,12]. In this study, all surrogates are constructed using our in-house MATLAB codes, referring to Ref. [54].

Secondly, raw data-based and prediction-based PCC metrics [16] are computed as

$$r = f_{\text{corr}}\left(\boldsymbol{Y}_{\text{S,LF}}^{\text{common}},\ \boldsymbol{Y}_{\text{S,HF}}^{\text{common}}\right) \tag{13}$$

and

$$\widetilde{r} = f_{\text{corr}}\left(\widetilde{\boldsymbol{Y}}_{\text{LF}},\ \widetilde{\boldsymbol{Y}}_{\text{MF}}\right) \tag{14}$$

respectively, where Pearson correlation function f_{corr}, represented as

$$f_{\text{corr}}(\boldsymbol{Y}_{\text{A}},\ \boldsymbol{Y}_{\text{B}}) = \left(\frac{\sum_{k=1}^{n}\left(\boldsymbol{Y}_{\text{A}_k} - \overline{\boldsymbol{Y}}_{\text{A}}\right)\left(\boldsymbol{Y}_{\text{B}_k} - \overline{\boldsymbol{Y}}_{\text{B}}\right)}{\sqrt{\sum_{k=1}^{n}\left(\boldsymbol{Y}_{\text{A}_k} - \overline{\boldsymbol{Y}}_{\text{A}}\right)^2}\sqrt{\sum_{k=1}^{n}\left(\boldsymbol{Y}_{\text{B}_k} - \overline{\boldsymbol{Y}}_{\text{B}}\right)^2}} \right) \tag{15}$$

where $\boldsymbol{Y}_{\text{A}}$ and $\boldsymbol{Y}_{\text{B}}$ are output sets of n sampled points from each data group, with the upper bar indicating the mean of the respective sets. In addition, $\boldsymbol{Y}_{\text{S,LF}}^{\text{common}}$ and $\boldsymbol{Y}_{\text{S,HF}}^{\text{common}}$ are the outputs for sampled common LF and HF inputs, respectively, and $\widetilde{\boldsymbol{Y}}_{\text{LF}}(=\widehat{y}_{\text{LF}}(\boldsymbol{S}_{\text{D}}))$ and $\widetilde{\boldsymbol{Y}}_{\text{MF}}(=\widehat{y}_{\text{MF}}(\boldsymbol{S}_{\text{D}}))$ are the predicted values for LF and MF surrogates over the entire design area, respectively, where $\boldsymbol{S}_{\text{D}}$ represents the sample points over the entire design domain. In this study, $\boldsymbol{S}_{\text{D}}$ is generated using LHS with a sample size that is 100 times the number of design variables. In particular, when obtaining PCC for the candidate ensemble combination-based MF surrogates in Section 3.2.3, $\boldsymbol{f}_{\text{aug}}(\boldsymbol{x})^{\text{T}}\widehat{\boldsymbol{\beta}}_{\text{aug}}$ and $\widehat{y}(\boldsymbol{x})$ in Eq. (7) are used for calculating $\widetilde{\boldsymbol{Y}}_{\text{LF}}$ and $\widetilde{\boldsymbol{Y}}_{\text{MF}}$, respectively. This is because the sum of the weighted LF information (i.e., $\boldsymbol{f}_{\text{aug}}(\boldsymbol{x})^{\text{T}}\widehat{\boldsymbol{\beta}}_{\text{aug}}$) reflects the trend of the MF surrogate.

Thirdly, the NCVE value (i.e., ϵ_{NCVE}) [55], especially for leave-one-out error, is calculated as

$$\epsilon_{\mathrm{NCVE}} = \frac{\sqrt{\frac{1}{n_{\mathrm{HF}}}\sum_{j=1}^{n_{\mathrm{HF}}}\left(y_j - \widehat{y}_{-j}\left(\boldsymbol{S}_{\mathrm{HF},j}\right)\right)^2}}{\max(\boldsymbol{Y}_{\mathrm{HF}}) - \min(\boldsymbol{Y}_{\mathrm{HF}})} \tag{16}$$

where $\boldsymbol{S}_{\mathrm{HF}}$ and $\boldsymbol{Y}_{\mathrm{HF}}$ are the inputs and outputs of HF data, respectively; n_{HF} is the number of HF data; y_j is the output of $\boldsymbol{S}_{\mathrm{HF},j}$; $\widehat{y}_{-j}\left(\boldsymbol{S}_{\mathrm{HF},j}\right)$ is the predicted HF output of $\boldsymbol{S}_{\mathrm{HF},j}$ from the surrogate built without $\boldsymbol{S}_{\mathrm{HF},j}$ and y_j; $\max(\boldsymbol{Y}_{\mathrm{HF}})$ and $\min(\boldsymbol{Y}_{\mathrm{HF}})$ are the maximum and minimum outputs of the observed HF data, respectively.

Appendix B. Computational complexity of the proposed method

This section discusses the computational complexity of the proposed method. To obtain a rough estimate in Big-O notation, assume that there are N_{S} LF sources, each with n_{LF} samples, and that the HF dataset has n_{HF} samples. Under these assumptions, the computational complexity of conventional EHK is $N_{\mathrm{S}} \times O\left(n_{\mathrm{LF}}^3\right) + O\left(n_{\mathrm{HF}}^3\right)$, because each LF surrogate is trained once and then calibrated with the HF data. In the proposed method, four stages are considered, as mentioned earlier in Section 3.2. In Stage 1, N_{S} LF surrogates are constructed, thereby requiring a computational complexity of $N_{\mathrm{S}} \times O\left(n_{\mathrm{LF}}^3\right)$. Stage 2 requires $N_{\mathrm{S}} \times O\left(n_{\mathrm{HF}}^3\right)$ for basis screening. The dominant cost comes from computing NCVE for the N basis functions, which can be efficiently evaluated via the inversion of partitioned matrix [36]. In Stage 3, $(2 + N_{\mathrm{S}})$ candidate surrogates are systematically constructed as listed in Table 12, yielding a computational complexity of $(2 + N_{\mathrm{S}}) \times O\left(n_{\mathrm{HF}}^3\right)$. In addition, the computational cost of Stage 4 is negligible since all key metrics required to select the best surrogate have already been computed in Stage 3. Therefore, the total computational complexity of the proposed method is $N_{\mathrm{S}} \times O\left(n_{\mathrm{LF}}^3\right) + (2 + 2N_{\mathrm{S}}) \times O\left(n_{\mathrm{HF}}^3\right)$. Here, given that n_{HF} is generally small, the influence of the additional $(2N_{\mathrm{S}}) \times O\left(n_{\mathrm{HF}}^3\right)$ in the proposed method decreases as n_{LF} increases relative to n_{HF}. However, when the importance ranking of the basis functions is unknown and all possible cases are considered, the advantage of the proposed method becomes more evident. Specifically, as indicated in Table 12, if all combinations are considered, the computational complexity becomes $N_{\mathrm{S}} \times O\left(n_{\mathrm{LF}}^3\right) + \left(N_{\mathrm{S}}! + 2^{N_{\mathrm{S}}}\right) \times O\left(n_{\mathrm{HF}}^3\right)$, so $\left(N_{\mathrm{S}}! + 2^{N_{\mathrm{S}}}\right) \times O\left(n_{\mathrm{HF}}^3\right)$ term can no longer be ignored. The detailed counts for each combination are presented in Table 12, and the comparison is visualized in Fig. 9. These results show that while the number of candidate surrogate combinations in all cases increases factorially with respect to N_{S}, the proposed method increases linearly, making it much more efficient. In other words, when the importance of the LF basis functions is not given, exploring all possible cases makes the computation almost prohibitive due to the exponential increase in the number of combinations.

Table 12
The number of candidate surrogate combinations when LF source # $= N_{\mathrm{S}}$ ($N_{\mathrm{S}} \geq 2$).

	Conventional	All	Proposed
SF surrogate	–	1	1
Hierarchical-based MF surrogate	–	$N_{\mathrm{S}}!$	1
Ensemble-based MF surrogate	1	$2^{N_{\mathrm{S}}} - 1$ $(={}_{N_{\mathrm{S}}}\mathrm{C}_1 + {}_{N_{\mathrm{S}}}\mathrm{C}_2 + \ldots + {}_{N_{\mathrm{S}}}\mathrm{C}_{N_{\mathrm{S}}})$	N_{S}
Total	1	$N_{\mathrm{S}}! + 2^{N_{\mathrm{S}}}$	$2 + N_{\mathrm{S}}$

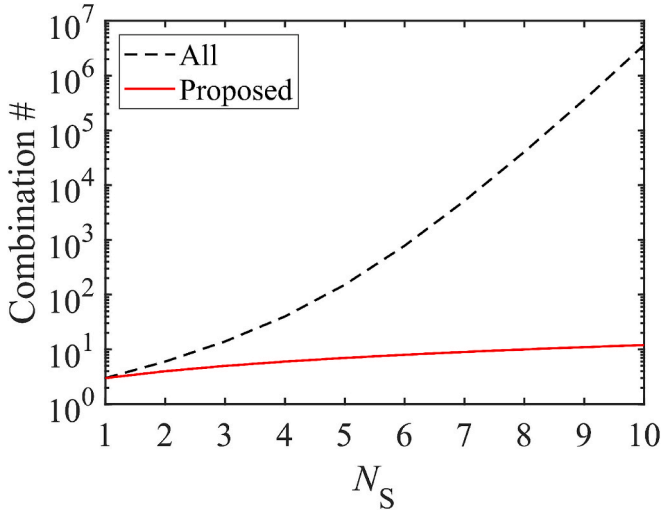


Fig. 9. Comparison of candidate surrogate combinations with respect to the number of LF sources (logarithmic scale on y-axis)

Appendix C. Conventional MF dataset selection algorithm

This section explains the conventional MF dataset selection algorithm [16,19], which can be summarized as follows:

(1) Select the LF data source with the highest $|r|$ value from the available LF data sources.
(2) If the chosen $|r|$ exceeds the predefined threshold $r_{\min}$, designate the corresponding MF surrogate as the best model. Otherwise, select the SF surrogate as the best model.

Appendix D. Benchmark test functions

The benchmark test functions used in this study are adopted or modified from Refs. [28] and [55] and their expressions are summarized in Table 13.

Table 13
Formulations of the analytical test functions.

Test function #	Expression
Test function 1	$y_{HF}(x) = \sin(x) + 0.2x + (x-5)^2/16 + 0.5$ $y_{LF,1}(x) = (x-0.5)(x-4)(x-9)/20 + 2$ $y_{LF,2}(x) = \sin(x) + 0.2x + 0.5$ $x \in [0,\ 10]^1$
Test function 2	$y_{HF}(x) = (6x-2)^2\sin(12x-4)$ $y_{LF,1}(x) = (x-0.5)(x-4)(x-9)/20 + 2$ $y_{LF,2}(x) = \left(\cos(18x)\left(2x^2-1\right)\right)/15$ $x \in [0,\ 1]^1$
Test function 3	$y_{HF}(x) = 0.4(6x-2)^2\sin(12x-4) + 10\left(0.4x^3 + 0.1x^2 + 0.2x + 0.2\right)$ $y_{LF,1}(x) = 0.4(6x-2)^2\sin(12x-4) + 4x^3 - 10e^x$ $y_{LF,2}(x) = 9\left(0.1x^2 + 0.2x + 0.2\right) + 8e^x$ $x \in [0,\ 1]^1$
Test function 4	$y_{HF}(\boldsymbol{x}) = 4x_1^2 - 2.1x_1^4 + x_1^6/3 + x_1x_2 - 4x_2^2 + 4x_2^4$ $y_{LF,1}(\boldsymbol{x}) = y_{HF}(0.7x_1,\ 0.8x_2) + x_1x_2 - 65$ $y_{LF,2}(\boldsymbol{x}) = y_{HF}(0.8x_1,\ 0.6x_2) - x_1^4 + 32$ $\boldsymbol{x} \in [-2,\ 2]^2$
Test function 5	$y_{HF}(\boldsymbol{x}) = \sum_{i=1}^{dm}\frac{x_i^2}{4000} - \prod_{i=1}^{dm}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$ $y_{LF,1}(\boldsymbol{x}) = \prod_{i=1}^{dm} x_i\sin\left(\frac{x_i}{\sqrt{i}}\right) + 1$ $y_{LF,2}(\boldsymbol{x}) = \sum_{i=1}^{dm}\frac{x_i^2}{7} - \prod_{i=1}^{dm}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$ $y_{LF,3}(\boldsymbol{x}) = \sum_{i=1}^{dm}\frac{x_i^2}{4000} - \prod_{i=1}^{dm}\cos\left(\frac{x_i}{i}\right) + 1$ $y_{LF,4}(\boldsymbol{x}) = \sum_{i=1}^{dm}\frac{x_i^2}{2000} - \prod_{i=1}^{dm}0.2\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$ $\boldsymbol{x} \in [-2,\ 2]^4$
Test function 6	$y_{HF}(\boldsymbol{x}) = [100\left(x_2 - x_1^2\right)^2 + (x_1-1)^2 + 100\left(x_3 - x_2^2\right)^2 + (x_2-1)^2 + 100\left(x_4 - x_3^2\right)^2 + (x_3-1)^2$ $+100\left(x_5 - x_4^2\right)^2 + (x_4-1)^2 + 100\left(x_6 - x_5^2\right)^2 + (x_5-1)^2]/100000$ $y_{LF,1}(\boldsymbol{x}) = \left[\left(x_2^4 + 50x_1^2\right) + \left(x_3^4 + 50x_2^2\right) + \left(x_4^4 + 50x_3^2\right) + \left(x_5^4 + 50x_4^2\right) + \left(x_6^4 + 50x_5^2\right)\right]/100000$ $y_{LF,2}(\boldsymbol{x}) = \sum_{i=1}^{dm} ix_i^2$ $y_{LF,3}(\boldsymbol{x}) = \sum_{i=1}^{dm}\frac{x_i^2}{4000} - \prod_{i=1}^{dm}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$ $\boldsymbol{x} \in [-5,\ 10]^6$
Test function 7	$y_{HF}(\boldsymbol{x}) = y_{sub}(\boldsymbol{x}) + 0.01x_1^2x_8 + x_1x_7/x_3 + x_1x_6/x_2 + x_1^2x_4$ $y_{LF,1}(\boldsymbol{x}) = 0.01x_1^2x_8 + x_1x_7/x_3 + x_1x_6/x_2 + x_1^2x_4$ $y_{LF,2}(\boldsymbol{x}) = 0.0001y_{sub}(\boldsymbol{x}) + x_1x_5/4x_2 + x_1x_8/x_4 + x_1^3x_7$ where $y_{sub}(\boldsymbol{x}) = 2\pi x_3(x_4 - x_6)/\left(\log(x_2/x_1)(1 + 2x_7x_4)/\log(x_2/x_1)x_1^2x_8\right) + x_3/x_5$ $x_1 \in [0.05,\ 0.15];\ x_2 \in [100,\ 50000];\ x_3 \in [63070,\ 115600];$ $x_4 \in [990,\ 1110];\ x_5 \in [63.1,\ 116];\ x_6 \in [700,\ 820];$ $x_7 \in [1120,\ 1680];\ x_8 \in [9855,\ 12045];$
Test function 8	$y_{HF}(\boldsymbol{x}) = \sum_{i=1}^{dm/4}\left[(x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4\right]$ $y_{LF,1}(\boldsymbol{x}) = \sum_{i=1}^{dm/4}\left[(x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 8(x_{4i-3} - x_{4i})^4\right]$ $y_{LF,2}(\boldsymbol{x}) = \sum_{i=1}^{dm/4}\left[(x_{4i-3} + 10x_{4i-2})^2 + 4(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 5(x_{4i-3} - x_{4i})^4\right]$ $\boldsymbol{x} \in [-5,\ 5]^{12}$

Data availability

Data will be made available on request.

References

[1] A.I. Forrester, A.J. Keane, Recent advances in surrogate-based optimization, Prog. Aerosp. Sci. 45 (1–3) (2009) 50–79.

[2] L. Chen, H. Qiu, L. Gao, C. Jiang, Z. Yang, Optimization of expensive black-box problems via Gradient-enhanced Kriging, Comput. Methods Appl. Mech. Eng. 362 (2020) 112861.

[3] H. Jo, K. Lee, M. Lee, Y. Jung, I. Lee, Optimization-based model calibration of marginal and joint output distributions utilizing analytical gradients, Struct. Multidiscip. Optim. 63 (2021) 2853–2868.

[4] M. Lee, Y. Noh, I. Lee, A novel sampling method for adaptive gradient-enhanced Kriging, Comput. Methods Appl. Mech. Eng. 418 (2024) 116456.

[5] H. Long, J. Hao, W. Ye, Z. Zhu, M. Shu, Integrating small data and shape prior knowledge with gradient-enhanced Kriging through adaptive knowledge sampling, Comput. Ind. Eng. 110660 (2024).

[6] L. Brunel, M. Balesdent, L. Brevault, R. Le Riche, B. Sudret, A survey on multi-fidelity surrogates for simulators with functional outputs: unified framework and benchmark, Comput. Methods Appl. Mech. Eng. 435 (2025) 117577.

[7] M. Kim, Y. Jung, M. Lee, I. Lee, An expected uncertainty reduction of reliability: adaptive sampling convergence criterion for Kriging-based reliability analysis, Struct. Multidiscip. Optim. 65 (7) (2022) 206.

[8] S. Yang, M. Lee, Y. Jung, H. Cho, W. Hu, I. Lee, An effective active learning strategy for reliability-based design optimization under multiple simulation models, Struct. Saf. 107 (2024) 102426.

[9] K. Cheng, Z. Lu, C. Ling, S. Zhou, Surrogate-assisted global sensitivity analysis: an overview, Struct. Multidiscip. Optim. 61 (2020) 1187–1213.

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

[11] Forrester, A. I., Sóbester, A., & Keane, A. J. (2007). Multi-fidelity optimization via surrogate modelling. *Proceedings of the royal society a: mathematical, physical and engineering sciences*, *463*(2088), 3251-3269.

[12] M. Lee, Y. Jung, J. Choi, I. Lee, A reanalysis-based multi-fidelity (RBMF) surrogate framework for efficient structural optimization, Comput. Struct. 273 (2022) 106895.

[13] J. Lee, M. Lee, B.J. Lee, I. Lee, A comprehensive multi-fidelity surrogate framework based on Gaussian process for datasets with heterogeneous responses, Knowl.-Based Syst. 295 (2024) 111827.

[14] H. Lee, M. Lee, J. Jung, I. Lee, S. Ryu, Enhancing Injection Molding Optimization for SFRPs Through Multi-Fidelity Data-Driven Approaches Incorporating Prior Information in Limited Data Environments, Adv. Theor. Simul. 7 (8) (2024) 2400130.

[15] M. Lee, Y. Jung, C. Hwang, M. Kim, M. Kim, U. Lee, I. Lee, An efficient multi-fidelity design optimization framework for a thermoelectric generator system, Energ. Conver. Manage. 315 (2024) 118788.

[16] M. Lee, M.G. Jeong, J. Lee, B.J. Lee, I. Lee, Efficient and robust thermal battery design optimization leveraging physically similar data, Appl. Therm. Eng. 269 (2025) 126009.

[17] Y. Guo, D. Yuan, B. Wei, Y. Hu, Structural damage identification method of concrete dam based on multi-fidelity surrogate model collaboratively corrected by monitoring and simulation information, Adv. Eng. Inf. 67 (2025) 103559.

[18] L. You, S. Xing, J. Yi, S. Yuan, J. Yang, H. Pu, J. Luo, A harmonic domain regressor with dynamic task weighting strategy for multi-fidelity surrogate modeling in engineering design, Adv. Eng. Inf. 64 (2025) 102999.

[19] M. Giselle Fernández-Godino, C. Park, N.H. Kim, R.T. Haftka, Issues in deciding whether to use multifidelity surrogates, AIAA J. 57 (5) (2019) 2039–2054.

[20] Z.H. Han, S. Görtz, Hierarchical kriging model for variable-fidelity surrogate modeling, AIAA J. 50 (9) (2012) 1885–1896.

[21] J. Hu, Q. Zhou, P. Jiang, X. Shao, T. Xie, An adaptive sampling method for variable-fidelity surrogate models using improved hierarchical kriging, Eng. Optim. 50 (1) (2018) 145–163.

[22] Q. Zhou, Y. Wu, Z. Guo, J. Hu, P. Jin, A generalized hierarchical co-Kriging model for multi-fidelity data fusion, Struct. Multidiscip. Optim. 62 (2020) 1885–1904.

[23] X. Lai, Y. Pang, F. Liu, W. Sun, X. Song, A multi-fidelity surrogate model based on design variable correlations, Adv. Eng. Inf. 59 (2024) 102248.

[24] Y. Wang, Y. Pang, T. Xue, S. Zhang, X. Song, Ensemble learning based hierarchical surrogate model for multi-fidelity information fusion, Adv. Eng. Inf. 60 (2024) 102535.

[25] Y. Zhang, N.H. Kim, C. Park, R.T. Haftka, Multifidelity surrogate based on single linear regression, AIAA J. 56 (12) (2018) 4944–4952.

[26] M. Xiao, G. Zhang, P. Breitkopf, P. Villon, W. Zhang, Extended Co-Kriging interpolation method based on multi-fidelity data, Appl. Math Comput. 323 (2018) 120–131.

[27] M. Cheng, P. Jiang, J. Hu, L. Shu, Q. Zhou, A multi-fidelity surrogate modeling method based on variance-weighted sum for the fusion of multiple non-hierarchical low-fidelity data, Struct. Multidiscip. Optim. 64 (2021) 3797–3818.

[28] L. Zhang, Y. Wu, P. Jiang, S.K. Choi, Q. Zhou, A multi-fidelity surrogate modeling approach for incorporating multiple non-hierarchical low-fidelity data, Adv. Eng. Inf. 51 (2022) 101430.

[29] V. Pham, M. Tyan, T.A. Nguyen, J.W. Lee, Extended Hierarchical Kriging Method for Aerodynamic Model Generation Incorporating Multiple Low-Fidelity Datasets, Aerospace 11 (1) (2023) 6.

[30] H. Chen, L. Ouyang, L. Liu, Y. Ma, A multi-fidelity surrogate modeling method in the presence of non-hierarchical low-fidelity data, Aerosp. Sci. Technol. 146 (2024) 108928.

[31] M. Xiong, H. Huang, S. Xie, Y. Duan, A novel multi-fidelity surrogate modeling framework integrated with sequential sampling criterion for non-hierarchical data, Struct. Multidiscip. Optim. 67 (2) (2024) 13.

[32] Z. Guo, L. Song, C. Park, J. Li, R.T. Haftka, Analysis of dataset selection for multi-fidelity surrogates for a turbine problem, Struct. Multidiscip. Optim. 57 (2018) 2127–2142.

[33] D.J. Toal, Some considerations regarding the use of multi-fidelity Kriging in the construction of surrogate models, Struct. Multidiscip. Optim. 51 (2015) 1223–1245.

[34] Z.Z. Foumani, M. Shishehbor, A. Yousefpour, R. Bostanabad, Multi-fidelity cost-aware Bayesian optimization, Comput. Methods Appl. Mech. Eng. 407 (2023) 115937.

[35] J. Sacks, W.J. Welch, T.J. Mitchell, H.P. Wynn, Design and analysis of computer experiments, Stat. Sci. 4 (1989) 409–423.

[36] C.E. Rasmussen, C.K.I. Williams, *Gaussian processes for machine learning*, MIT Press, 2006.

[37] M.A. Bouhlel, N. Bartoli, A. Otsmane, J. Morlier, Improving kriging surrogates of high-dimensional design models by Partial Least Squares dimension reduction, Struct. Multidiscip. Optim. 53 (2016) 935–952.

[38] Joseph, V. R., Hung, Y., & Sudjianto, A. (2008). Blind Kriging: A new method for developing metamodels. *Journal of Mechanical Design*, *130*(3), 031102-1.

[39] L. Zhao, K.K. Choi, I. Lee, Metamodeling method using dynamic kriging for design optimization, AIAA J. 49 (9) (2011) 2034–2046.

[40] K. Kang, C. Qin, B. Lee, I. Lee, Modified screening-based Kriging method with cross validation and application to engineering design, App. Math. Model. 70 (2019) 626–642.

[41] P. Ramu, P. Thananjayan, E. Acar, G. Bayrak, J.W. Park, I. Lee, A survey of machine learning techniques in structural and multidisciplinary optimization, Struct. Multidiscip. Optim. 65 (9) (2022) 266.

[42] J. Lee, D. Park, M. Lee, H. Lee, K. Park, I. Lee, S. Ryu, Machine learning-based inverse design methods considering data characteristics and design space size in materials design and manufacturing: a review, Mater. Horiz. (2023).

[43] S. Oh, Y. Jung, S. Kim, I. Lee, N. Kang, Deep generative design: integration of topology optimization and generative models, J. Mech. Des. 141 (11) (2019) 111405.

[44] M. Lee, Y. Park, H. Jo, K. Kim, S. Lee, I. Lee, Deep generative tread pattern design framework for efficient conceptual design, J. Mech. Des. 144 (7) (2022) 071703.

[45] H.K. Yong, L. Wang, D.J. Toal, A.J. Keane, F. Stanley, Multi-fidelity Kriging-assisted structural optimization of whole engine models employing medial meshes, Struct. Multidiscip. Optim. 60 (2019) 1209–1226.

[46] D. Wang, K. Cai, Optimizing the static–dynamic performance of the body-in-white using a modified non-dominated sorting genetic algorithm coupled with grey relational analysis, Eng. Optim. 50 (4) (2018) 615–633.

[47] F. Xiong, D. Wang, S. Chen, Q. Gao, S. Tian, Multi-objective lightweight and crashworthiness optimization for the side structure of an automobile body, Struct. Multidiscip. Optim. 58 (2018) 1823–1843.

[48] H. Kim, T.H. Lee, T. Kwon, Normalized neighborhood component feature selection and feasible-improved weight allocation for input variable selection, Knowl.-Based Syst. 218 (2021) 106855.

[49] D. Mundo, R. Hadjit, S. Donders, M. Brughmans, P. Mas, W. Desmet, Simplified modelling of joints and beam-like structures for BIW optimization in a concept phase of the vehicle design process, Finite Elem. Anal. Des. 45 (6–7) (2009) 456–462.

[50] W.H. Choi, J.M. Kim, G.J. Park, Comparison study of some commercial structural optimization software systems, Struct. Multidiscip. Optim. 54 (3) (2016) 685–699.

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

[52] K. Li, Q. Li, L. Lv, X. Song, Y. Ma, I. Lee, A nonlinearity integrated bi-fidelity surrogate model based on nonlinear mapping, Struct. Multidiscip. Optim. 66 (9) (2023) 196.

[53] D. Zhan, H. Xing, Expected improvement for expensive optimization: a review, J. Glob. Optim. 78 (3) (2020) 507–544.

[54] Lophaven, S. N., Nielsen, H. B., & Søndergaard, J. (2002). DACE-A Matlab Kriging toolbox, version 2.0.

[55] K. Kang, I. Lee, Efficient high-dimensional metamodeling strategy using recursive decomposition coupled with sequential sampling method, Struct. Multidiscip. Optim. 63 (2021) 375–390.