Proceedings of the Eighteenth International Conference on
Civil, Structural and Environmental Engineering Computing
Edited by: P. Iványi, J. Kruis and B.H.V. Topping
Civil-Comp Conferences, Volume 10, Paper 4.1
Civil-Comp Press, Edinburgh, United Kingdom, 2025
ISSN: 2753-3239, doi: 10.4203/ccc.10.4.1

A Novel Kriging-Based Multi-Fidelity Surrogate Model and Optimization Strategy

S. Cho, J. Kim and G. Noh

Department of Mechanical Engineering, Korea University, Seoul, Republic of Korea

Abstract

Multi-fidelity surrogate modelling methods have gained significant attention in engineering optimization, as they can achieve high accuracy at a reduced computational cost. However, to construct a high-performing multi-fidelity surrogate model, it is essential to accurately capture the correlation between the high-fidelity and low-fidelity models. In this paper, we propose a novel Kriging-based multi-fidelity surrogate model. The proposed method tunes the low-fidelity Kriging model to capture both the linear and nonlinear correlations with the high-fidelity data, enabling a close approximation to the high-fidelity model. Then, a suitable surrogate model for the discrepancy data is selected from among the Kriging model and the polynomial regression model. The basis functions for the Kriging model and the discrepancy Kriging model are included as part of the hyperparameter optimization. All hyperparameters are optimized simultaneously using a metaheuristic algorithm to ensure that all complex relationships between hyperparameters are considered. The proposed method demonstrates superior performance and robustness in analytical problems.

Keywords: multi-fidelity, surrogate model, optimization, kriging, metaheuristic, leave one out cross validation, correlation coefficient.

1 Introduction

With the advancement of simulation technologies, high-fidelity (HF) simulations are widely used to solve complex engineering problems. However, due to their high computational cost, they are often unsuitable for tasks such as optimization. To address this, surrogate models have been widely adopted as computationally efficient alternatives to HF simulations. Ensuring the accuracy of a surrogate model typically requires a sufficient amount of HF data. However, in many cases, limited resources make it difficult to obtain enough HF data, which in turn hampers the accuracy of the surrogate model. To overcome this issue, multi-fidelity (MF) modelling approaches have been proposed and extensively studied in recent years [1].

Among various MF modelling approaches, the comprehensive framework is generally known to offer the best performance. In this framework, low-fidelity (LF) model is first constructed, and then a model is trained on the discrepancy between the HF data and the LF model. The HF model is composed of the LF model added with the discrepancy model. A widely used formulation of the comprehensive framework is as follows:

$$\hat{y}_{HF}(x) = \rho(x)\hat{y}_{LF}(x) + \hat{\delta}(x) \tag{1}$$

where $\hat{y}_{HF}(x)$ and $\hat{y}_{LF}(x)$ represent the HF model and LF model, respectively. $\rho(x)$ represents the scaling factor, and $\hat{\delta}(x)$ is the discrepancy model. This framework relies on both the LF model and the discrepancy model being accurately constructed to ensure the overall performance of the HF model. However, if the correlation between the LF and HF models is not properly captured, the discrepancy model may be poorly constructed, making it difficult to ensure the accuracy of the HF model. Therefore, accurately capturing the correlation between the LF and HF models is essential in multi-fidelity surrogate (MFS) modelling, as it enables the proper construction of both the LF model and the discrepancy model.

In comprehensive framework, it is necessary to optimize the hyperparameters scaling factor and discrepancy model. The performance of the MFS model is highly dependent on how these values are optimized. Mainly, Bayesian-based MFS models [2, 3, 4, 5] use maximum likelihood estimation (MLE) for optimization metric. In this case, the scaling factor ρ and hyperparameter θ_δ of discrepancy model are determined through the following equation [4]:

$$\left[\hat{\rho}, \boldsymbol{\theta}_\delta\right] = \underset{\rho,\theta_\delta}{\arg\max} - \frac{N}{2}\ln\left(\hat{\sigma}_\delta^2\right) - \frac{1}{2}\ln\left|\mathbf{R}_\delta\right| \tag{2}$$

where N is the number of the HF sample data, $\mathbf{R}_\delta$ is the correlation matrix and $\hat{\sigma}_\delta^2$ is the process variance.

Another widely used estimation for determining the scaling factor in MFS models is the mean squared error (MSE) [3]. The common approach for optimizing scaling factor ρ is to minimize the error between the HF data and the scaled LF model, which can be expressed as follows:

$$\hat{\rho} = \underset{\rho}{\arg\min} \sum_{i=1}^{N}\left[\rho\hat{y}_{LF}\left(x^i\right) - y_{HF}\left(x^i\right)\right]^2 \tag{3}$$

where $y_{HF}(\cdot)$ represents the HF output, and $\hat{y}_{LF}(\cdot)$ represent the prediction of the LF model. N is the number of the HF data, and x^i represents the i-th input. However, simply optimizing the MLE or MSE does not sufficiently account for the bumpiness of the discrepancy model, which can lead to poor accuracy [6].

2 Methods

In this paper, we propose a new formulation based on the comprehensive framework by introducing additional hyperparameters to more effectively tune the LF model to closely match the HF model. For the LF model, the Kriging is used to train the LF data. The Kriging trained on LF data is optimized to maximize the MLE with LF data to derive the hyperparameter $\boldsymbol{\theta}_{LF}$. The $\boldsymbol{\theta}_{LF}$ in LF Kriging is tailored to the information in the LF data and may be different from $\boldsymbol{\theta}_{HF}$, which is optimized by training the Kriging on efficient number of HF data. In general, the LF data may differ from the HF data but follow a similar trend. Thus, the trend of $\boldsymbol{\theta}_{LF}$ in the Kriging model trained on LF data is expected to be similar to the trend of $\boldsymbol{\theta}_{HF}$ in the Kriging model trained on HF data. To explore the relationship between $\boldsymbol{\theta}_{HF}$ and $\boldsymbol{\theta}_{LF}$, we utilize one of the benchmark test functions, the six-dimensional Rosenbrock function. The HF function y_h and the LF function y_l are defined in Table 1. The $\boldsymbol{\theta}_{HF}$ of the Kriging trained on HF data and the $\boldsymbol{\theta}_{LF}$ of the Kriging trained on LF data in the Rosenbrock function are plotted as shown in Figure 1.

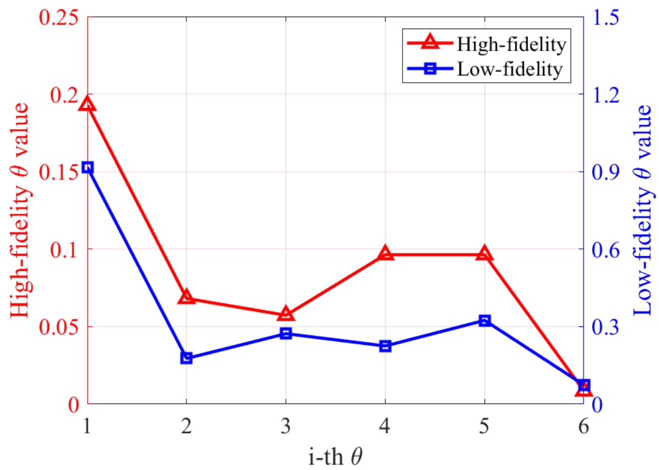

Figure 1: Kriging hyperparameter $\boldsymbol{\theta}$ values by fidelity of the Rosenbrock function.

Note that the scale of the i-th $\boldsymbol{\theta}$ values for $\boldsymbol{\theta}_{HF}$ and $\boldsymbol{\theta}_{LF}$ is different, but the trend follows a similar path. Therefore, by scaling $\boldsymbol{\theta}_{LF}$, it becomes possible to adjust the smoothness of the LF Kriging model to better align with the HF data. Based on this, we propose a new formulation that allows additional tuning of the Kriging hyperparameter $\boldsymbol{\theta}$ as follows:

$$\hat{y}_{HF}(x) = \rho \hat{y}^{n}_{LF}\left(x, \lambda \boldsymbol{\theta}_{LF}\left(\boldsymbol{\theta}_0\right)\right) + \hat{\delta}^{m}(x) \tag{4}$$

where $\hat{y}_{HF}(\cdot)$ represents the HF model, $\hat{y}_{LF}(\cdot)$ is the LF Kriging model, and $\hat{\delta}(\cdot)$ denotes the discrepancy model. $\boldsymbol{\theta}_0$ is the initial LF Kriging hyperparameter, and $\boldsymbol{\theta}_{LF}(\boldsymbol{\theta}_0)$ represents the LF Kriging hyperparameter optimized using MLE by training the LF data, with $\boldsymbol{\theta}_0$ as the starting point. The λ is a hyperparameter that scale the $\boldsymbol{\theta}_{LF}$, and ρ is a hyperparameter that scales the LF Kriging model. The n represents the type of basis function used in the LF Kriging model, and m denotes the type of surrogate model used in the discrepancy model, both serving as hyperparameters. The type of surrogate model is selected from the discrepancy model library, which includes Kriging model with different basis functions and the polynomial regression models. Thus, there are five hyperparameters to be optimized in the proposed method. All hyperparameters are optimized using genetic algorithm (GA), one of the metaheuristic algorithms suitable for solving complex problems [5]. We optimized the hyperparameters using leave one out cross validation (LOOCV) $\frac{MAE}{r^2}$ as a new objective function. Here, MAE denotes the mean absolute error, and r^2 represents the square of the Pearson correlation coefficient. Optimizing the hyperparameters to minimize this objective function yields optimized hyperparameters that minimize MAE and maximize r^2 between the HF samples and the proposed method.

3 Results

In this paper, we validate the proposed method through four analytical problems. The sample points are generated using latin hypercube sampling (LHS), which ensures effective coverage of the design space. The number of samples is determined based on a cost, where the total computational budget is defined as 40+5D, with D representing the number of design variables [7]. The cost of one HF sample is set to 1, and that of an LF sample is set to 0.2 [7]. In this study, the cost allocation ratio between HF and LF data is set to 0.6:0.4, which falls within the recommended range for effective MFS modelling as suggested in [8]. To account for the stochastic nature of LHS, 20 different DOE sets are generated and used to evaluate the MFS models. For performance evaluation, a validation set composed of 1000 HF samples is employed. The predictive performance is assessed using the coefficient of determination (R^2) and the normalized root mean squared error (NRMSE). The formulation of R^2 and NRMSE are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y^i - \hat{y}^i\right)^2}{\sum_{i=1}^{N}\left(y^i - \bar{y}\right)^2} \tag{5}$$

$$NRMSE = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y^i - \hat{y}^i\right)^2}}{y_{max} - y_{min}} \tag{6}$$

where N is the total number of test samples, y^i and $\hat{y}^i$ denote the actual output and the predicted value of surrogate model, respectively, and $\bar{y}$ is mean of actual outputs. The $y_{\max}$ and $y_{\min}$ are the maximum and minimum values of the actual outputs, respectively.

We demonstrate the performance of the proposed method by comparing the state-of-the-art methods and major widely used methods in of MFS modelling with the proposed method through the analytical problems. We selected the following MFS modelling methods for comparison: co-kriging (CO-KRG) [9], output scaling-MFS (OS-MFS) [5], multi-fidelity neural network (MF-NN) [10], and linear regression MFS (LR-MFS) [11]. Additionally, a single-fidelity kriging (SF-KRG) using only HF data was created and compared with the same cost of sample data used in the MFS models.

We validate the performance of proposed method using four analytical problems. These problems range from low to high dimensions and include various levels of nonlinearity to evaluate the effectiveness of MFS models. As shown in Table 1, selected examples include test functions commonly used in various research papers [12, 13].

	HF/LF	Test functions	Domain
Bird (2D)	HF	$y_h = \sin(x_1) e^{(1-\cos(x_2))^2} + \cos(x_2) e^{(1-\sin(x_1))^2} + (x_1 - x_2)^2$	$[-2\pi, 2\pi]$
	LF	$y_l = 0.9 y_h - 2 x_1 x_2$	
Rastrigin (3D)	HF	$y_h = 10d + \sum_{i=1}^{d-1} \left[x_i^2 - 10\cos(2\pi x_i) \right]$	$[-1, 1]$
	LF	$y_l = 10d + \sum_{i=1}^{d-1} 0.8 \left[x_i^2 - 10\cos(1.7\pi x_i) \right]$	
Rosenbrock (6D)	HF	$y_h = \sum_{i=1}^{d-1} \left[100\left(x_{i+1} - x_i^2\right)^2 + (x_i - 1)^2 \right]$	$[-5, 10]$
	LF	$y_l = \sum_{i=1}^{d-1} \left[100\left(x_{i+1} - 0.1x_i^3\right)^2 + 3(x_i - 3)^3 \right]$	
Styblinski-Tang (10D)	HF	$y_h = \sum_{i=1}^{10} x_i^4 - 16x_i^2 + 5x_i$	$[-3, 3]$
	LF	$y_l = \sum_{i=1}^{10} x_i^3 - 16x_i^2 + 5x_i$	

Table 1: The expression of analytical problems.

The results of evaluating the performance metrics for both the proposed method and the comparison method are shown in Figure 2 as a box plot, and the averages of the performance metrics are shown in Table 2. From the results, we can see that the proposed method performs well on all analytical problems. CO-KRG, which is a popular MFS model, is the next best performer. MF-NN failed to learn the Rosenbrock function and Styblinski-Tang function, which is likely due to the difficulty of securing the performance of neural networks with limited data [14]. This confirms that the proposed method performs robustly on examples with various nonlinearities in multiple dimensions

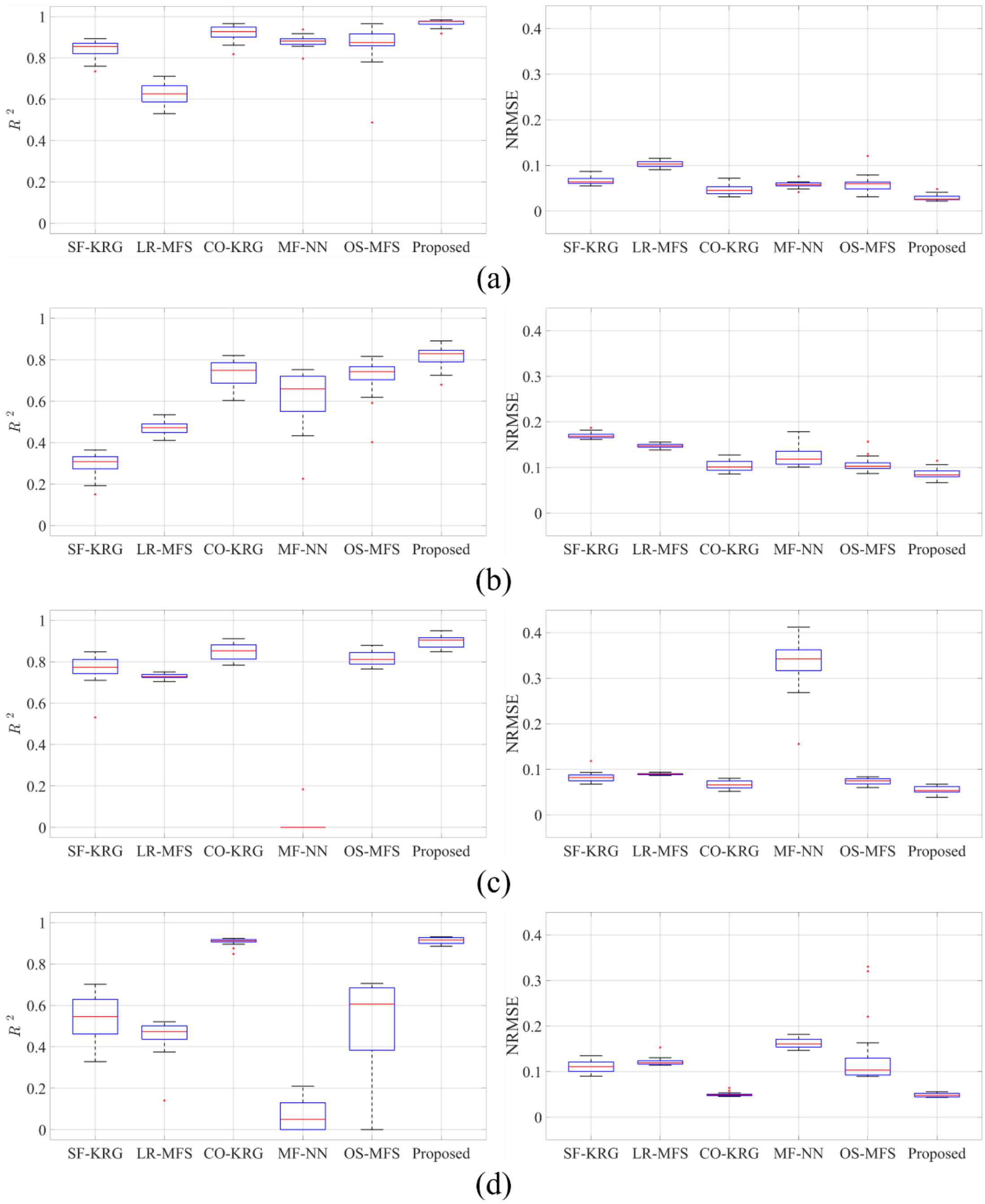

Figure 2: Box plot of performance results for the analytical problems; (a) Brid (2D), (b) Rastrigin (3D), (c) Rosenbrock (6D), (d) Styblinski-Tang (10D).

Test functions	Metrics	SF-KRG	LR-MFS	CO-KRG	MF-NN	OS-MFS	Proposed
Bird (2D)	R^2	0.8395	0.6251	0.9184	0.8789	0.8680	**0.9678**
	NRMSE	0.0670	0.1031	0.0471	0.0583	0.0585	**0.0295**
Rastrigin (3D)	R^2	0.2960	0.4713	0.7346	0.6263	0.7198	**0.8114**
	NRMSE	0.1699	0.1473	0.1037	0.1223	0.1061	**0.0872**
Rosenbrock (6D)	R^2	0.7655	0.7293	0.8479	0.0091	0.8169	**0.8966**
	NRMSE	0.0829	0.0898	0.0668	0.3403	0.0736	**0.0549**
Styblinski-Tang (10D)	R^2	0.5407	0.4498	0.9073	0.0695	0.4817	**0.9128**
	NRMSE	0.1110	0.1220	0.0500	0.1620	0.1342	**0.0485**

Table 2: Results of performance metrics in analytical problems.

4 Conclusions and Contributions

In this paper, we propose a multi-fidelity Kriging-based surrogate modelling approach with a novel formulation. The method involves five key hyperparameters: initial Kriging hyperparameter ($\boldsymbol{\theta}_0$), scaling factor (ρ, λ), type of basis function in LF Kriging model (n) and type of surrogate model in discrepancy model (m). All hyperparameters were optimized using GA to optimize for complex interactions. This ensures that the LF Kriging model aligns as closely as possible with the HF model and that the best model for the discrepancy data is selected. In doing so, we found that the proposed method robustly outperforms other MFS models on a variety of analytical problems.

In this paper, a new formulation is proposed to further optimize $\boldsymbol{\theta}$, the main hyperparameter that determines the performance of Kriging, in a comprehensive framework. In addition, LOOCV $\frac{MAE}{r^2}$ is used as an optimization objective function to optimize various hyperparameters more robustly. The proposed formulation and objective function can be easily modified and applied to other studies.

The proposed method has a high computational cost compared to other techniques because it optimizes multiple hyperparameters through GA. In future work, it is necessary to exploit the hyperparameter optimization strategy to reduce this computational cost.

Acknowledgements

This work was partly supported by the Young Scientist Grants-Outstanding Track (Grant No. NRF-2021R1C1C1011494) through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT.

References

[1] A.I. Forrester, A.J. Keane, Recent advances in surrogate-based optimization, Progress in aerospace sciences, 45 (2009) 50-79, https://doi.org/10.1016/j.paerosci.2008.11.001.

[2] A.I. Forrester, A. Sóbester, A.J. Keane, Multi-fidelity optimization via surrogate modelling, Proceedings of the royal society a: mathematical, physical and engineering sciences, 463 (2007) 3251-3269, https://doi.org/10.1098/rspa.2007.1900.

[3] S.-H. Son, D.-H. Choi, The effects of scale factor and correction on the multi-fidelity model, Journal of Mechanical Science and Technology, 30 (2016) 2075-2081, https://doi.org/10.1007/s12206-016-0414-0.

[4] H. Bu, L. Song, Z. Guo, J. Li, Selecting scale factor of Bayesian multi-fidelity surrogate by minimizing posterior variance, Chinese Journal of Aeronautics, 35 (2022) 59-73, https://doi.org/10.1016/j.cja.2022.05.012.

[5] Y. Zhao, J. Liu, Z. He, A general multi-fidelity metamodeling framework for models with various output correlation, Structural and Multidisciplinary Optimization, 66 (2023), https://doi.org/10.1007/s00158-023-03537-5.

[6] L. Shu, P. Jiang, X. Song, Q. Zhou, Novel approach for selecting low-fidelity scale factor in multifidelity metamodeling, AIAA Journal, 57 (2019) 5320-5330, https://doi.org/10.2514/1.J057989.

[7] R. Pellegrini, J. Wackers, R. Broglia, A. Serani, M. Visonneau, M. Diez, A multi-fidelity active learning method for global design optimization problems with noisy evaluations, Engineering with Computers, 39 (2023) 3183-3206, https://doi.org/10.1007/s00366-022-01728-0.

[8] D.J.J. Toal, Some considerations regarding the use of multi-fidelity Kriging in the construction of surrogate models, Structural and Multidisciplinary Optimization, 51 (2014) 1223-1245, https://doi.org/10.1007/s00158-014-1209-5.

[9] A.I. Forrester, A. Sóbester, A.J. Keane, Multi-fidelity optimization via surrogate modelling, Proceedings of the royal society a: mathematical, physical and engineering sciences, 463 (2007) 3251-3269, https://doi.org/10.1098/rspa.2007.1900.

[10] X. Meng, G.E. Karniadakis, A composite neural network that learns from mult i-fidelity data: Application to function approximation and inverse PDE proble ms, Journal of Computational Physics, 401 (2020), https://doi.org/10.1016/j.jcp.2019.10902.

[11] Y. Zhang, N.H. Kim, C. Park, R.T. Haftka, Multifidelity Surrogate Based on Single Linear Regression, AIAA Journal, 56 (2018) 4944-4952, https://doi.org/10.2514/1.J057299.

[12] M. Jamil, X.-S. Yang, A literature survey of benchmark functions for global optimisation problems, International Journal of Mathematical Modelling and Numerical Optimisation, 4 (2013) 150-194, https://doi.org/10.1504/IJMMNO.2013.055204.

[13] G.H. Cheng, A. Younis, K. Haji Hajikolaei, G. Gary Wang, Trust region based mode pursuing sampling method for global optimization of high dimensional

design problems, Journal of Mechanical Design, 137 (2015) 021407, https://doi.org/10.1115/1.4029219.
[14] X. Wang, Y. Xiao, W. Li, M. Wang, Y. Zhou, Y. Chen, Z. Li, Kriging-based surrogate data-enriching artificial neural network prediction of strength and permeability of permeable cement-stabilized base, Nature Communications, 15 (2024) 4891, https://doi.org/10.1038/s41467-024-48766-4.