RESEARCH PAPER

Multi-fidelity uncertainty propagation using polynomial chaos and Gaussian process modeling

Fenggang Wang[1] · Fenfen Xiong[1] (ID) · Shishi Chen[2] · Jianmei Song[1]

Abstract

The polynomial chaos (PC) method has been widely studied and applied for uncertainty propagation (UP) due to its high efficiency and mathematical rigor. However, the straightforward application of PC on the computationally expensive and highly complicated model for UP might be too costly. Therefore, a multi-fidelity PC approach using the Gaussian process modeling theory is developed in this work, by which the classic multi-level co-kriging multi-fidelity modeling framework in the deterministic domain is extended to the stochastic one. Meanwhile, taking advantage of the Gaussian process modeling theory, the strategies for response models with hierarchical and non-hierarchical fidelity are both addressed within the proposed multi-fidelity PC approach. The effectiveness and relative merit of the proposed method are demonstrated by comparative studies on several numerical examples for UP. It is noticed that the proposed approach can significantly improve the accuracy and robustness of UP compared to the commonly used addition correction-based multi-fidelity PC method; compared to co-kriging, the accuracy and robustness are generally also improved, especially for problems with unsymmetric distributed random input and large variation. An engineering robust aerodynamic optimization problem further verifies the applicability of the proposed multi-fidelity PC method.

Keywords Uncertainty propagation · Multi-fidelity · Polynomial chaos · Gaussian process modeling

Nomenclature

GP	Gaussian process
HF	The high-fidelity model
LF	The low-fidelity model
PC	Polynomial chaos
b_i	The ith coefficient of PC model
$\mathbf{d}$	Response data
d	Dimension of random inputs
h	Exponential of the exponential correlation function
n_i	The number of sample points for the ith-level fidelity model
s	The highest level fidelity
t	The ith-level fidelity

$\mathbf{x}$	Random input vector
y	Stochastic response value
$\mathbf{B}$	Polynomial coefficient matrix
$\mathbf{D}_t$	Input sites
$\mathbf{E}$	Unknown covariance matrix between lower-fidelity models
R	Correlation function
$\boldsymbol{\alpha}$	Multi-indices for PC
$\delta(x)$	Correction function
μ	Mean value
$\boldsymbol{\theta}$	Hyper-parameters vector
ρ	Scaling factor
σ	Standard deviation value
ξ	Random vector in standard random space
Φ	Orthogonal polynomial
Γ	Input space

✉ Fenfen Xiong
  fenfenx@bit.edu.cn

[1] School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China

[2] Beijing Electro-Mechanical Engineering Institute, 100074 Beijing, People's Republic of China

1 Introduction

Uncertainty propagation (UP) methods, which quantify uncertainty in system output performance based on random or noisy inputs, are of great importance for design under uncertainty

like robust design and reliability-based design (Du and Chen 2002; Chen et al. 2015). A wide variety of UP techniques have been developed (Lee and Chen 2009) among which the polynomial chaos (PC) technique is one of the most popular approach due to its mathematically rigorous concept, strong theoretical basis, and inherent ability to converge to computer calculation precision (Eldred 2009). With PC, a stochastic quantity can be represented as a polynomial chaos expansion, based on which the statistical moments and reliability can be conveniently obtained. Oftentimes, the analysis models in practical engineering are highly nonlinear and computationally expensive, such as computational fluid dynamics (CFD) for aerodynamic analysis, resulting in intensively computational cost in implementing UP via PC, which becomes more serious for high-dimensional problems. Therefore, the straightforward application of PC on the expensive model for UP might be too costly and infeasible in practical application.

Generally, a complicated physical process can be modeled using several methods with different levels of fidelity, or a computer code for a complex problem can be run at different levels of fidelity. For example, the aircraft aerodynamic analysis can be simulated with different reduced physical order (e.g., Euler model vs. potential flow model) or different numerical solver (e.g., finite difference method vs. finite element analysis). A high-fidelity (HF) model takes more computational time but offers higher accuracy, whereas a low-fidelity (LF) model is faster at the cost of accuracy. The exploitation of the availability of multiple models within a hierarchy of fidelity is popular in assisting the process of optimization in engineering (Gratiet and Cannamela 2012; Huang et al. 2006; Gratiet et al. 2014). This scenario has been extended to UP via PC for improving computational efficiency recently, which has received considerable interest (Shah et al. 2015; Ng and Eldred 2012; Zhu et al. 2017; Zhu et al. 2014). The earliest work about the multi-fidelity PC method was proposed by Ng and Eldred (Ng and Eldred 2012), in which the stochastic collocation technique was employed to construct the PC model, and the LF and correction PC expansions are integrated in the form of addition, multiplication, or a combination of the two into a single expansion to match the HF model values. As stated by Ng and Eldred in their work (Ng and Eldred 2012), for the multiplicative and combinative correction forms, the calculation of the multi-fidelity polynomial coefficients is much more complicated and the accuracy is generally worse or comparable, and thus the additive form is widely studied and applied. This method with the additive form has been applied to UP for a vertical axis wind turbine under extreme gusts (Santiago Padron et al. 2014; Palar et al. 2018). Another similar technique is the multi-fidelity stochastic collocation that relies on Lagrange-polynomial interpolation, in which a greedy procedure based on the information from the LF model is used to collect "important" sample points for the HF simulations (Zhu et al. 2014, 2017).

For the widely studied method proposed by Ng and Eldred, the roots of the orthogonal polynomials are employed as the collocation points for the PC expansion. Therefore, the number and location of collocation points cannot be arbitrary, resulting in less flexibility in performing UP for the user with a limited computational budget. To address this issue, based on the work of Ng and Eldred, the multi-fidelity PC approach using regression has been developed (Pramudita et al. 2016), which has been applied to multidisciplinary design optimization under uncertainty (West and Gumbert 2017) and aerodynamic robust optimization (Palar et al. 2015). For all the above PC based multi-fidelity modeling approaches for UP, it is required that the HF sample points should be a subset of the LF ones (i.e., nested sample points). To relax this constraint and improve the flexibility, Berchier has proposed to calculate the correction expansion term using the low-fidelity PC model rather than the LF sample points (Matteo 2016).

In the deterministic case, the most well-known multi-fidelity modeling method is the multi-level co-kriging approach proposed by Kennedy and O'Hagan (short for KOH in this work), in which the discrepancy-based autoregressive multi-fidelity modeling formulation and Gaussian process (GP) modeling technique are employed (Kennedy and O'Hagan 2000). It has been widely recognized that KOH is more accurate and flexible compared to the classic additive and multiplicative correction forms for multi-fidelity modeling (Laurenceau and Sagaut 2008; Toal et al. 2011; Han et al. 2012; Huang et al. 2013; Toal and Keane 2015). One main reason is that the KOH framework employs the Gaussian process modeling method that can flexibly capture the nonlinearity of the model, and the scaling factor on the LF model that is more beneficial to improve the accuracy of the correction term and avoid the bumpy issue (Fernández-Godino et al. 2016). This prompts us to think about whether the KOH multi-fidelity modeling framework can be extended to the stochastic domain for UP, within which the multi-fidelity PC can be implemented to improve the performance of UP. As is well known, the basic theoretical foundation of KOH is the GP modeling method, and the predicted response in KOH can be represented as a GP. Recently, the PC method has been extended to polynomial-chaos-kriging (PC-Kriging) by adding a GP term to the original PC model to more accurately capture the local variability of the response model (Schobi et al. 2015). It has been demonstrated that the PC-Kriging approach is more accurate than PC as well as kriging in performing UP (Schobi et al. 2015; Kersaudy et al. 2015). Clearly, the success that introducing GP into PC for UP lays a foundation for the extension of KOH framework to multi-fidelity PC. Therefore, a multi-fidelity PC approach will be developed and studied using the KOH framework in this paper.

In addition, almost all the works of multi-fidelity UP focus on model fusion within a hierarchical fidelity of the models.

However, in many applications, it is not possible to rank models by their levels of fidelity a priori, exhibiting non-hierarchical fidelity. For example, the models of climate system are developed from different research groups to understand and predict its behavior, based on disparate theories or mechanisms to incorporate the physics and chemistry of the atmosphere, ocean, and land surface (Allaire and Willcox 2012). The non-hierarchical multi-fidelity modeling approach from the deterministic point of view has been developed by Chen et al. using the spatial random process (Chen et al. 2016), which will be extended to multi-fidelity PC within a non-hierarchal fidelity in this work.

It is the objective of this work to explore the applicability and effectiveness of the Gaussian process modeling theory on multi-fidelity UP via the PC technique. For the hierarchical fidelities, the well-known KOH framework is extended to multi-fidelity UP, in which the lowest-fidelity model and all the correction terms are respectively represented as a PC-Kriging model. For the non-hierarchical fidelities, the weighted summation method proposed by Chen et al. is extended, in which all the lower-fidelity models and the correction term are respectively represented as a PC-Kriging model. Meanwhile, for high-dimensional problems, the hyperbola truncation scheme is employed to reduce the number of the orthogonal polynomials during the construction of PC term in the PC-Kriging model, and thus to reduce the computational cost.

The remainder of this paper is organized as follows. A brief review of the PC-Kriging method combining PC and Gaussian process modeling for UP in given in Section 2. The proposed multi-fidelity UP method using PC and the Gaussian process modeling technique is presented, in which multi-fidelity UP strategies respectively for hierarchical and non-hierarchical fidelities are explained in detail. Comparative studies on numerical problems are presented in Section 4, where the commonly used co-kriging method (Kennedy and O'Hagan 2000) and multi-fidelity PC method (Ng and Eldred 2012) are also tested for comparison. In Section 5, the proposed method is applied to an aerodynamic robust optimization problem to further verify its effectiveness and applicability in dealing with practical problems. Conclusions are drawn in Section 6.

2 Review of PC-kriging

The polynomial-chaos-kriging (PC-Kriging) method is a newly developed approach for UP by adding a GP term to the PC model. As is well known, the PC method formulated as a weighted sum of a set of orthogonal polynomials can efficiently capture the global behavior of the analysis model. The introduction of the GP term helps to capture the local variability, thus to improve the accuracy of PC for UP. With PC-Kriging, a stochastic response $y = g(\mathbf{x})$ with a d-dimensional input vector $\mathbf{x} = [x_1, ..., x_d]$ can be represented as follows:

$$y \approx M^{(PCK)}(\mathbf{x}) = \sum_{i=0}^{P} b_i \Phi_i(\mathbf{x}(\xi)) + \sigma^2 Z(\mathbf{x}) \tag{1}$$

where $\sum_{i=0}^{P} b_i \Phi_i(\mathbf{x}(\xi))$ is a PC model with order p describing the mean value of the Gaussian process, ξ is a standard random vector generated by mapping the original random vector $\mathbf{x}$ to the standard random space according to the distribution parameters of $\mathbf{x}$, b_i is the coefficient of PC model; σ is the prior standard deviation of the Gaussian process, and $Z(\mathbf{x})$ is a zero-mean and unit-variance stationary Gaussian process with the autocorrelation function R.

The commonly used formulation of R in the literature is:

$$R\left(\mathbf{x}, \mathbf{x}', \boldsymbol{\theta}, h\right) = \exp\left(-\sum_{k=1}^{d} \theta_k \left|x_k - x_k'\right|^h\right) \tag{2}$$

Building a PC-Kriging stochastic metamodel consists of two parts: (i) the construction of $\Phi_i(\mathbf{x}(\xi))$ in the PC term, and (ii) the estimation of hyper-parameters $(\boldsymbol{\theta}, h, \sigma)$ and $\boldsymbol{b} = [b_0, ..., b_P]^T$. For the first part, the commonly used method is the direct tensor product technique. For high-dimensional problems, the least angle regression method (Wang et al. 2016) or the hyperbolic truncation scheme (Blatman and Sudret 2011) shown in (3) can be employed to remove some unimportant orthogonal polynomials, and thus to reduce the computational cost of PC:

$$A_q^{d,\omega} = \left\{ \boldsymbol{\alpha} \in \mathbb{N}^d : \|\boldsymbol{\alpha}\|_q = \left(\sum_{i=1}^{d} \alpha_i^q\right)^{\frac{1}{q}} \leq \omega \right\} \tag{3}$$

where A is the truncation set of multi-indices $\boldsymbol{\alpha}$ for PC, q is the sparse factor, ω is the highest order of polynomial $\Phi(\mathbf{x}(\xi))$, and α_i is the degree of the ith random variable in $\Phi(\mathbf{x}(\xi))$.

For the second part, the maximum likelihood estimation (MLE) method or the cross validation (CV) method can be employed (Schobi et al. 2015) by maximizing the following functions, respectively:

$$G_{ML}(\boldsymbol{\theta}, h) = \arg\min_{\theta, h} \left[\frac{1}{N}(\boldsymbol{y} - \mathbf{F}\boldsymbol{b})^T \mathbf{R}^{-1}(\boldsymbol{y} - \mathbf{F}\boldsymbol{b})(\det\mathbf{R})^{\frac{1}{N}}\right] \tag{4}$$

$$G_{CV}(\boldsymbol{\theta}, h) = \arg\min_{\theta, h} \left[\boldsymbol{y}^T \mathbf{R}^{-1} diag\left(\mathbf{R}^{-1}\right)^{-2} \mathbf{R}^{-1} \boldsymbol{y}\right] \tag{5}$$

where N is the number of sample points, $\boldsymbol{y} = [y_1, ..., y_N]^T$ is the response vector at the input sample points, $\mathbf{R}$ is the autocorrelation matrix with the ith row and j^{th} column element as $R_{ij} = R(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta}, h)$, and $\mathbf{F}$ is a $(P + 1) \times N$ matrix with $F_{ij} = \Phi_i(\boldsymbol{x}_j(\xi))(i = 0, 1, ..., P; j = 1, ..., N)$; the PC coefficient vector $\boldsymbol{b}$ can be represented as follows:

$$\boldsymbol{b}(\theta, h) = \left(\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F}\right)^{-1} \mathbf{F} \mathbf{R}^{-1} y \tag{6}$$

Once all the parameters are obtained, the predicted response at any new input site $\boldsymbol{x}_p$ can be calculated by (Rasmussen and Williams 2006):

$$y_{(PCK)}(\mathbf{x}_p) = \Phi\big(\mathbf{x}_P(\boldsymbol{\xi})\big)^T \boldsymbol{b} + R\big(\mathbf{x}_p, \mathbf{X}, \boldsymbol{\theta}, h\big)\mathbf{R}^{-1}(\boldsymbol{y} - \mathbf{F}\boldsymbol{b}) \quad (7)$$

where $\mathbf{X}$ is the matrix stacking all the collected input sample points.

Then, Monte Carlo simulation (MCS) can be directly employed on the PC-Kriging metamodel to obtain the statistical moments and probabilistic distribution of random response y.

3 The proposed multi-fidelity UP method

3.1 Multi-fidelity UP for hierarchical fidelity

It is assumed that there exist s analysis models with responses as $y^t(\mathbf{x})|_{t=1,\dots,s}$, and a larger t corresponds to a higher level of fidelity and larger computational cost. Therefore, $y^1(x)$ is the lowest-fidelity and cheapest model, while $y^s(\mathbf{x})$ is the highest fidelity and most expensive one. $\mathbf{x} = [x_1, x_2, \dots, x_d] \in \mathbb{R}^d$ represents a d-dimensional random input vector. A step-by-step description of the proposed multi-fidelity UP using PC and the Gaussian process modeling technique for the hierarchical fidelity scenario is presented as follows.

Step 1. According to the distribution information of random input vector $\mathbf{x}$, generate input sample points using methods such as the Latin Hypercube sampling or Gaussian quadrature points rule etc., and calculate the corresponding model responses with different levels of fidelity.

For the tth level ($t = 1,\dots,s$) model, it is supposed that a set of response observations $\mathbf{d}_t = \big[y^t(\mathbf{x}_1^t), \dots, y^t\big(\mathbf{x}_{n_t}^t\big)\big]^T$ at input sites $\mathbf{D}_t = \Big[(\mathbf{x}_1^t)^T, (\mathbf{x}_2^t)^T, \dots, \big(\mathbf{x}_{n_t}^t\big)^T\Big]^T$ have been collected. Let $\mathbf{d} = \big[\mathbf{d}_1^T, \dots, \mathbf{d}_s^T\big]^T$ denote all of the collected response data from all models at the input space $\boldsymbol{\Gamma} = [\mathbf{D}_1; \mathbf{D}_2; \dots; \mathbf{D}_s]$. Generally, the number of sample points n_t is decreased with the increase of t considering the computational cost.

Step 2. Construct the multi-fidelity PC-Kriging metamodel in replacement of the highest fidelity model $y^s(\mathbf{x})$ by extending the KOH framework to UP.

The KOH formulation in constructing multi-level co-kriging (Kennedy and O'Hagan 2000) is:

$$y^t(\mathbf{x}) = \rho_{t-1}y^{t-1}(\mathbf{x}) + \delta^t(\mathbf{x}), t = 2, \dots, s \quad (8)$$

where ρ_{t-1} represents the scaling factor between model responses $y^t(\mathbf{x})$ and $y^{t-1}(\mathbf{x})$, and the correction function $\delta^t(\mathbf{x})$ is a Gaussian process denoting the discrepancy between $y^t(\mathbf{x})$ and $\rho_{t-1}y^{t-1}(\mathbf{x})$.

Accordingly, the highest fidelity output response $y^s(\mathbf{x})$ can be expressed as below based on (8):

$$y^s(\mathbf{x}) = \left(\prod_{i=1}^{s-1}\rho_i\right)y^1(x) + \left(\prod_{i=2}^{s-1}\rho_i\right)\delta^2(\mathbf{x}) + \dots$$
$$+ \rho_{s-1}\delta^{s-1}(\mathbf{x}) + \delta^s(\mathbf{x}) \quad (9)$$

It is assumed that $y^1(\mathbf{x})$, $\delta^2(\mathbf{x})$, ..., $\delta^s(\mathbf{x})$ can be respectively modeled by a GP, which are represented as PC-Kriging metamodels as below, respectively:

$$\begin{cases} y^1(\mathbf{x}) = \sum_{i=0}^{P} b_i^1\Phi_i^1(\mathbf{x}(\boldsymbol{\xi})) + \sigma_1^2 Z^1(\mathbf{x}) \sim \mathcal{GP}\big(M^1(\mathbf{x}), V^1\big(\mathbf{x}, \mathbf{x}'\big)\big) \\ \delta^2(\mathbf{x}) = \sum_{i=0}^{P} b_i^2\Phi_i^2(\mathbf{x}(\boldsymbol{\xi})) + \sigma_2^2 Z^2(\mathbf{x}) \sim \mathcal{GP}\big(M^2(\mathbf{x}), V^2\big(\mathbf{x}, \mathbf{x}'\big)\big) \\ \quad\quad\quad \vdots \\ \delta^s(\mathbf{x}) = \sum_{i=0}^{P} b_i^s\Phi_i^s(\mathbf{x}(\boldsymbol{\xi})) + \sigma_s^2 Z^s(\mathbf{x}) \sim \mathcal{GP}\big(M^s(\mathbf{x}), V^s\big(\mathbf{x}, \mathbf{x}'\big)\big) \end{cases} \quad (10)$$

where $M(\mathbf{x}) = \sum_{i=0}^{P} b_i\Phi_i(\mathbf{x}(\boldsymbol{\xi}))$ is the mean function expressed as the weighted sum of $\Phi(\mathbf{x}(\boldsymbol{\xi}))$ (i.e., PC term), $V(\mathbf{x}, \mathbf{x}') = \sigma^2 R(\mathbf{x}, \mathbf{x}', \boldsymbol{\theta}, h)$ is the covariance function, representing the spatial covariance between any two inputs $\mathbf{x}$ and $\mathbf{x}'$ of the GP.

Then, additively, the highest fidelity model $y^s(\mathbf{x})$ can be further expressed by a GP based on (9) and (10) as follows:

$$y^s(\mathbf{x}) \sim \mathcal{GP}\Big(M^1(\mathbf{x})\Big(\prod_{i=1}^{s-1}\rho_i\Big) + M^2(\mathbf{x})\Big(\prod_{i=2}^{s-1}\rho_i\Big) + \dots + M^s(\mathbf{x}),$$
$$\Big(\prod_{i=1}^{s-1}\rho_i\Big)^2 V^1\big(\mathbf{x}, \mathbf{x}'\big) + \Big(\prod_{i=2}^{s-1}\rho_i\Big)^2 V^2\big(\mathbf{x}, \mathbf{x}'\big) + \dots + V^s\big(\mathbf{x}, \mathbf{x}'\big)\Big) \quad (11)$$

Step 3. Estimate all the unknown hyper-parameters by the maximum likelihood estimation method.

The unknown hyper-parameters in (11) are: $\boldsymbol{\Delta} = \{\mathbf{B}, \boldsymbol{\sigma}, \boldsymbol{\Theta}, \boldsymbol{\rho}, \boldsymbol{h}\}$, where $\mathbf{B} = [(\boldsymbol{b}^1)^T, \dots, (\boldsymbol{b}^i)^T, \dots, (\boldsymbol{b}^s)^T]^T$ is the polynomial coefficient matrix with each element as $\boldsymbol{b}^i = [b_0, b_1, \dots, b_P]^T$, and the rest can be expressed as $\boldsymbol{\sigma} = [\sigma^1, \dots, \sigma^s]^T$, $\boldsymbol{\Theta} = [\boldsymbol{\theta}^1, \dots, \boldsymbol{\theta}^s]^T$, $\boldsymbol{\rho} = [\rho^1, \dots, \rho^{s-1}]^T$, $\boldsymbol{h} = [h^1, \dots, h^s]^T$.

In this work, the method considering the full correlation of all the response models (Liu et al. 2018) is employed for parameter estimation. Then, one can obtain that all the collected data $\mathbf{d}$ follow a multivariate normal distribution based on the assumption of GP, i.e.:

$$\mathbf{d} \sim \mathcal{N}(\mathbf{HB}, \mathbf{V}_d) \quad (12)$$

where $\mathbf{H}$ is defined as:

$$
\mathbf{H} = \begin{bmatrix}
\Phi^1(\mathbf{D}_1(\xi)) & \mathbf{0} & \cdots & \mathbf{0} \\
\rho_1\Phi^1(\mathbf{D}_2(\xi)) & \Phi^2(\mathbf{D}_2(\xi)) & \mathbf{0} & \mathbf{0} \\
\rho_1\rho_2\Phi^1(\mathbf{D}_3(\xi)) & \rho_2\Phi^2(\mathbf{D}_3(\xi)) & \Phi^3(\mathbf{D}_3(\xi)) & \vdots \\
\vdots & \vdots & \vdots & \mathbf{0} \\
\left(\prod_{i=1}^{s-1}\rho_i\right)\Phi^1(\mathbf{D}_s(\xi)) & \left(\prod_{i=2}^{s-1}\rho_i\right)\Phi^2(\mathbf{D}_s(\xi)) & \cdots & \Phi^s(\mathbf{D}_s(\xi))
\end{bmatrix}
$$
(13)

and $\Phi^t(\mathbf{D}_j(\xi))$ is a matrix of size $n_j \times (P+1)$ $(j, t = 1,2, \ldots,s)$, formulated as:

$$
\Phi^t(\mathbf{D}_j(\xi)) = \left[\Phi_0^t(\mathbf{D}_j(\xi)), \Phi_1^t(\mathbf{D}_j(\xi)), \ldots, \Phi_P^t(\mathbf{D}_j(\xi))\right]
$$
(14)

The matrix $\mathbf{V}_d$ of size $(n_1+, \ldots, +n_s) \times (n_1+, \ldots, +n_s)$ is given as:

$$
\mathbf{V}_d = \begin{pmatrix}
V_{1,1} & \cdots & V_{1,s} \\
\vdots & \ddots & \vdots \\
V_{s,1} & \cdots & V_{s,s}
\end{pmatrix}
$$
(15)

where the tth diagonal block $(n_t \times n_t)$ is defined as:

$$
V_{t,t} = \sigma_t^2 R_t(\mathbf{D}_t) + \sigma_{t-1}^2\rho_{t-1}^2 R_{t-1}(\mathbf{D}_t) + \ldots
$$
$$
+ \sigma_1^2\left(\prod_{i=1}^{t-1}\rho_i^2\right)R_1(\mathbf{D}_t)
$$
(16)

with $R_i(\mathbf{D}_t) = R_i(\mathbf{D}_t, \mathbf{D}_t, \theta^i, h^i)$, and the off-diagonal block of size $n_t \times n_{t'}$ is given by:

$$
V_{t,t'} = \left(\prod_{i=t}^{t'-1}\rho_i\right)\left(\sigma_t^2 R_t(\mathbf{D}_t, \mathbf{D}_{t'}) + \ldots + \sigma_1^2\left(\prod_{i=1}^{t-1}\rho_i^2\right)R_1(\mathbf{D}_t, \mathbf{D}_{t'})\right)
$$
$$
1 \leq t < t' \leq s
$$
(17)

The maximum likelihood function is defined as follows:

$$
\mathcal{L}(\Delta|\mathbf{d}) \propto |\mathbf{V}_d|^{-1/2}|\mathbf{W}|^{1/2}exp\left\{-\frac{1}{2}(\mathbf{d}-\mathbf{HB})^T\mathbf{V}_d^{-1}(\mathbf{d}-\mathbf{HB})\right\}
$$
(18)

where $\mathbf{W} = \left(\mathbf{H}^T\mathbf{V}_d^{-1}\mathbf{H}\right)^{-1}$, and the PC coefficient matrix $\mathbf{B}$ in the PC-Kriging model can be derived using the first order optimal condition:

$$
\widehat{\mathbf{B}} = \mathbf{WHV}_d^{-1}\mathbf{d}
$$
(19)

The rest parameters σ, Θ, ρ and h can be obtained using genetic algorithm or simulated annealing algorithm by maximizing (18).

Step 4. MCS is conducted on the constructed multi-fidelity PC-Kriging metamodel above to obtain the stochastic property of random output y.

Based on the GP modeling theory, all the collected data $\mathbf{d}$ together with the to-be-predicted responses $\mathbf{y}(\mathbf{x}_p) = \left[y(\mathbf{x}_1), \ldots, y(\mathbf{x}_{n_p})\right]^T$ follow a multivariate normal distribution:

$$
\begin{bmatrix}\mathbf{d} \\ y(\mathbf{x}_p)\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\mathbf{H} \\ \mathbf{H}_p\end{bmatrix}\mathbf{B}, \begin{bmatrix}\mathbf{V}_d & \mathbf{T}_p^T \\ \mathbf{T}_p & \mathbf{V}_p\end{bmatrix}\right)
$$
(20)

Then, the final prediction of $y(x_p)$ can be calculated by:

$$
\widehat{y}(\mathbf{x}_p) = \mathbf{H}_p\widehat{\mathbf{B}} + \mathbf{T}_p\mathbf{V}_d^{-1}\left(\mathbf{d}-\mathbf{H}\widehat{\mathbf{B}}\right)
$$
(21)

where:

$$
\mathbf{H}_p = \begin{pmatrix}\left(\prod_{i=1}^{s-1}\rho_i\right)\Phi^1(\mathbf{x}_p(\xi)), \left(\prod_{i=2}^{s-1}\rho_i\right)\Phi^2(\mathbf{x}_p(\xi)), \ldots, \\ \rho_{s-1}\Phi^{s-1}(\mathbf{x}_p(\xi)), \Phi^s(\mathbf{x}_p(\xi))\end{pmatrix}
$$
(22)

$$
\mathbf{T}_p = \left(t_1(x_p, \mathbf{D}_1)^T, \ldots, t_s(x_p, \mathbf{D}_s)^T\right)^T
$$
(23)

$$
t_1(\mathbf{x}_p, \mathbf{D}_1) = \left(\prod_{i=1}^{s-1}\rho_i\right)\sigma_1^2 R_1(\mathbf{x}_p, \mathbf{D}_1)^T
$$
(24)

$$
t_t(\mathbf{x}_p, \mathbf{D}_t) = \rho_{t-1}t_t(\mathbf{x}_p, \mathbf{D}_t) + \left(\prod_{i=t}^s\rho_i\right)\sigma_t^2 R_t(\mathbf{x}_p, \mathbf{D}_t)^T, t = 2, \ldots, s
$$
(25)

$$
\mathbf{V}_p = V_d^s(\mathbf{x}_p, \mathbf{x}_p) + \rho_{s-1}^2 V_d^{s-1}(\mathbf{x}_p, \mathbf{x}_p)
$$
$$
+ \rho_{s-1}^2\rho_{s-2}^2 V_d^{s-2}(\mathbf{x}_p, \mathbf{x}_p) + \ldots
$$
$$
+ \left(\prod_{t=2}^{s-1}\rho_t\right)V_d^2(\mathbf{x}_p, \mathbf{x}_p)
$$
(26)

MCS is employed directly on (21) to calculate the mean, standard deviation and probabilistic distribution, etc., of the random response y.

3.2 Multi-fidelity UP for non-hierarchical fidelity

For the non-hierarchical fidelity case, it is assumed that there are m lower-fidelity models with responses as $y^q(\mathbf{x})|_{q = 1, \ldots m}$, of which the fidelity cannot be ranked in advance, and one high-fidelity model $y^H(\mathbf{x})$.

Step 1. Similar to the step 1 during hierarchical multi-fidelity modeling in Section 3.1, collect input data $\mathbf{D}_q = \left[(\mathbf{x}_1^q)^T, (\mathbf{x}_2^q)^T, \ldots, \left(\mathbf{x}_{n_q}^q\right)^T\right]^T$ and output response data $\mathbf{d}_q = \left[y^q(\mathbf{x}_1^q), \ldots, y^q\left(\mathbf{x}_{n_q}^q\right)\right]^T$ for each lower-fidelity model. Let $\mathbf{d}_L = \left[\mathbf{d}_1^T, \ldots, \mathbf{d}_m^T\right]^T$ denote the

collected response data from all lower-fidelity models at the input sites $\Gamma_L = [\mathbf{D}_1; \mathbf{D}_2; ...; \mathbf{D}_m]$, and $\mathbf{d}_H = \left[y^H\left(\mathbf{x}_1^H\right), ..., y^H\left(\mathbf{x}_{n_H}^H\right) \right]^T$ denote the collected response data from high-fidelity model at input sites $\mathbf{D}_H = \left[\left(\mathbf{x}_1^H\right)^T, ..., \left(\mathbf{x}_{n_H}^H\right)^T \right]^T$.

Step 2. Construct the multi-fidelity PC-Kriging metamodel in replacement of $y^H(\mathbf{x})$.

According to the weighted summation method (Chen et al. 2016), represent $y^H(\mathbf{x})$ as an addition of the weighted summation of all the lower-fidelity models $y^q(\mathbf{x})|_{q=1, ...m}$ and a residual deviation function $\delta(\mathbf{x})$:

$$y^H(\mathbf{x}) = \sum_{q=1}^{m} \rho^q y^q(\mathbf{x}) + \delta(\mathbf{x}) \tag{27}$$

where ρ^q denotes the weighting coefficient of model $y^q(\mathbf{x})$.

It is assumed that all the lower-fidelity models $y^q(\mathbf{x})|_{q=1, ...m}$ and $\delta(x)$ are priori independent and can be respectively represented as a GP to simplify the model fusion process. Construct the stochastic metamodel for each lower-fidelity model $y^q(\mathbf{x})|_{q=1, ...m}$ using the PC-Kriging method, during which the same Gaussian correlation function $R(x, x', \theta, h)$ is employed considering that all the lower-fidelity models describe the same physical process. Similarly, construct the PC-Kriging metamodel for $\delta(x)$ with the Gaussian correlation function $R^\delta(\mathbf{x}, \mathbf{x}', \boldsymbol{\theta}^\delta, h^\delta)$.

Based on (27), $y^H(\mathbf{x})$ can be further expressed as a GP additively as follows:

$$y^H(\mathbf{x}) \sim \mathcal{GP}\left(M(\mathbf{x}), V\left(\mathbf{x}, \mathbf{x}'\right) \right) \tag{28}$$

where $M(\mathbf{x}) = \sum_{q=1}^{m} \boldsymbol{\Phi}^q(\mathbf{x})^T \boldsymbol{b}^q \rho^q + \boldsymbol{\Phi}^\delta(\mathbf{x})^T \boldsymbol{b}^\delta$, $V(\mathbf{x}, \mathbf{x}') = \boldsymbol{\rho}^T \mathbf{E} \boldsymbol{\rho} R(\mathbf{x}, \mathbf{x}') + \sigma_\delta^2 R^\delta(\mathbf{x}, \mathbf{x}')$, $\boldsymbol{\rho} = [\rho^1, ..., \rho^m]^T$, $\mathbf{E} = \begin{bmatrix} \mathbf{E}_{1,1} & \cdots & \mathbf{E}_{1,m} \\ \vdots & & \vdots \\ \mathbf{E}_{m,1} & \cdots & \mathbf{E}_{m,m} \end{bmatrix}$, $\mathbf{E}_{i,j}$ is the unknown covariance between lower-fidelity models $y^i(\mathbf{x})$ and $y^j(\mathbf{x})$ calculated by $\mathbf{E}_{i,j} = c_{i,j} \sqrt{\mathbf{E}_{i,i} \mathbf{E}_{j,j}}$, and $c_{i,j} \in [-1, 1]$ is the unknown correlation coefficient.

Step 3. Estimate the hyper-parameters by the maximum likelihood estimation method.

The hyper-parameters to be estimated are $\Delta = \{b^1, ..., b^m, b^\delta, E, \rho, \theta, \theta^\delta, h, h^\delta\}$. Similarly to the step 3 during the hierarchical multi-fidelity modeling, all the collected response data $d = [d_L; d_H]$ follow a multivariate Gaussian distribution as shown in (12), and some related matrices are re-defined as follows:

$$\mathbf{B} = \left[\left(\boldsymbol{b}^1\right)^T, ..., \left(\boldsymbol{b}^m\right)^T, \left(\boldsymbol{b}^\delta\right)^T \right]^T \tag{29}$$

$$\mathbf{H} = \begin{bmatrix} \boldsymbol{\Phi}^1(\mathbf{D}_1(\xi)) & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \boldsymbol{\Phi}^m(\mathbf{D}_m(\xi)) & \mathbf{0} \\ \rho^1 \boldsymbol{\Phi}^1(\mathbf{D}_H(\xi)) & \cdots & \rho^m \boldsymbol{\Phi}^m(\mathbf{D}_H(\xi)) & \boldsymbol{\Phi}^\delta(\mathbf{D}_H(\xi)) \end{bmatrix} \tag{30}$$

$$\mathbf{V}_d = \begin{bmatrix} e_1^T \mathbf{E} e_1 R(\mathbf{D}_1, \mathbf{D}_1) & \cdots & e_1^T \mathbf{E} e_m R(\mathbf{D}_1, \mathbf{D}_m) & e_1^T \mathbf{E} \boldsymbol{\rho} R(\mathbf{D}_1, \mathbf{D}_H) \\ \vdots & \ddots & \vdots & \vdots \\ e_m^T \mathbf{E} e_1 R(\mathbf{D}_m, \mathbf{D}_1) & \cdots & e_m^T \mathbf{E} e_m R(\mathbf{D}_m, \mathbf{D}_m) & e_m^T \mathbf{E} \boldsymbol{\rho} R(\mathbf{D}_m, \mathbf{D}_H) \\ \boldsymbol{\rho} \mathbf{E} e_1 R(\mathbf{D}_H, \mathbf{D}_m) & \cdots & \boldsymbol{\rho} \mathbf{E} e_m R(\mathbf{D}_H, \mathbf{D}_m) & \begin{matrix} \boldsymbol{\rho}^T \mathbf{E} \boldsymbol{\rho} R(\mathbf{D}_H, \mathbf{D}_H) \\ + \sigma_\delta^2 R^\delta(\mathbf{D}_H, \mathbf{D}_H) \end{matrix} \end{bmatrix} \tag{31}$$

where e_i is an m-dimensional unit column vector, with the ith element as 1, while the others as zeros.

Then, using the same MLE method in Section 3.1, the hyper-parameters can be estimated.

Step 4. Predict response value at the sample points $\mathbf{x}_p$ similarly to the step 4 during the hierarchical multi-fidelity modeling in Section 3.1. Meanwhile, some matrices are re-defined as follows:

$$\mathbf{H}_p = \left[\rho^1 \boldsymbol{\Phi}^1\left(\mathbf{x}_p(\xi)\right), ..., \rho^m \boldsymbol{\Phi}^m\left(\mathbf{x}_p(\xi)\right), \boldsymbol{\Phi}^\delta\left(\mathbf{x}_p(\xi)\right) \right] \tag{32}$$

$$\mathbf{T}_p = \left[\boldsymbol{\rho}^T \mathbf{E} e_1 R(\mathbf{x}_p, \mathbf{D}_1), ..., \boldsymbol{\rho}^T \mathbf{E} e_m R(\mathbf{x}_p, \mathbf{D}_m), \boldsymbol{\rho}^T \mathbf{E} \boldsymbol{\rho} R(\mathbf{x}_p, \mathbf{D}_H) + \sigma_\delta^2 R^\delta(\mathbf{x}_p, \mathbf{D}_H) \right] \tag{33}$$

$$\mathbf{V}_p = \boldsymbol{\rho}^T \mathbf{E} \boldsymbol{\rho} R\left(\mathbf{x}_p, \mathbf{x}_p\right) + \sigma_\delta^2 R^\delta\left(\mathbf{x}_p, \mathbf{x}_p\right) \tag{34}$$

4 Comparative studies

The effectiveness of the proposed two multi-fidelity modeling approaches is tested by some mathematical examples with different nonlinearity and random input dimension for UP in this section.

4.1 Test for hierarchical fidelity

For hierarchical fidelity, the existing multi-fidelity PC method with the addition form proposed by Ng and Eldred (Ng and Eldred 2012) (denoted as MF-PC) and co-kriging that has been commonly employed for metamodeling in the deterministic domain (Kennedy and O'Hagan 2000) are also employed for UP and compared to the proposed method in this work (denoted as MF-PCK). The tested examples are shown in Table 1, in which $\mathcal{B}, \mathcal{U}$, and $\mathcal{N}$ respectively represent beta, uniform, and normal distribution. Examples 1 and 3–6 are adopted from the papers about multi-fidelity modeling in

Table 1 Tested examples with hierarchical fidelity

No.	Functions	Distributions
1 (Park et al. 2016; Ren et al. 2016)	$y^2 = (6x-2)^2 \sin(12x-4)$ $y^1 = 0.5(6x-2)^2 \sin(12x-4) + 8(x-0.5) + 5$	Case1: $x \sim \mathcal{N}$ $(0.5, 0.5)$ Case2: $x \sim \mathcal{B}(5, 0.75)$
2	$y^2 = \sin(x) + \cos(2x)$ $y^1 = \sin(x) + \cos(2x) + 0.1x$	$x \sim \mathcal{N}(2.5, 0.83)$
3 (Liu et al. 2018)	$y^3(x) = \sin(x) + 0.2x + (x-0.5)^2/16 + 0.5$ $y^2(x) = \sin(x) + 0.8x + (x-0.5)^2/45 + 0.5$ $y^1(x) = \sin(x) + 0.2x + 0.5$	$x \sim \mathcal{N}(5, 1.33)$
4 (Liu et al. 2018)	$y^3 = \left(1 - e^{\frac{-1}{2x_2}}\right)\left(\frac{1000tf \times x_1^3 + 1900x_1^2 + 2092x_1 + 60}{1000tl \times x_1^3 + 500x_1^2 + 4x_1 + 20}\right)$ $y^2 = \left(1 - e^{\frac{-1}{2x_2}}\right)\left(\frac{1000tf \times x_1^3 + 1900x_1^2 + 2092x_1 + 60}{1000tl \times x_1^3 + 500x_1^2 + 4x_1 + 20}\right) + \frac{5e^{-tf}x_1^{th/2}}{x_2^{2+th} + 1}$ $y^1 = \left(1 - e^{\frac{-1}{2x_2}}\right)\left(\frac{1000tf \times x_1^3 + 1900x_1^2 + 2092x_1 + 60}{1000tl \times x_1^3 + 500x_1^2 + 4x_1 + 20}\right) + \frac{5e^{-tf}x_1^{th/2}}{x_2^{2+th} + 1}$ $\quad + \frac{10x_1^2 + 4x_2^2}{50x_1 x_2 + 10}$ $tf = 0.2; th = 0.3; tl = 0.1;$	$x_1 \sim \mathcal{U}(0,1);$ $x_2 \sim \mathcal{U}(0,1)$
5 (Li et al. 2018)	$y^2 = 25(x_1-2)^2 + (x_2-2)^2 + (x_3-1)^2 + (x_4-4)^2$ $\quad + (x_5-1)^2 + (x_6-4)^2$ $y^1 = 15(x_1-2)^2 + 0.85(x_2-2)^2 + 0.6(x_3-1)^2 + 1.35(x_4-4)^2$ $\quad + 0.6(x_5-1)^2 + 0.6(x_6-4)^2$	$x_i \sim \mathcal{U}(0, 10)$ $i = 1, 2, \ldots, 6$
6 (Park et al. 2016; Ren et al. 2016)	$y^2 = \dfrac{2\pi T_u (H_u - H_l)}{\log(r/r_\omega)\left[1 + \frac{2LT_u}{\log(r/r_\omega)r_\omega^2 K_\omega} + T_u/T_l\right]}$ $y^1 = \dfrac{5\pi T_u (H_u - H_l)}{\log(r/r_\omega)\left[1.5 + \frac{2LT_u}{\log(r/r_\omega)r_\omega^2 K_\omega} + T_u/T_l\right]}$	See Table 2

literature, and example 2 is created by authors that have a simple correction term to explore the effect of the scale factor in MF-PCK. Meanwhile, cases with nested ($D_{i+1} \subseteq D_i$, $i = 1, \ldots, s$-1) and non-nested ($D_{i+1} \not\subseteq D_i$, $i = 1, \ldots, s$-1) sample points are both tested to explore the impact of sample property on the accuracy of multi-fidelity modeling. The results generated by conducting MCS on the original high-fidelity response function (denoted as direct MCS for simplicity) is employed as the benchmark to validate the effectiveness of the proposed methods.

Considering that high-fidelity sample points are often much fewer than low-fidelity ones, for examples 1, 2, 5, and 6 with a two-level fidelity, the order of PC terms for the low-fidelity model $y^1(\mathbf{x})$ and the correction term between two models are set as 5 and 3 (for $\delta^2(x)$), respectively. For examples 3 and 4 with a three-level fidelity, the order of PC terms for lower-fidelity model $y^1(\mathbf{x})$ and correction terms between three models are respectively set as 5, 3 (for $\delta^2(x)$), and 2 (for $\delta^3(x)$). The number of sample points employed for each response model during multi-fidelity modeling for all the examples is listed in Table 3,

Table 2 Random variables and distributions for example 6

Random variable	Distribution	Random variable	Distribution
r_ω	$\mathcal{U}(0.05, 0.15)$	T_l	$\mathcal{U}(63.1, 116)$
r	$\mathcal{U}(100, 50000)$	H_l	$\mathcal{U}(700, 820)$
T_u	$\mathcal{U}(63700, 115600)$	L	$\mathcal{U}(1120, 1680)$
H_u	$\mathcal{U}(990, 1100)$	K_ω	$\mathcal{U}(9855, 12045)$

Table 3 The number of sample points for each response model

Example	1	2	3	4	5	6
n_1	15	15	20	30	50	100
n_2	6	6	10	25	40	80
n_3	\	\	5	10	\	\

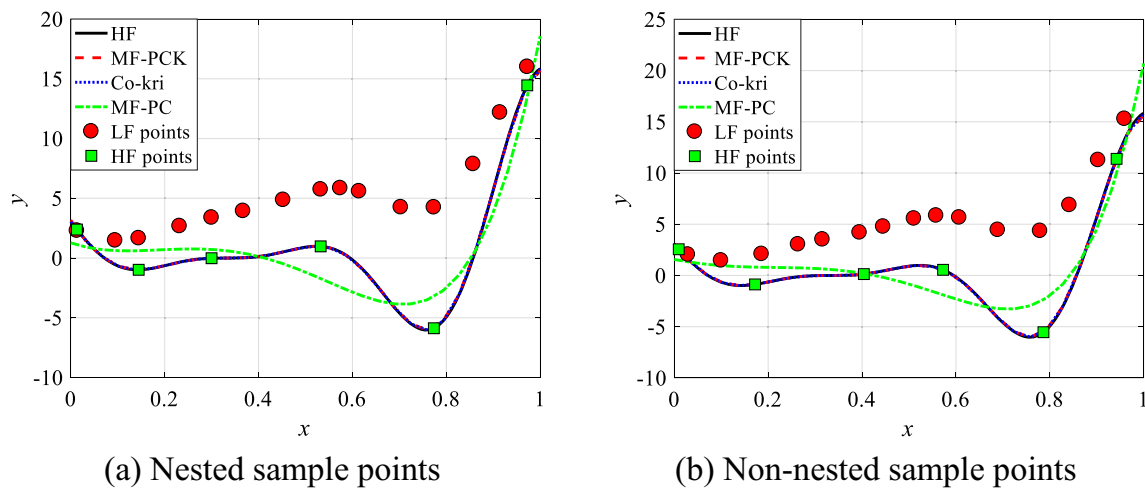(a) Nested sample points

(b) Non-nested sample points

Fig. 1 Function response curves of example 1 (case 1)

considering the varying nonlinearity and complexity of each example.

Meanwhile, as it is relatively a high-dimensional problem ($d = 6$), the sparse index in (3) is set as $q = 0.5$ for example 5 and $q = 0.25$ for example 6 during the construction of all PC models/terms for both MF-PCK and MF-PC to save computational cost.

The first four order statistical moments (mean, variance, skewness, and kurtosis) of the function response for the three multi-fidelity UP methods (MF-PCK, co-kriging, and MF-PC) are calculated through conducting MCS on the constructed multi-fidelity metamodel, which are then compared to those obtained by direct MCS. Considering that the sample points for each model are generated randomly during the tests, the results of UP may vary each time. Therefore, simulation is repeatedly conducted on each example for 20 times to avoid the impact of sample randomness on UP. Taking one of the 20 simulations for illustration, the curves/surfaces of the function responses produced by the three methods for examples 1–4 (one- or two-dimensional problems) are respectively shown in Figs. 1, 3, 5, 7, and 9, in which HF, MF-PCK, Co-kri, and MF-PC respectively denote the curve generated by the high-fidelity model, the proposed MF-PCK method, co-kriging, and the existing MF-PC method. The errors of the first four order statistical moments relative to those of direct MCS are

calculated as well, of which the mean values of the 20 simulations for all the examples are shown in Tables 4, 5, 6, 7, 8, 9, and 10, respectively.

During our tests, it is found that the proposed MF-PCK method can produce results that are very close to those of direct MCS, exhibiting high accuracy. However, with the same sample points as the three methods, the errors of the existing MF-PC method are evidently much larger than those of the other two approaches (MF-PCK and co-kriging) for most examples. Therefore, another UP test is conducted on MF-PC by increasing the number of high-fidelity sample points, during which only examples 1 and 2 are tested considering the space limit of this paper and the number of high-fidelity sample points are increased to 10 and 12, respectively. Correspondingly, the orders of PC for low-fidelity model and correction term are increased to 10 and 8, respectively. The results of this test are shown in Tables 4, 5, and 6, in which "A," "B," "C," and "CI" respectively denote the results produced by MF-PCK, co-kriging, MF-PC, and MF-PC with increased high-fidelity sample points; "1" and "2" respectively denote the case with nested and non-nested sample points for simplicity; e_m, e_v, e_s, and e_k, respectively denote the relative error of mean, variance, skewness, and kurtosis of the output response with respect to MCS.

Table 4 Relative errors of statistical moments for example 1 (case 1)

Methods Errors	A-1	B-1	C-1	CI-1	A-2	B-2	C-2	CI-2
e_m (%)	0.21	0.47	142.12	1.33	1.30	1.30	42.54	2.32
e_v (%)	0.32	0.32	12.93	0.01	0.17	0.47	17.95	0.79
e_s (%)	0.08	0.65	40.88	3.28	0.50	0.90	42.58	9.79
e_k (%)	0.16	1.34	45.40	6.64	0.68	1.68	47.60	12.27

Table 5 Relative errors of statistical moments for example 1 (case 2)

Methods Errors	A-1	B-1	C-1	CI-1	A-2	B-2	C-2	CI-2
e_m (%)	0.14	0.58	7.38	0.51	0.24	1.09	7.59	2.97
e_v (%)	0.11	1.88	1.40	2.42	0.58	2.75	0.13	5.48
e_s (%)	1.88	12.29	211.90	25.93	1.92	17.48	226.86	38.36
e_k (%)	0.07	0.46	9.42	1.52	0.13	0.64	10.58	0.84

Table 6 Relative errors of statistical moments for example 2

Methods Errors	A-1	B-1	C-1	CI-1	A-2	B-2	C-2	CI-2
e_m (%)	2.23e-7	0.01	0.04	0.02	0.01	0.05	0.10	0.04
e_v (%)	0.07	0.13	0.23	0.25	0.02	0.18	6.94	0.12
e_s (%)	0.03	0.03	16.24	0.33	**0.24**	**0.18**	16.76	0.07
e_k (%)	**0.11**	**0.07**	18.73	0.23	**0.41**	**0.23**	19.39	0.05

Table 8 Relative errors of statistical moments for example 4

Methods Errors	A-1	B-1	C-1	A-2	B-2	C-2
e_m (%)	0.02	0.24	0.01	0.06	0.35	1.50
e_v (%)	0.13	0.40	1.45	0.45	0.65	6.97
e_s (%)	1.43	4.79	1.94	2.74	7.42	58.14
e_k (%)	0.49	0.96	1.48	1.76	2.32	13.95

To more clearly show the robustness of the three multi-fidelity UP methods, the calculated statistical moments for all the examples are shown as boxplots in Figs. 2, 4, 6, 8, 10, 11, and 12. In the boxplots, the results of direct MCS are marked as MCS, and the black dot represents the mean value of the obtained data from the repeated 20 times UP and the red star represents the abnormal value (Figs. 6, 8, and 12). The top and bottom edges of the box are the upper and lower quartile of the data from the 20 simulations, respectively. And the line within the box represents the median. Generally, the smaller the distance between the top and bottom edges of the box is, the more concentrated the data is, and the more robust this method exhibits. From these tables (relative errors of statistical moments) and figures (function response curves/surfaces and boxplots of statistical moments), some noteworthy observations can be made.

Firstly, it is noticed that for both nested and non-nested sample points with the same computational cost, the proposed MF-PCK method is generally the most accurate, followed by co-kriging and then MF-PC, and MF-PC produces relatively large errors compared to the other two approaches.

MF-PCK vs. co-kriging MF-PCK is generally more accurate than co-kriging. Especially for examples 1, 4, 5, and 6, MF-PCK is evidently much more accurate. The interpretation is that although both approaches employ the GP modeling theory within the KOH framework to construct metamodel for UP, a PC model and a constant number are respectively adopted to capture the global trend of the stochastic output response for MF-PCK and co-kriging. As is well known, the PC model can deal with various random input distribution types including symmetric and unsymmetric ones and high nonlinearity of output response. Therefore, the conjunction of PC and GP in MF-PCK is more accurate in doing UP. Specially, for example 1 (case 2), as beta distribution (unsymmetric) is considered that is difficult to be handled by co-kriging, the results of the proposed MF-PCK method are much more accurate. For example 6, it is a high-dimensional problem and the variation of output response is large, which is very difficult to be approximated accurately by co-kriging. Therefore, co-kriging produces very large errors (see e_v, e_s, and e_k). However, it is observed that for example 2, although both approaches can produce accurate results that are very close to those of direct MCS, some errors (e_s and e_k) of co-kriging are slightly smaller than MF-PCK (see the bold numbers with underline in Table 6). The reason is that for example 2, normal distribution (symmetrical) is considered and the variation of output response is relatively small, and thus it is easier for co-kriging to describe the output response function with high accuracy. Meanwhile, for MF-PCK, as PC is included, more parameters should be estimated, which may induce certain numerical error during parameter estimation.

MF-PCK vs. MF-PC The proposed MF-PCK method is clearly more accurate than the existing MF-PC method. With the fusion of some lower-fidelity data, the accuracy of predicted response curves/surfaces as well as UP can be evidently improved with only a few HF data within the proposed MF-PCK multi-fidelity framework. Although

Table 7 Relative errors of statistical moments for example 3

Methods Errors	A-1	B-1	C-1	A-2	B-2	C-2
e_m (%)	0.03	0.10	0.03	0.07	0.10	8.46
e_v (%)	0.09	0.24	0.34	0.09	0.25	36.16
e_s (%)	0.12	0.42	2.03	0.26	0.66	20.24
e_k (%)	0.08	0.59	6.08	0.26	1.02	34.21

Table 9 Relative errors of statistical moments for example 5

Methods Errors	A-1	B-1	C-1	A-2	B-2	C-2
e_m (%)	0.07	1.18	4.22	0.19	0.41	20.21
e_v (%)	0.14	7.15	7.29	0.44	1.43	58.14
e_s (%)	0.09	5.71	10.60	0.65	6.73	8.64
e_k (%)	0.27	6.55	8.01	0.23	7.98	6.12

Table 10 Relative errors of statistical moments for example 6

Methods Errors	A-1	B-1	C-1	A-2	B-2	C-2
e_m (%)	0.08	0.17	9.83	0.10	0.19	9.95
e_v (%)	0.65	99.97	42.24	0.73	96.65	42.34
e_s (%)	7.57	2007.57	21.88	9.86	1.40e4	21.94
e_k (%)	3.64	1.83e6	1.51	5.28	4.15e6	1.61

MF-PC can be expected to produce comparable results (see the errors of CI-1 and CI-2 with increased HF sample points in Tables 4, 5, and 6) to MF-PCK, the number of HF sample points should be clearly increased. The interpretation is that the highly efficient KOH framework is employed for model fusion for MF-PCK, in which a scale factor ρ can be optimized. The introduction of ρ can clearly reduce the nonlinearity and bumpy issue of the correction term $\delta^t(x)$ (Park et al. 2016; Ren et al. 2016), which can certainly improve the accuracy of multi-fidelity metamodel. However, for MF-PC, ρ is actually set as a

constant $\rho = 1$. Taking examples 1 and 2 for illustration here, as ρ calculated by MF-PCK is $\rho = 2$ for example 1, the correction term $\delta^t(x)$ is completely linear, which is easy to be approximated by PC-Kriging. However, for MF-PC, as $\rho = 1$, a nonlinear correction term $\delta^t(x)$ is produced, which would evidently increase the difficulty in approximation. For example 2, as ρ employed by MF-PCK and MF-PC are both $\rho = 1$, the correction terms $\delta^t(x)$ approximated by the two approaches are close to each other. Therefore, the accuracy of MF-PC is clearly improved compared to that of example 1. In addition, a GP term is added to the PC model to enhance the local approximation ability of PC for the proposed MF-PCK method, which can also improve the accuracy compared to MF-PC.

Secondly, from the boxplots shown in Figs. 2, 4, 6, 8, 10, and 11, it is observed that for MF-PCK and co-kriging with nested and non-nested sample points, the top and bottom edges of the box generally are very close to each other, and some even seem overlapped (Figs. 6 and 8), and the distances between the two edges are much smaller

(a) Mean

(b) Variance

(c) Skewness

(d) Kurtosis

Fig. 2 Boxplots of statistical moments for example 1 (case 1)

(a) Nested sample points
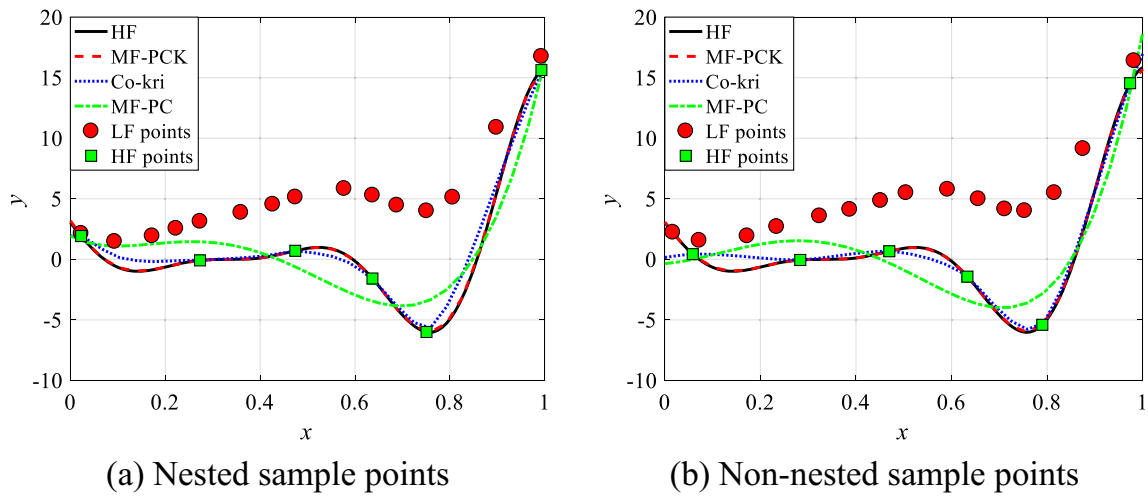
(b) Non-nested sample points
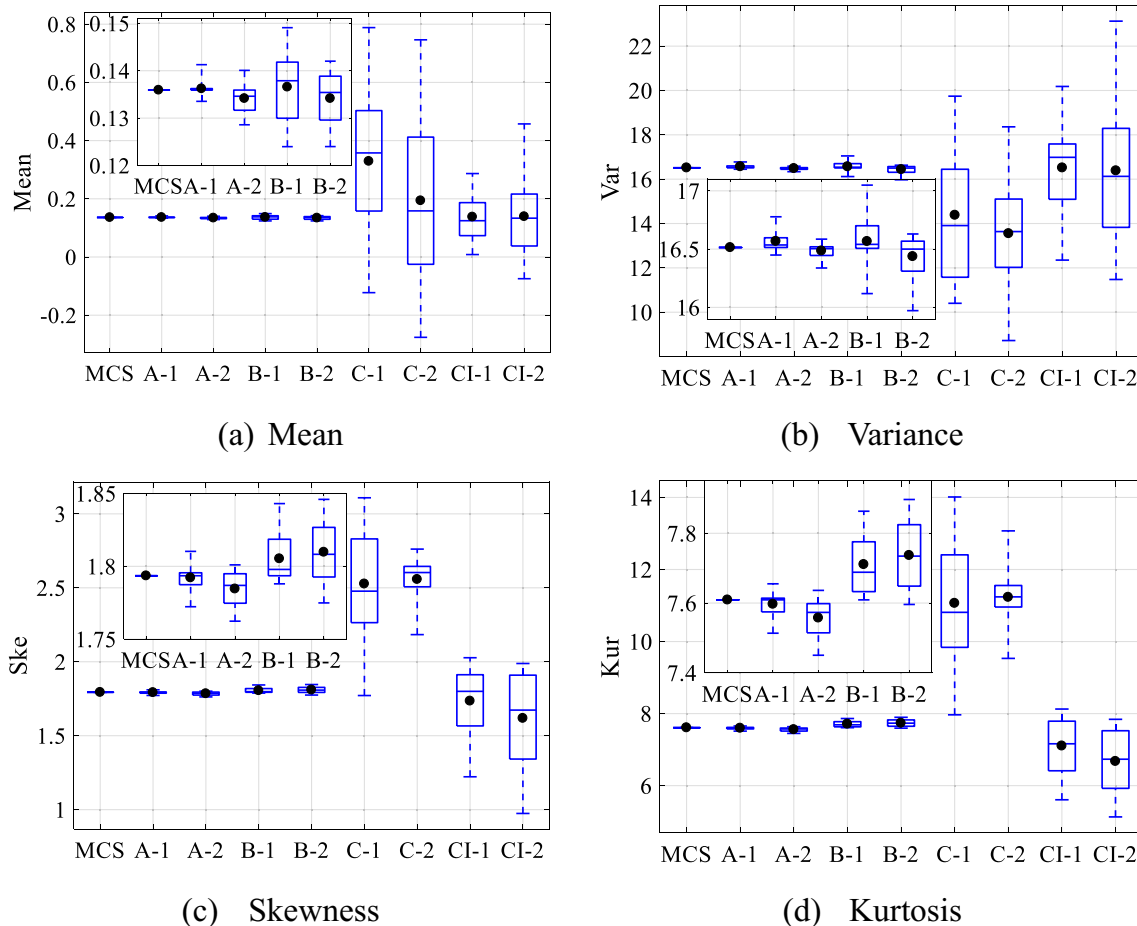
Fig. 3 Function response curves of example 1 (case 2)

than those of MF-PC with the same sample points. This indicates that MF-PCK and co-kriging are more robust and stable than MF-PC for UP as the data from the 20 simulations for the two approaches are more concentrated, which is attributed to the employment of the KOH framework in the two approaches. Although, MF-PC can also produce concentrated

and comparable results, the order of PC and number of high-fidelity sample points are required to be increased greatly (see the boxplots of CI-1 and CI-2 for examples 1 and 2). Meanwhile, it is also observed that the distances between the two edges for MF-PCK are basically smaller than those of co-kriging, indicating that MF-PCK is more robust and stable

(a) Mean

(b) Variance

(c) Skewness

(d) Kurtosis

Fig. 4 Boxplots of statistical moments for example 1 (case 2)

(a) Nested sample points   (b) Non-nested sample points
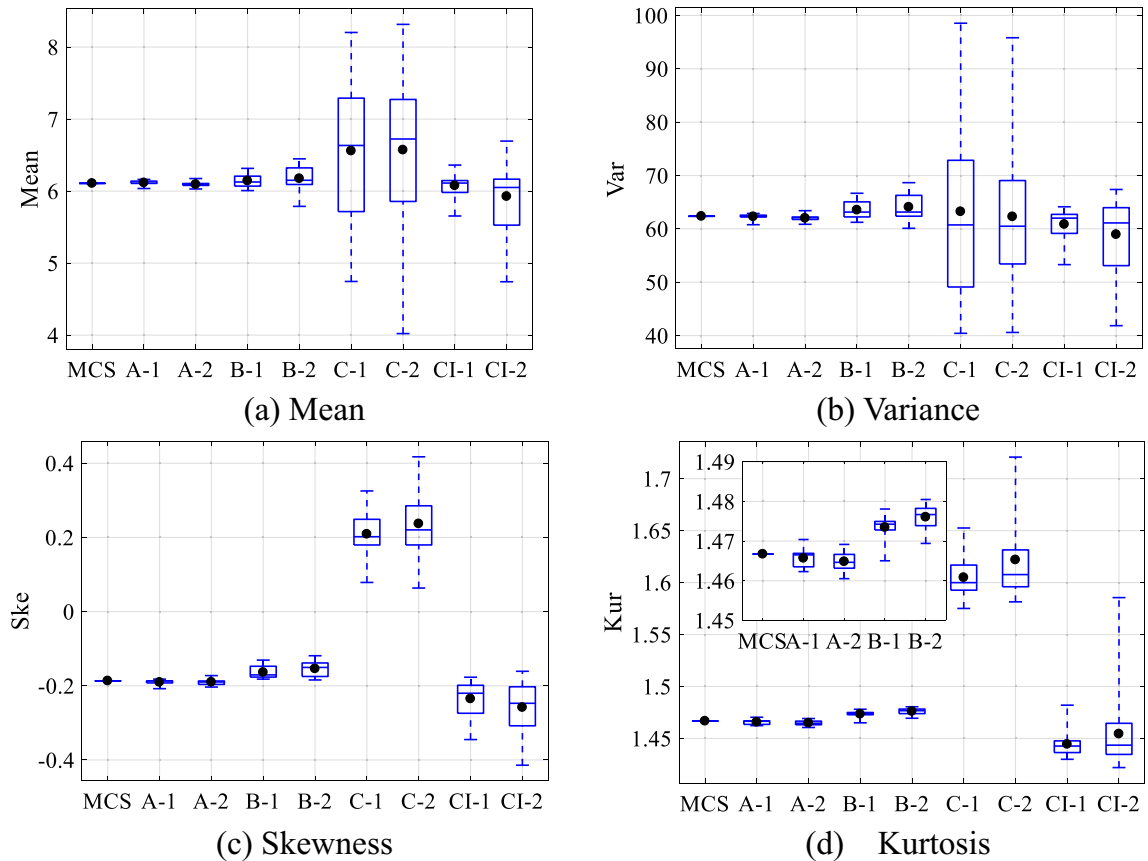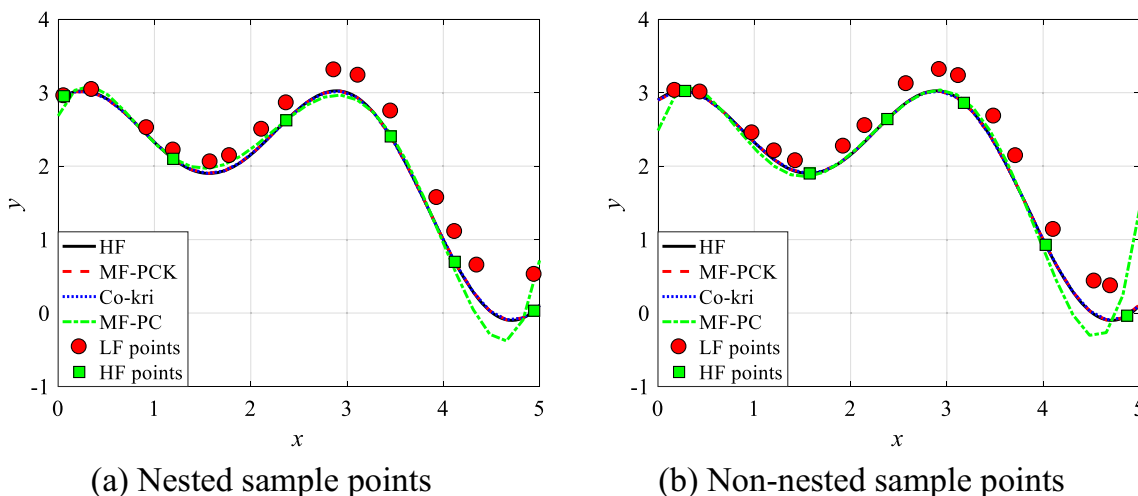
Fig. 5 Function response curves for example 2

than co-kriging. However, for example 2, such distance for co-kriging is smaller than MF-PCK (see Fig. 6). As has been stated above that this example is relatively easier to be approximated, the impact of instability existing in parameter estimation for MF-PCK that is weak in common cases becomes more evident.

Thirdly, compared to the results with non-nested sampling points, the accuracy with nested ones is generally slightly better for MF-PCK and co-kriging; while it is much better for MF-PC. The interpretation lies in two aspects. For MF-PC, the direct addition framework is employed, in which the PC model in the correction term $\delta(x)$ is directly constructed

(a) Mean   (b) Variance

(c) Skewness   (d) Kurtosis

Fig. 6 Boxplots of statistical moments for example 2
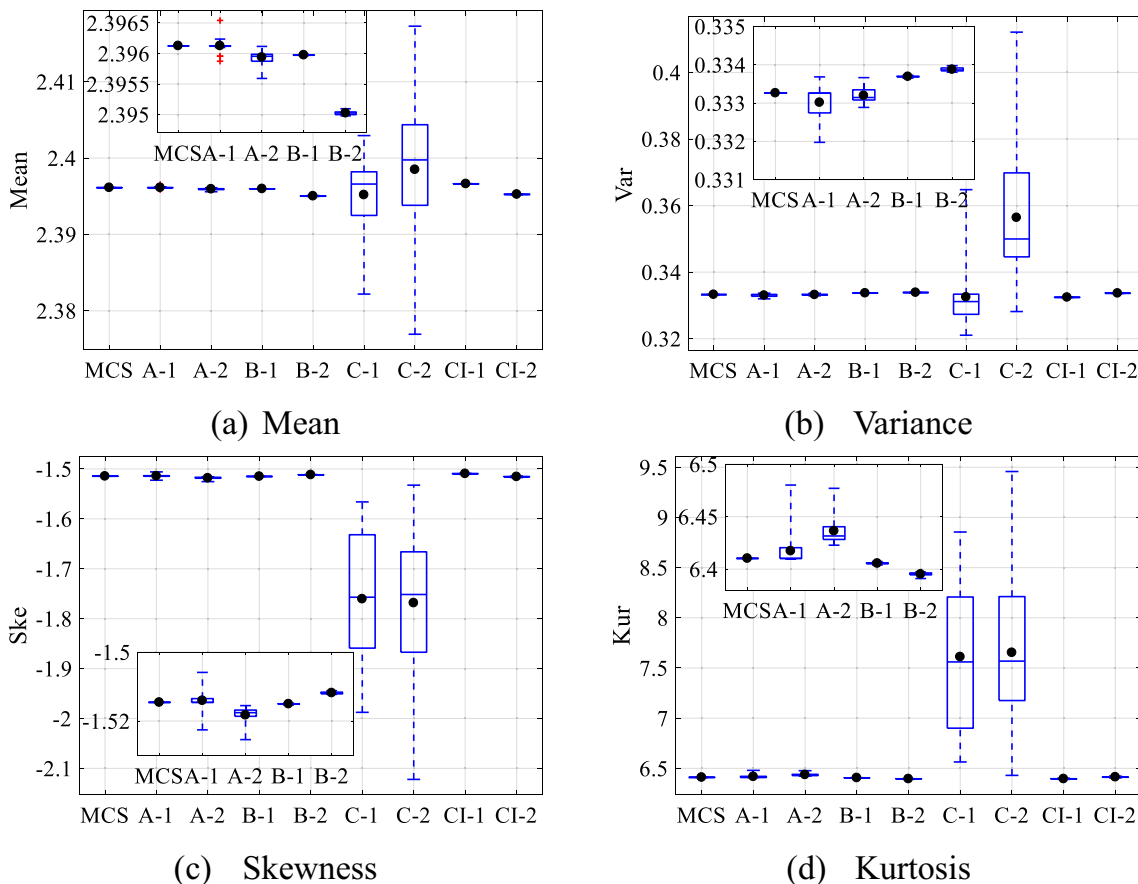
(a) Nested sample points   (b) Non-nested sample points
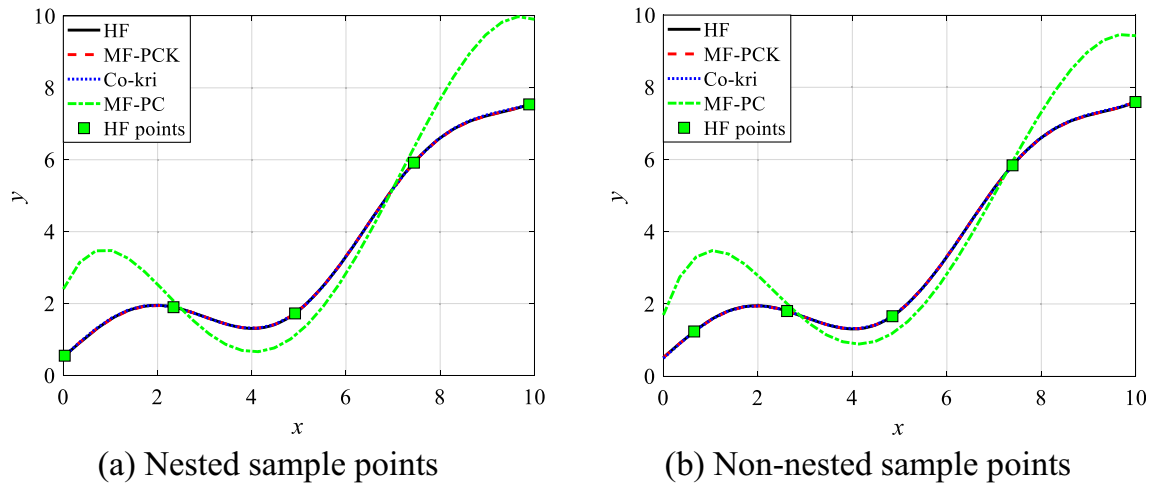
Fig. 7 Function response curves of example 3

based on the high-fidelity input sample points x^H and the corresponding response difference values $y^H(x^H) - y^L(x^H)$. For the nested case, $y^L(x^H)$ is the exact response from the low-fidelity model, while for the non-nested one, $y^L(x^H)$ is predicted by PC at x^H, which inevitably would induce error. Such error may be large when PC for the low-fidelity model is

inaccurate. However, for the other two approaches, the KOH framework is employed, in which the PC-Kriging/Kriging model for the correction term $\delta(x)$ is not directly and explicitly constructed based on the response difference values as that of MF-PC, but estimated using MLE based on all the collected response data from the whole simulation models. Therefore,

(a) Mean   (b) Variance

(c) Skewness   (d) Kurtosis

Fig. 8 Boxplots of statistical moments for example 3

(a) Nested sample points　　　　　　(b) Non-nested sample points
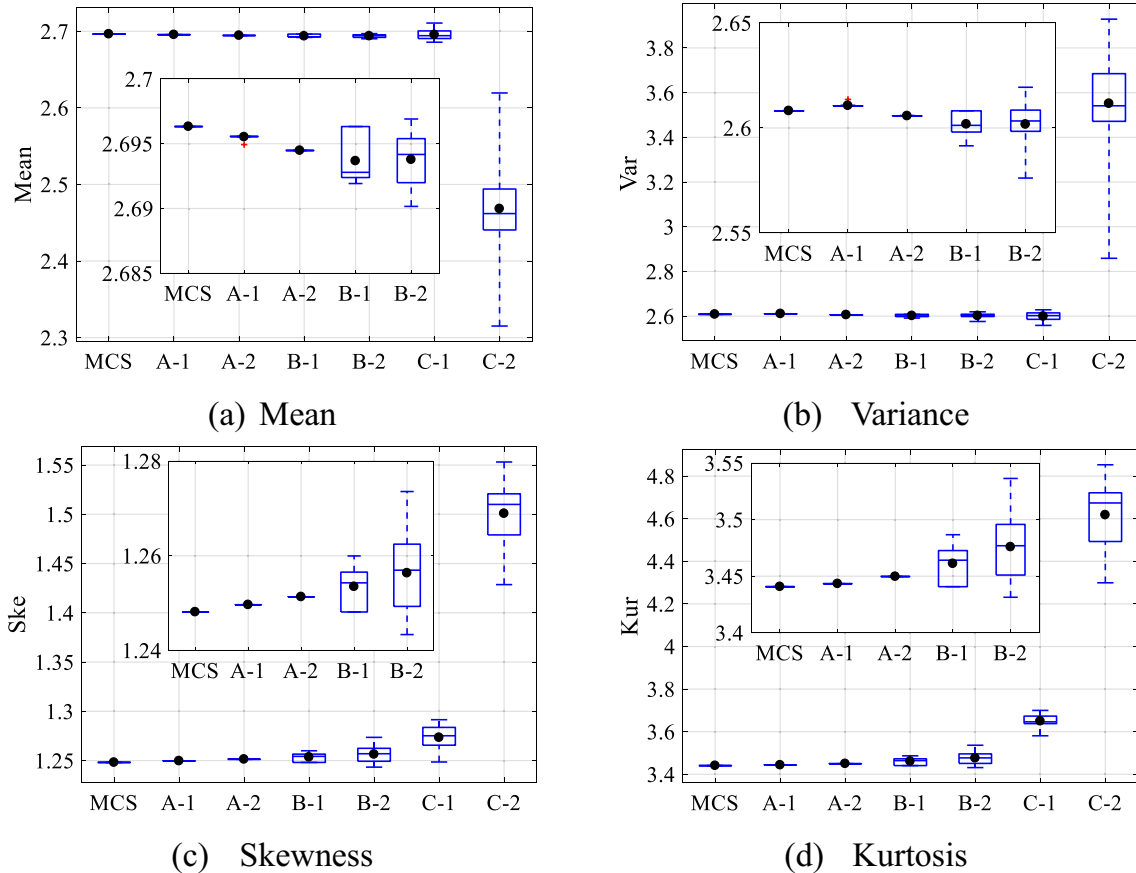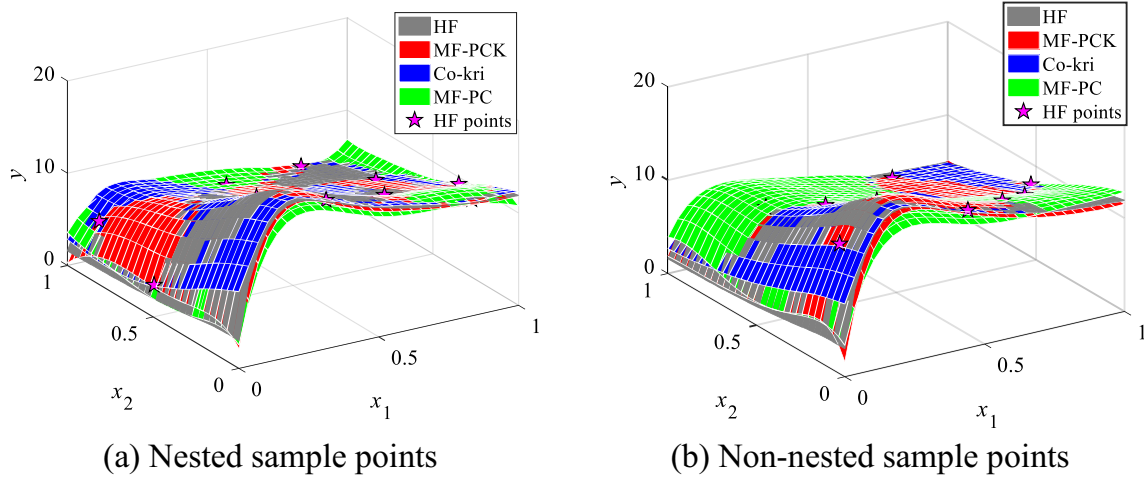
Fig. 9 Function response curves of example 4

the prediction error is avoided, which is helpful for ensuring the accuracy. In addition, the covariance among data with different fidelities is fully considered during the hyper-parameter estimation for MF-PCK and co-kriging (Liu et al.

2018), which can reduce the adverse impact of the non-nested sample points to some extent. However, for MF-PC, the PC coefficients for the LF model and additive correction terms are directly calculated by regression, based on which the PC

(a)　Mean　　　　　　　　　　　　(b)　Variance

(c)　Skewness　　　　　　　　　　(d)　Kurtosis

Fig. 10 Boxplots of statistical moments for example 4

(a) Mean

(b) Variance

(c) Skewness

(d) Kurtosis
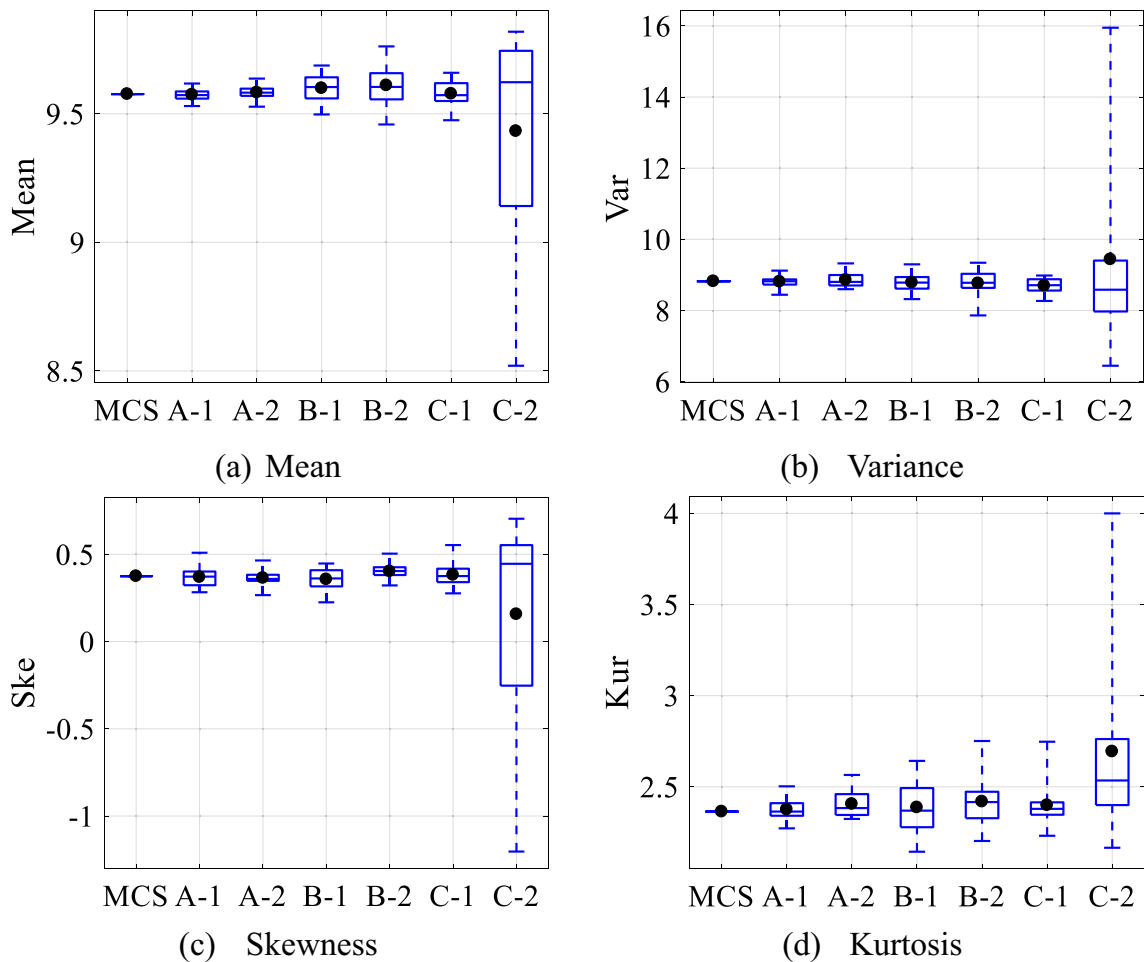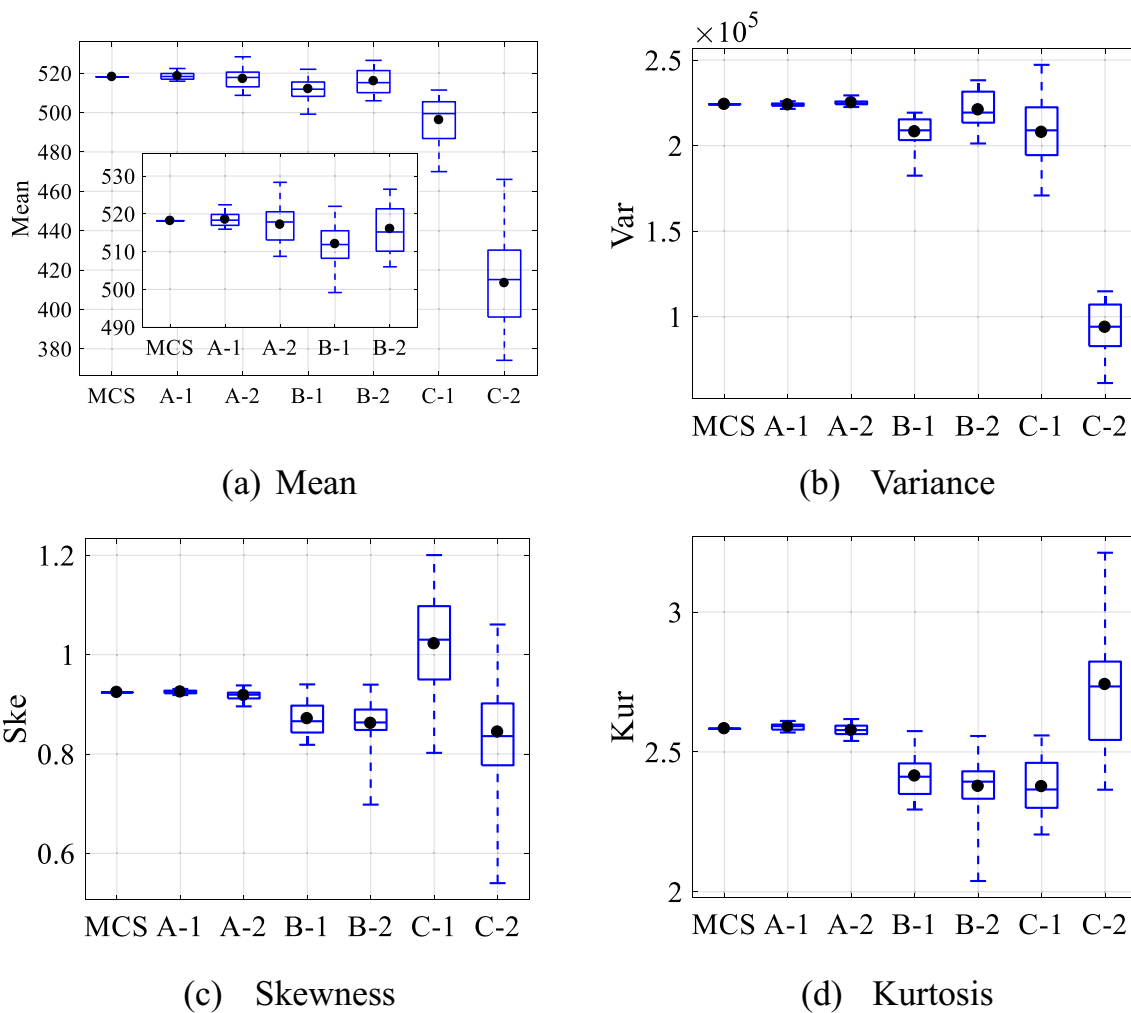
Fig. 11 Boxplots of statistical moments for example 5

model for the HF model is constructed. It is also noticed that the results of MF-PC with nested sample points are much more accurate than those with non-nested ones for examples 3 and 4. The reason is that there are three multi-fidelity models to fuse, involving two correction terms, which would amplify the impact of prediction error on UP for MF-PC.

Fourthly, it is found from Figs. 1, 5, and 7 that the response curves generated by MF-PC is so far from the HF response at the first and last HF point. The explanation is as follows. The weighted stochastic response surface method (WSRSM) (Xiong et al. 2011) is employed to construct the PC model for MF-PC, in which the Latin Hypercube sampling method is employed to generate sample points and values of the joint probability density function at the sample points are employed as the weights. It has been demonstrated that WSRSM considering sample weight is more accurate than the stochastic response surface method (SRSM) for UP. For example 1 (case 1,

Fig. 1), example 2 (Fig. 5), and example 3 (Fig. 7), the random input follows normal distribution, and thus the smallest weights are assigned to the first and last sample points as they are furthest from the mean point. Therefore, the improvement in accuracy of the stochastic response surface at the first and last sample points is relatively smaller than the points in the middle region.

As can be seen from Fig. 7, this phenomenon is more obvious for example 3. The reason is that three multi-fidelity models with random inputs of normal distribution and three weighted stochastic response surfaces (two for the correction terms and one for the lowest-fidelity model) are involved for example 3, which would amplify the impact of above prediction errors. For problem with uniform distributed random input, this phenomenon is not very obvious. For the proposed MF-PCK method, as the PC coefficients are calculated by the maximum likelihood estimation method rather than WSRSM, this phenomenon does not exist.
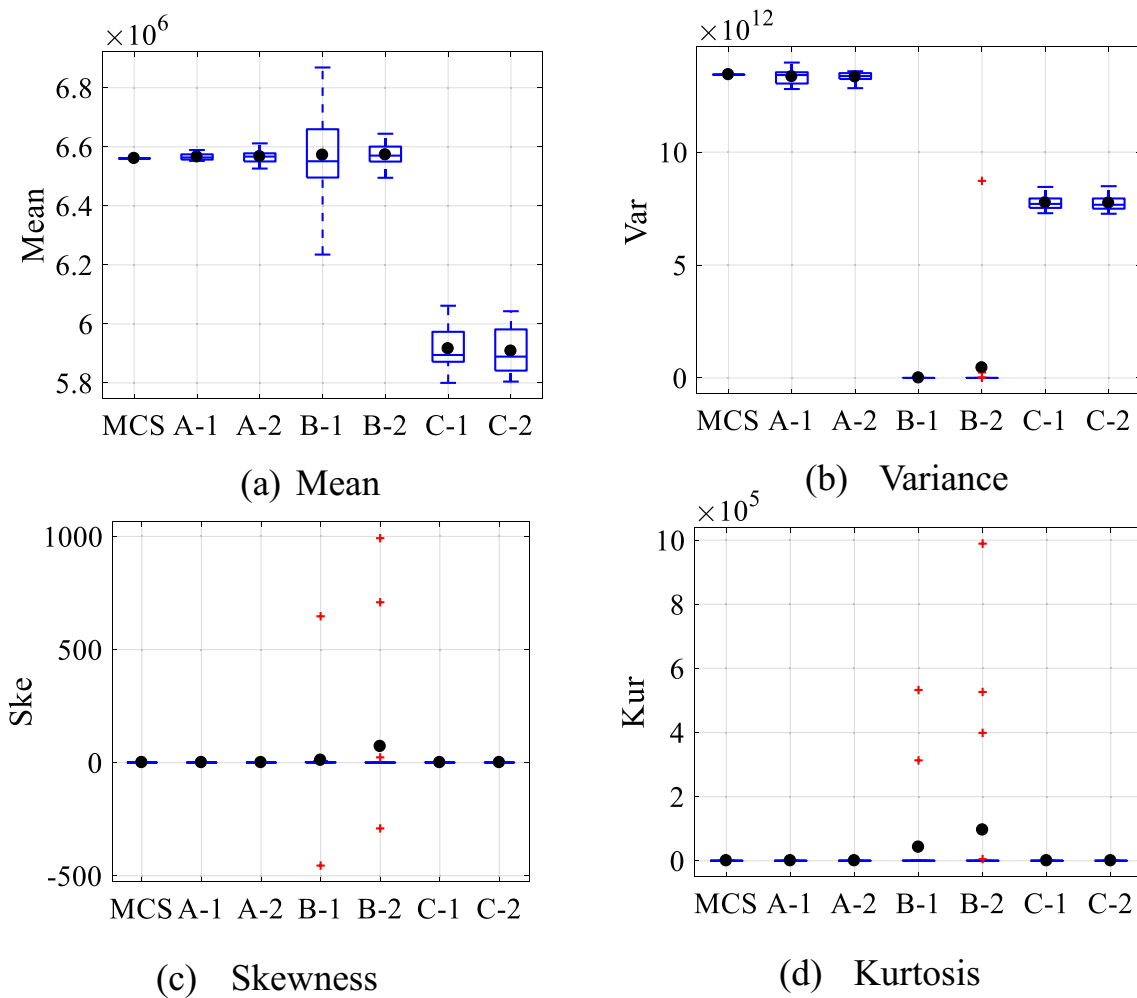
(a) Mean

(b)   Variance

(c)   Skewness

(d)   Kurtosis

Fig. 12 Boxplots of statistical moments for example 6

4.2 Test for non-hierarchical fidelity

For models with non-hierarchical fidelity, cases with nested $(D_H \subseteq D_i,\ i = 1,\ \dots, m)$ and non-nested $(D_H \not\subseteq D_i,\ i = 1,\ \dots, m)$ sample points are also tested to explore the impact of sample property on the accuracy of multi-fidelity modeling. The tested examples are displayed in Table 11. The MF-PC method can only be used in the case that the accuracy of the multi-fidelity models can be ranked (it is named as the hierarchical fidelity). However, for the case of non-hierarchical fidelity

Table 11 Tested examples with non-hierarchical fidelity

No.	Functions	Distributions
1	$y^H = e^{-0.05x^2}\cos(0.5x)$ $y^1 = e^{-0.05x^2}\cos(0.5x) - 0.5e^{-0..02(x-5)^2}$ $y^2 = e^{-0.05x^2}\cos(0.5x) + 0.5e^{-0..02(x-5)^2}$	$x{\sim}\mathcal{U}\,(-8,12)$
2	$y^H = \left(1 - e^{\frac{-1}{2x_2}}\right)\left(\dfrac{1000tf \times x_1^3 + 1900x_1^2 + 2092x_1 + 60}{1000tl \times x_1^3 + 500x_1^2 + 4x_1 + 20}\right)$ $y^1 = \left(1 - e^{\frac{-1}{2x_2}}\right)\left(\dfrac{1000tf \times x_1^3 + 1900x_1^2 + 2092x_1 + 60}{1000tl \times x_1^3 + 500x_1^2 + 4x_1 + 20}\right) + \dfrac{5e^{-tf}x_1^{th/2}}{x_2^{2+th} + 1}$ $y^2 = \left(1 - e^{\frac{-1}{2x_2}}\right)\left(\dfrac{1000tf \times x_1^3 + 1900x_1^2 + 2092x_1 + 60}{1000tl \times x_1^3 + 500x_1^2 + 4x_1 + 20}\right) - \dfrac{5e^{-tf}x_1^{th/2}}{x_2^{2+th} + 1}$ $tf = 0.2; th = 0.3; tl = 0.1;$	$x_{1,2}{\sim}\mathcal{U}\,(0,1)$

(a) Nested sample points
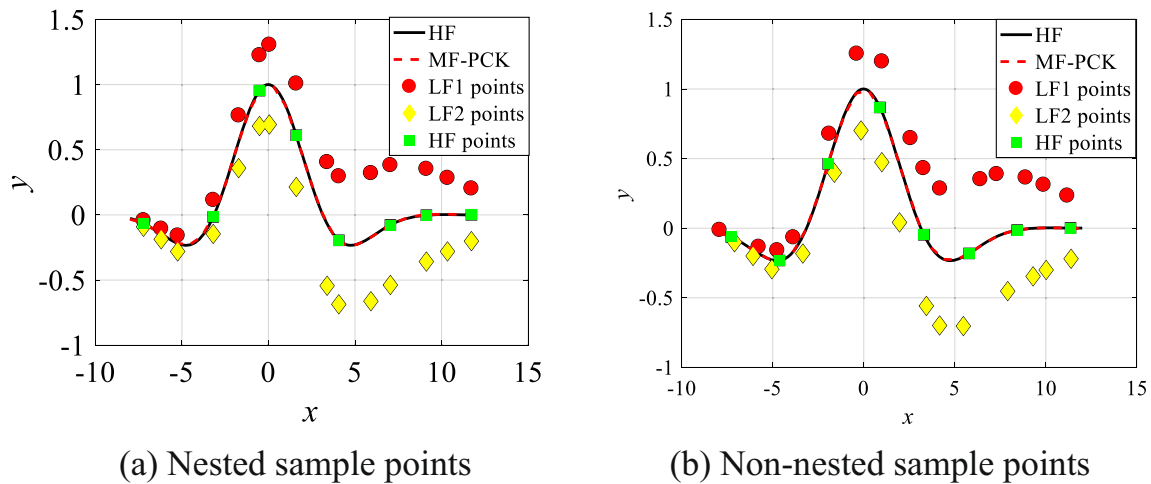
(b) Non-nested sample points

Fig. 13 Function response curves of example 1

that the accuracy of the multi-fidelity models cannot be ranked, MF-PC is no longer applicable, as it constructs the metamodel based on the fidelity level and the higher-fidelity model is represented as the sum of the adjacent lower-fidelity PC model and the PC correction model of the difference between them. Therefore, only the proposed MF-PCK method is tested.

Similar to the cases with hierarchical fidelity, 20 simulations are repeatedly done. For all the examples, the order of PC for lower-fidelity models and correction term are set as 5 and 3, respectively. For example 1, the input sample points for the three models $y^1(\mathbf{x})$, $y^2(\mathbf{x})$, and $y^3(\mathbf{x})$ are set as $n_1 = 15$, $n_2 = 15$, and $n_3 = 8$, and for example 2, $n_1 = 20$, $n_2 = 20$, and $n_3 = 10$. The existing MF-PC method is not tested here, as it cannot deal with models with non-hierarchical fidelity. Therefore, only the MF-PCK method is tested.

The response function curves, errors of the first four order statistical moments relative to those of MCS, and boxplots of the statistical moments are respectively shown in Figs. 13 and 15, Tables 12 and 13, and Figs. 14 and 16. From these results, it is observed that for the non-hierarchical case, the proposed non-hierarchical MF-PCK approach can also produce accurate enough and robust results. The approximated response function curves by MF-PCK are very close to those of the high-fidelity one. Meanwhile, the statistical moments produced by MF-PCK show great agreements to those of MCS, and exhibit high

robustness. However, the existing MF-PC method cannot work in these cases.

In this work, the sample sizes for the multi-fidelity response models $\{n_1, n_2, \ldots\}$ in each example are determined by testing the accuracy of UP corresponding to different combinations of sample sizes. The combination that yields accurate UP results for MF-PCK and the minimum amount of computational cost is selected. The other approaches (co-kriging and MF-PC) just use the same sample points to do UP for comparison. As the tested examples are simple mathematical problems, it is relatively easy to obtain a good combination of sample sizes. However, it is impossible to employ this way in practice especially for black-box-type problems. This paper tries to develop a new multi-fidelity modeling approach for UP to reduce the computational cost, and the focus lies in how to fusion models with different fidelities. How to efficiently determine the sample size for each response model is another research topic and is out of the scope of this work. In literature, many works related to this topic have been done. Some works aims to develop a sequential sampling strategy that can obtain an optimal combination of the sample sizes for a given computational budget and improve the accuracy of a multi-fidelity metamodel as far as possible (Guo et al. 2018; Zhang et al. 2018). Some works propose the scenario of resource allocation to determine such combination, so as to reduce the

Table 12 Relative errors of statistical moments for example 1	Methods Errors	A-1	A-2
	e_m (%)	2.49	3.60
	e_v (%)	0.56	1.25
	e_s (%)	1.66	1.96
	e_k (%)	1.83	1.99

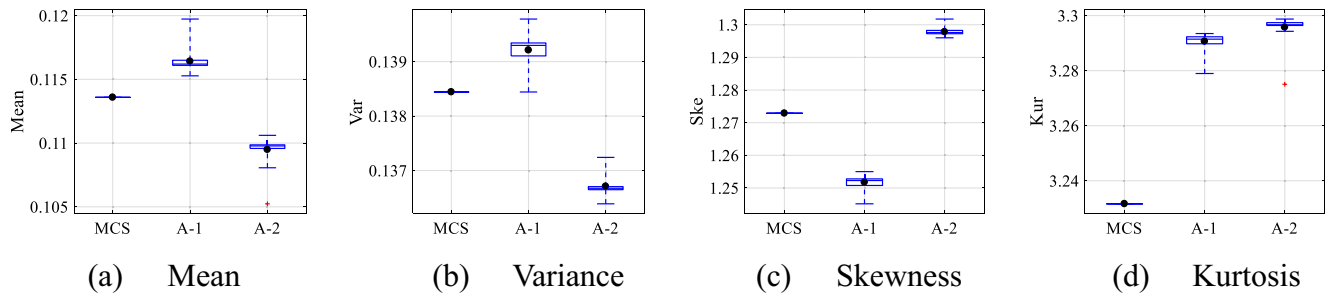| Table 13 Relative errors of statistical moments for example 2	Methods Errors	A-1	A-2
	e_m (%)	0.01	0.73
	e_v (%)	1.42	2.91
	e_s (%)	5.53	2.25
	e_k (%)	3.02	6.00

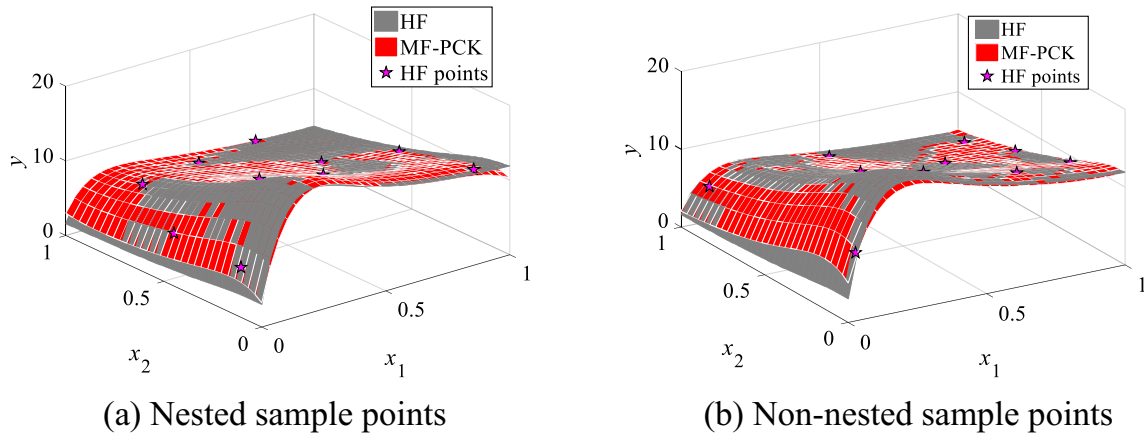Fig. 14 Boxplots of statistical moments for example 1

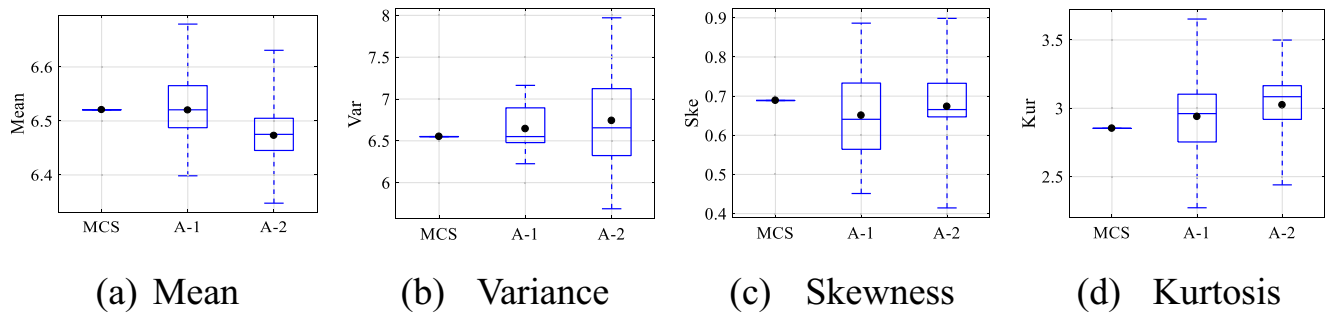Fig. 15 Function response curves of example 2

Fig. 16 Boxplots of statistical moments for example 2

epistemic uncertainty of model as far as possible (Jiang et al. 2016; Hu and Mahadevan 2018).

5 Application to airfoil optimization

The proposed MF-PCK method is applied to a benchmark aerodynamic design problem involving inviscid and viscous transonic past airfoil shapes, developed by the AIAA aerodynamic design optimization discussion group (Farin 1993). It aims at maximizing the lift-to-drag ratio of the modified NACA 0012 airfoil section at a free-stream Mach number of $M_a = 0.7$ and an angle of attack $\alpha = 3°$, subject to the thickness constraint.

In this work, the B-spline curve (Guo et al. 2018) with 10 control points is employed for the shape parameterization of the airfoil, where the horizontal locations of the 10 control points are fixed as $x = [0.1\ 0.3\ 0.5\ 0.7\ 0.9\ 0.9\ 0.7\ 0.5\ 0.3\ 0.1]$ and the vertical locations y are free to move, as shown in Fig. 17.

The determinate optimization (DO) is formulated as:

$$\max_{y} f = L(y)/D(y) \qquad \text{s.t.} \quad t_{\max}(y) \geq 0.1043 \qquad (35)$$

where y is the vector of design variables; $L(y)$ and $D(y)$ are lift and drag, respectively; and $t_{\max}$ is the airfoil maximum thickness and equals 0.1043 for the original baseline airfoil.

Robust optimization (RO) is performed on this problem considering uncertainties from the flight condition, i.e., the Mach number M_a and the attack angle α. M_a and α are assumed to follow uniform distribution with variations ± 0.1 and $\pm 1°$ around their nominal values, respectively. The robust airfoil optimization is formulated as follows:

$$\max_{y} \ F = \mu_f - k\sigma_f \qquad \text{s.t.} \quad t_{\max}(y) \geq 0.1043 \qquad (36)$$

where μ_f and σ_f are the mean and standard deviation of the lift-to-drag ratio of airfoil, and the weighting factor k is set as $k = 3$.

During the optimization, the computational fluid dynamics (CFD) flow field solver using Fluent17.0 is employed to obtain the aerodynamic data, and the turbulence model employed is the k-omega two-path turbulence model. The steady-state density solver with Roe-FDS is employed. Two multi-fidelity models are considered according to the number of nodes of grid employed during the CFD analysis. Table 14 shows the convergence condition with different grid density for the NACA0012 airfoil, from which it is found that the density of the mesh significantly impacts the simulation time cost and the accuracy of CFD analysis result. When the density of mesh is increased largely enough, the lift-drag ratio basically remains unchanged (please see the results in the last two rows of Table 14). Therefore, in this work, the CFD analysis with 150 nodes in the remote field of the grid, and 300 in the airfoil boundary is considered as the HF model, while the CFD analysis with 100 nodes in the remote field, and 200 in the airfoil boundary is employed as the LF model.

For the MF-PCK model, 5 HF sample points and 10 LF sample points are generated using the Latin hypercube sampling technique to construct the MF-PCK model for uncertainty propagation. Meanwhile, the order of the PC
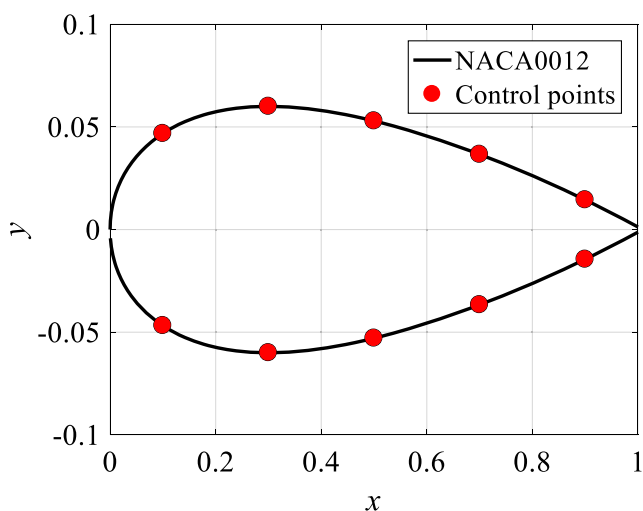
term for the LF model and correction term are set as 2 equally. To test the accuracy of the MF-PCK method for UP using these sample points, the MF-PC model based on 8 HF sample points and 10 LF sample points generated by the Latin hypercube sampling method is also constructed for comparison and the order of PC term is set as 3. The results of MCS are used as the reference values. Table 15 shows the mean and standard deviation of the lift-drag ratio produced by the proposed MF-PCK method, MF-PC and MCS at the original NACA0012 airfoil considering uncertainties from M_a and α. It is observed that the results of MF-PCK are very close to those of MCS, demonstrating its accuracy and effectiveness. Moreover, MF-PCK is basically as accurate as MF-PC, while the function calls of the HF model are clearly reduced. Therefore, during the following robust optimization, MF-PCK constructed with 5 HF and 10 LF sample points is employed for UP during each optimization iteration.

The robust optimization (RO) with MF-PCK and MF-PC, and deterministic optimization (DO) without considering any uncertainties are conducted for comparison, of which the results are shown in Figs. 18 and 19 and Table 16. For the sake of simplicity, the results of RO using the proposed MF-PCK and the existing MF-PC are denoted as RO-P and RO-E, respectively.

Figure 18 shows the airfoils obtained by the two ROs and DO, as well as that of the baseline. It is observed that the two ROs using the proposed MF-PCK UP and the MF-PC UP methods can produce very similar airfoils. The leading edges

Fig. 17 B-spine parameterization for the airfoil

Table 14 Convergence condition with different grid density for the NACA0012 airfoil

Grid size		L/D	Simulation time (s)
Remote field	Airfoil boundary		
80	160	26.7997	148.8601
100	200	30.6793	186.2760
120	240	32.8425	236.0699
150	300	33.2187	291.0449
180	360	33.2362	351.9661

Table 15 Uncertainty propagation results of the NACA0012 airfoil

| Method	Number of samples		$\mu_{L/D}$	$\sigma_{L/D}$
	HF	LF		
MF-PCK	5	10	27.5105	12.5738
MF-PC	8	10	28.1315	12.5072
MCS	100	0	27.7369	13.4137
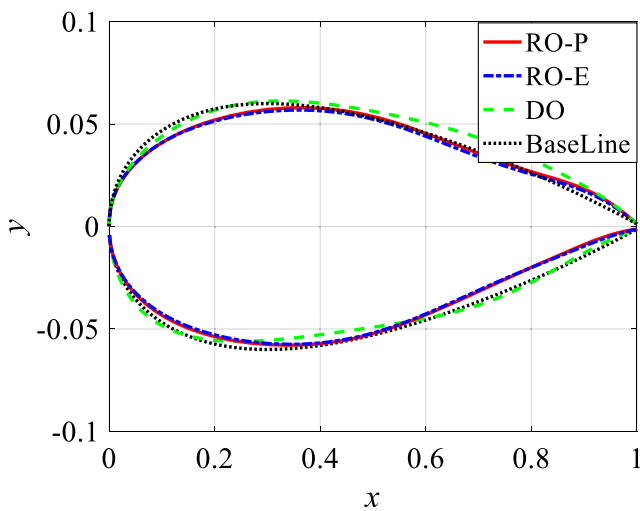
Fig. 18 Comparison of three optimized and baseline airfoils

produced by the two ROs are almost overlapped. Meanwhile, the thickness of the leading edge of optimization design is obviously reduced than the original baseline airfoil, which can improve the critical Mach number and weaken the shock wave area, thus reduce the drag coefficient of the airfoil. These

results indicate that through optimization, the drag coefficient is reduced. In addition, it is noticed that the rear trailing edge produced by DO bends downwards exhibiting evident characteristics of supercritical airfoil, which can increase the lift coefficient. Therefore, compared to the two ROs, the obtained lift-drag ratio is larger for DO, which will be verified in the following analysis.

Figure 19 illustrates the static pressure nephograms obtained by the two ROs and DO, as well as the baseline. It is observed that there is a strong shock wave area (i.e., the junction of dark blue triangle area and light blue area on its right) on the upper surface for the baseline airfoil. The intensity of shock wave is proportional to the pressure difference between both sides of this junction. The stronger of the shock, the larger of the wave drag, and the smaller of the lift-drag ratio. Clearly, it is noticed that the pressure differences between both sides of the junction of the airfoils obtained by all the optimization designs are smaller than that of the baseline, especially for the RO. Therefore, the drag coefficient is clearly reduced through optimization, and it is reduced more for RO compared to that of DO. From the results of Figs. 18 and 19,

(a) RO-P

(b) RO-E

(c) DO

(d) Baseline
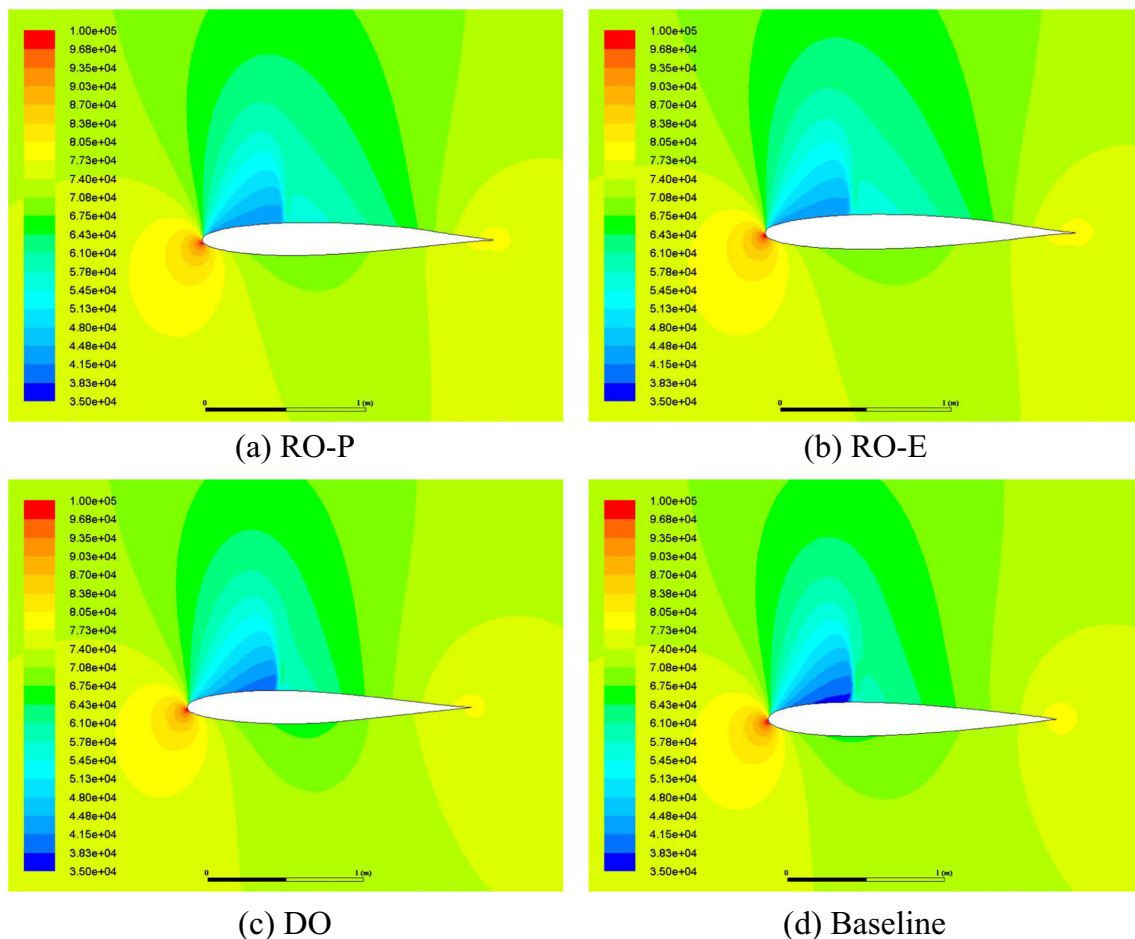
Fig. 19 Comparison of static pressure clouds

Table 16 Optimal results different methods

	μ_f	σ_f	$t_{\max}$	Number of function call		Total time (h)
				HF	LF	
RO-P	31.2091	10.0673	0.1043	2540	5080	415.29
RO-E	31.0689	10.1629	0.1043	3696	4620	489.72
DO	38.0623	13.9853	0.1046	470	0	38.00
Baseline	27.8256	13.0642	0.1043	/	/	/

it is concluded that the increase in the lift-drag ratio for DO is mainly caused by the increase in lift, while it is mainly caused by the decrease in drag for RO.

The optimal results and computational cost of the optimized airfoils by different methods are shown in Table 16, from which it is found that after optimization, the lift-drag ratio can be increased compared to that of the baseline airfoil. Compared to the results of DO, RO can clearly improve the robustness of design (smaller σ_f) that is less sensitive to uncertainties, while at a sacrifice of the performance (smaller μ_f). RO with the proposed MF-PCK method can produce very close results to those from RO with the existing MF-PC approach, while it clearly reduces the computational time (415.29 vs. 489.72). These results show great agreement to what have been observed above, and demonstrate the effectiveness and advantages of the proposed multi-fidelity UP method.

6 Conclusions

In this paper, a multi-fidelity PC approach using the Gaussian process modeling theory is developed to make the PC method more efficient and applicable to practical problems. With the proposed approach, the classic multi-level co-kriging modeling framework is extended from the deterministic domain to the stochastic one for UP, which can deal with analysis models with both hierarchical and non-hierarchical fidelities. Through comparative studies on several numerical examples for UP with the same computational cost, it is noticed that compared to the commonly used addition correction based multi-fidelity PC method, the proposed approach can consistently reduce the error to at least 5%. Compared to co-kriging, it generally can be reduced to about 50 to 12%, and it can be reduced to 10 to 0.1% for problems with unsymmetric distributed random input or large variation. Meanwhile, compared to both existing methods, the proposed approach can evidently enhance the robustness of UP. These results demonstrate the effectiveness and advantage of the proposed multi-fidelity PCK method. The application of the proposed multi-fidelity PCK approach to an

engineering robust aerodynamic optimization problem further verifies its effectiveness and applicability.

7 Replication of results

The results shown in the manuscript can be re-produced. Considering the size limit of the uploaded supplementary material, the codes for one of the mathematical example (Example 1) is uploaded as supplementary material. For the rest of the examples, it is very easy to implement by changing the response functions and sample points based on the codes provided to obtain the results shown in the manuscript.

Compliance with ethical standards

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

References

Allaire D, Willcox K. Fusing information from multifidelity computer models of physical systems. International Conference on Information Fusion. IEEE Jul, 9-12, 2012, Singapore, Singapore

Blatman G, Sudret B (2011) Adaptive sparse polynomial chaos expansion based on least angle regression. J Comput Phys 230(6):2345–2367

Chen Z, Peng S, Li X et al (2015) An important boundary sampling method for reliability-based design optimization using kriging model. Struct Multidiscip Optim 52(1):55–70

Chen S, Jiang Z, Yang S et al (2016) Nonhierarchical multi-model fusion using spatial random processes. Int J Numer Methods Eng 106(7): 503–526

Du X, Chen W (2002) Efficient uncertainty analysis methods for multi-disciplinary robust design. AIAA J 40(3):545–552

Eldred M. Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design. 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference May. 4-7, 2009, Palm Springs, California, USA

Farin G (1993) Curves and surfaces for computer aided geometric design. Academic Press, Boston

Fernández-Godino MG, Park C et al (2016) Review of multi-fidelity models. arXiv preprint arXiv:1609.07196

Gratiet LL, Cannamela C (2012) Kriging-based sequential design strategies using fast cross-validation techniques with extensions to multi-fidelity computer codes. Technometrics 57(3):418–427

Gratiet LL, Cannamela C, Iooss B (2014) A Bayesian approach for global sensitivity analysis of (multi-fidelity) computer codes. SIAM/ASA J Uncertain Quantif 2(1):336–363

Guo Z, Song L, Park C et al (2018) Analysis of dataset selection for multi-fidelity surrogates for a turbine problem. Struct Multidiscip Optim 57(6):2127–2142

Han Z, Zimmerman R, Görtz S (2012) Alternative cokriging method for variable-fidelity surrogate modeling. AIAA J 50(5):1205–1210

Hu Z, Mahadevan S (2018) Adaptive surrogate modeling for time-dependent multidisciplinary reliability analysis. J Mech Des 140(2):021401

Huang D, Allen TT, Notz WI et al (2006) Sequential kriging optimization using multiple-fidelity evaluations. Struct Multidiscip Optim 32(5):369–382

Huang L, Gao Z et al (2013) Research on multi-fidelity aerodynamic optimization methods. Chin J Aeronaut 26(2):279–286

Jiang Z, Chen S, Apley DW et al (2016) Reduction of epistemic uncertainty in simulation-based multidisciplinary design. J Mech Des 138(8):081403

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

Kersaudy P, Sudret B, Varsier N et al (2015) A new surrogate modeling technique combining kriging and polynomial chaos expansions-application to uncertainty analysis in computational dosimetry. J Comput Phys 286:103–117

Laurenceau J, Sagaut P (2008) Building efficient response surfaces of aerodynamic functions with kriging and cokriging. AIAA J 46(2):498–507

Lee SH, Chen W (2009) A comparative study of uncertainty propagation methods for black-box-type problems. Struct Multidiscip Optim 37(3):239–253

Li C, Wang P, Dong H (2018) Kriging-based multi-fidelity optimization via information fusion with uncertainty. J Mech Sci Technol 32(1):245–259

Liu Y, Chen S, Wang F et al (2018) Sequential optimization using multi-level cokriging and extended expected improvement criterion. Struct Multidiscip Optim 58(3):1155–1173

Matteo B (2016) Multi-fidelity surrogate modelling with polynomial chaos expansions. M.S. dissertation, Federal institute of technology in Zurich

Ng WT, Eldred M. Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation. AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, AIAA/ASME/AHS Adaptive Structures Conference April. 23-26, 2012, Honolulu, Hawaii, USA

Palar PS, Tsuchiya T, Parks G. Decomposition-based evolutionary aero-dynamic robust optimization with multi-fidelity point collocation non-intrusive polynomial chaos. AIAA Non-Deterministic Approaches Conference Jan. 5-9, 2015, Kissimmee, Florida, USA

Palar PS, Zuhal LR, Shimoyama K et al (2018) Global sensitivity analysis via multi-fidelity polynomial chaos expansion. Reliab Eng Syst Saf 170:175–190

Park C, Haftka RT, Kim NH (2016) Remarks on multi-fidelity surrogates. Struct Multidiscip Optim 55(3):1029–1050

Pramudita SP, Takeshi T, Geoffrey TP (2016) Multi-fidelity non-intrusive polynomial chaos based on regression. Comput Methods Appl Mech Eng 307:489–490

Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. The MIT Press, Cambridge

Ren J, Thelen AS, Amrit A et al. Application of multifidelity optimization techniques to benchmark aerodynamic design problems. 54th AIAA Aerospace Sciences Meeting, 4–8 January 2016, San Diego, California, USA

Santiago Padron A, Juan JA, Francisco P et al. Multi-fidelity uncertainty quantification: application to a vertical axis wind turbine under an extreme gust. 15th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference June. 16-20, 2014, Atlanta, GA

Schobi R, Sudret B, Wiart J (2015) Polynomial-chaos-based kriging. Int J Uncertain Quantif 5(2):171–193

Shah H, Hosder S, Koziel S et al. Multi-fidelity robust aerodynamic design optimization under mixed uncertainty. AIAA SciTech Forum, 17th AIAA Non-Deterministic Approaches Conference January, 5–9, 2015, Kissimmee, Florida, USA

Toal DJJ, Keane AJ (2015) Efficient multipoint aerodynamic design optimization via cokriging. J Aircr 48(5):1685–1695

Toal DJJ, Bressloff NW, Keane AJ et al (2011) The development of a hybridized particle swarm for kriging hyperparameter tuning. Eng Optim 43(6):675–699

Wang F, Xiong F, Yang S, Xiong Y. A sparse data-driven polynomial chaos expansion method for uncertainty propagation. ASME. International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 42nd Design Automation Conference, August 21–24, 2016, Charlotte, North Carolina, USA

West TK IV, Gumbert C. Multidisciplinary design under uncertainty with non-intrusive polynomial chaos. 2017 AIAA SciTech Forum Jan. 9–13, 2017, Grapevine, Texas, USA

Xiong FF, Chen W, Xiong Y, Yang SX (2011) Weighted stochastic response surface method considering sample weights. Struct Multidiscip Optim 43(6):837–849

Zhang Y, Han Z, Zhang K (2018) Variable-fidelity expected improvement method for efficient global optimization of expensive functions. Struct Multidiscip Optim 58(4):1431–1451

Zhu X, Narayan A, Xiu D (2014) Computational aspects of stochastic collocation with multifidelity models. SIAM/ASA J Uncertain Quantif 2:444–463

Zhu X, Linebarger EM, Xiu D (2017) Multi-fidelity stochastic collocation method for computation of statistical moments. J Comput Phys 341:386–396