



Determination of sample size for input variables in RBDO through bi-objective confidence-based design optimization under input model uncertainty

Yongsu Jung¹ · Hyunkyoo Cho² · Zunyi Duan¹ · Ikjin Lee¹ 

Received: 18 March 2019 / Revised: 5 June 2019 / Accepted: 8 July 2019 / Published online: 27 July 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Reliability analysis accounting for only randomness of input variables often shows a significant error due to the lack of knowledge and insufficient data in the real world. Confidence of reliability indicates that the reliability has uncertainty caused by epistemic uncertainties such as input statistical model and simulation model uncertainties, and these uncertainties can be reduced and manipulated by additional knowledge. In this paper, uncertainty of input statistical models prevailing due to limited resources in practical applications is mainly treated in the context of confidence-based design optimization (CBDO). The purpose of this research is to estimate the optimal sample size for input variables in reliability-based design optimization (RBDO) to reduce the overall cost incorporating development cost for collecting samples. There are two ways to increase the confidence of reliability to be satisfied: (1) shifting the design vector toward feasible domain and (2) supplementing more input data. Thus, it is essential to find a balanced optimum to minimize the overall cost accounting for trade-off between two operations, optimally distributing the resources to the operating cost of the design vector and the development cost of acquiring new data. In this study, two types of costs are integrated into a bi-objective function, and the probabilistic constraints for the confidence of reliability need to be satisfied. Since the sample size for input variables is also included in the design variables to be optimized, stochastic sensitivity analysis of confidence with respect to the sample size of input data is developed. As a result, the proposed bi-objective CBDO can provide the optimal sample size for input data estimated from the initial data. Then, the designers are able to decide the number of tests for collecting input samples according to the optimum of the bi-objective CBDO. In addition, the different optimal sample size for each input variable can be an indicator of which random variables are relatively important to satisfy the confidence of reliability. Numerical examples and an engineering application for the multi-scale composite frame are used to demonstrate the effectiveness of the developed bi-objective CBDO.

Keywords Reliability-based design optimization (RBDO) · Epistemic uncertainty · Confidence-based design optimization (CBDO) · Input model uncertainty

1 Introduction

Simulation-based design optimization considering various uncertainties has been widely developed pursuing reliable optimum for complex engineering systems. Representatively, two

sources of uncertainties have been usually quantified and managed in design optimization to prevent unpredictable failure of a system: aleatory uncertainty represents inherent and natural randomness that cannot be manipulated, and epistemic uncertainty is incurred by a lack of knowledge that can be reduced by additional data or more advanced theory (Der Kiureghian and Ditlevsen 2009). Both uncertainties have to be quantified and aggregated to guarantee the reliability of performances in a system. In other words, not only natural randomness of variables and parameters but also imperfect knowledge and lack of data have been taken into account through effective uncertainty quantification and propagation in recent developments.

Responsible Editor: Erdem Acar

✉ Ikjin Lee
ikjin.lee@kaist.ac.kr

¹ Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea

² Department of Mechanical Engineering, Mokpo National University, Muan-gun 58554, South Korea

Various methodologies accounting for epistemic uncertainty in reliability-based design optimization (RBDO) have been prominent in determining safe design. Conventional RBDOs rely on complete statistical models of random input variables and simulation models assumed to be accurate. Thus, researches concentrating on how to compute multi-dimensional integration (i.e., estimation of reliability) have been proposed such as most probable point (MPP)-based methods (Tu et al. 1999; Lee et al. 2012; Meng et al. 2015; Kang et al. 2017; Jung et al. 2019a) and sampling-based methods (Lee et al. 2011; Dubourg and Sudret 2014; Kang et al. 2019; Moustapha and Sudret 2019). These methods mainly treat the aleatory uncertainty of random variables; thus, uncertainty of a performance function propagated from the randomness of input variables is widely investigated.

However, simulation model results may not be consistent with experimental results due to missing underlying physics and unknown model parameters, and also statistical representations for input variables can be inaccurate due to insufficient samples. As a result, epistemic uncertainty in simulation-based design optimization can generally be categorized into two types: (1) simulation model uncertainty and (2) input model uncertainty. The simulation model uncertainty includes model parameter uncertainty, model bias, and random error and can be calibrated with experiment results using either Bayesian approaches or Gaussian process (Noh et al. 2011; Jiang et al. 2013; Pan et al. 2016; Moon et al. 2017, 2018; Xi 2019). On the other hand, the uncertainty of input statistical models is also critical since estimating true input statistical information is difficult due to cost limitation for testing samples to quantify the variability of random variables. Extensive researches associated with the distribution of reliability have been developed to handle the input model uncertainty representing the uncertainty of input distribution parameters and types (Gunawan and Papalambros 2006; Youn and Wang 2008; Picheny et al. 2010; Sankararaman and Mahadevan 2011; Lee et al. 2013; McDonald et al. 2013; Yoo and Lee 2014; Nannapaneni and Mahadevan 2016; Peng et al. 2017; Hao et al. 2017; Ito et al. 2018; Moon et al. 2019). Recently, conservative RBDO (CRBDO) using the Bayes' theorem has been proposed (Gelman et al. 2013; Cho et al. 2016), and then confidence-based design optimization (CBDO) to eliminate the double-loop MCS in CRBDO is developed (Jung et al. 2019b).

One of the notable features of the epistemic uncertainty in comparison with the aleatory uncertainty is that it can be reduced by gathering more experimental and computational data and acquiring more knowledge from experts. Thus, acquiring new data should be conducted that leads to the maximum impact on reducing the epistemic uncertainty using the minimum resource under the given

budget. In other words, resource allocation for each test is necessary to enable practical cost optimization considering epistemic and aleatory uncertainties. A few works to develop test resource allocation in the context of model uncertainty have been proposed. It can provide an optimal scheme to distribute resources to different simulation models especially in multidisciplinary design optimization (MDO) (Sankararaman et al. 2013). On the other hand, there has been research on the trade-off between the number of samplings for MCS and system objective since insufficient samples for numerical integration can lead to inaccurate reliability estimation (Bae et al. 2018, 2019). In spite of the aforementioned developments, research on the trade-off between the sample size for input variables and system cost considering confidence of reliability has been quite limited.

In this paper, costs for both system objective and collecting additional data are simultaneously taken into consideration to allocate the optimal number of tests. It is noted that the optimum of CBDO is more conservative than RBDO where input statistical models are assumed to be known. It means that insufficient input data cause loss of the objective function in CBDO. However, acquiring new samples for input variables also demands engineering resources. That is, there is a trade-off between shifting design vector corresponding to the system objective and reducing epistemic uncertainty of input statistical models corresponding to the cost for additional samples. Following the previous studies (Bae et al. 2018), the system objective is defined as an operating cost, and the cost for collecting new samples for input variables is defined as a development cost. Acquiring new data may decrease the operating cost by reducing the epistemic uncertainty, but the development cost increases. Hence, the objective function in CBDO should include the development cost as well to find a practically balanced optimum in real engineering applications. In this paper, the proposed approach is called a bi-objective CBDO to determine the optimal sample size for input variables by balancing the development and operating cost. Note that the operating cost is affected only by random design variables, but the development cost is involved in both random design variables and random parameters. Consequently, the following question can be answered from the proposed bi-objective CBDO: *How to determine sample size for input variables in RBDO to minimize the overall cost?*

The remainder of the paper is organized as follows: brief reviews for CBDO including quantification of input distribution parameters and estimation of confidence in parameter space are presented in Sect. 2. Then, the proposed bi-objective CBDO explained with uncertainty on input distribution type, formulation, and sensitivity analysis of confidence with respect to the sample size for input variables is developed in Sect. 3. In Sect. 4, the proposed bi-objective CBDO is applied to numerical examples under various configurations

to verify its feasibility and effectiveness. In addition, the bi-objective CBDO associated with a 4-beam multi-scale composite frame is given in Sect. 5. Finally, conclusions are presented in Sect. 6.

2 Review of confidence-based design optimization under insufficient input data

In this section, brief reviews for the confidence of reliability and CBDO using a reliability measure approach (RMA) under insufficient input samples to quantify the statistical models of random variables will be given (Jung et al. 2019b). The confidence is defined as a probability that reliability is larger than user-specified target reliability under epistemic uncertainty, and the reliability is defined as the probability that a limit-state function is less than zero to prevent failure of the system considering only aleatory uncertainty (e.g., randomness of input variables). CBDO quantifies the confidence of reliability induced by uncertainties of input statistical models such as input distribution types and parameters.

2.1 Uncertainty of input distribution model

Uncertain input statistical models have been taken into consideration using a Bayesian approach with uncertain input distribution parameters and types. However, in order to eliminate the uncertainty of input distribution types avoiding the discrete random variables, this research assumed that input distribution types are decided deterministically to concentrate on the proposed bi-objective optimization. Thus, only uncertainties of input distribution parameters affect the uncertainty of output reliability. In recent study on CBDO, when input distribution types are unknown, kernel density estimation (KDE) is exploited to construct the input distribution with varying parameters (Jung et al. 2019b). It will be briefly discussed in Sect. 3.1 for determination of input distribution type when only sparse data are available.

The variance and mean for input variables are quantified as inverse-gamma and normal distribution, respectively, based on the normality assumption as Cho et al. (2016)

$$\sigma_i^2 | \mathbf{x}_i \sim \text{IG} \left(\frac{ND_i - 1}{2}, \frac{(ND_i - 1)s_i^2}{2} \right) \tag{1}$$

and

$$\mu_i | \sigma_i^2, \mathbf{x}_i \sim N \left(\bar{x}_i, \sqrt{\frac{\sigma_i^2}{ND}} \right) \tag{2}$$

where \mathbf{x}_i , ND_i , s_i^2 , and \bar{x}_i are the given input dataset, the sample size, the sample variance, and the sample mean of the i th random variable, respectively. The normal distribution

of μ_i in (2) is conditional to realization of the variance from (1). As a result, variability of input distribution parameters is affected by the sample mean, variance, and the sample size. Note that the sample mean of random design variables is equivalent to the initial design point (i.e., mean vector) in CBDO to evaluate the cost function, and as the design moves, all the samples are shifted by the same amount meaning that the sample variance is fixed during the design optimization (Cho et al. 2016). However, the sample mean of random parameters is invariant and directly used in the optimization process.

2.2 Confidence of reliability

The confidence of reliability is computed using a cumulative distribution function (CDF) of reliability as a function of input distribution parameters when input distribution types are known. The confidence to satisfy the given reliability denoted as Re is expressed as Cho et al. (2016); Moon et al. (2018)

$$CL(Re | \mathbf{x}) \equiv 1 - F_{Re}(Re | \mathbf{x}) \tag{3}$$

$$= 1 - \int_0^{Re} \int_{\Omega_{\psi}} f(\phi, \psi | \mathbf{x}) d\psi d\phi$$

where ϕ is a variable corresponding to the reliability; ψ denotes a vector of input distribution parameters. The joint probability density function (PDF) in (3) is described using the Bayes' theorem as

$$f(\phi, \psi | \mathbf{x}) = f(\phi | \psi, \mathbf{x}) f(\psi | \mathbf{x}) \tag{4}$$

Therefore, the confidence of reliability can be estimated through uncertainty quantification and propagation from the uncertainty of input statistical models quantified by the given input dataset.

2.3 Reliability measure approach

Recently, reliability measure approach (RMA) of CBDO has been proposed to alleviate computational demands required to compute the confidence of reliability (Jung et al. 2019b). The key concept of RMA is identical to performance measure approach (PMA) of RBDO that the multi-dimensional integration is approximated on MPP. In RMA, MPP-search is performed in P-space where the input distribution parameters are transformed to the standard normal space through the Rosenblatt transformation (Rosenblatt 1952). The MPP of RMA is the point with the smallest reliability on the given hypersphere in P-space. The RMA formulation is given by

$$\begin{aligned}
 & \underset{\mathbf{p}}{\text{minimize}} \quad \text{Re}(G(\mathbf{X})|\boldsymbol{\mu}, \boldsymbol{\sigma}^2) & (5) \\
 & \text{subject to } \|\mathbf{p}\| = \Phi^{-1}(\text{CL}_{\text{Target}}) \\
 & \text{where } \quad \text{Re}(G(\mathbf{X})|\boldsymbol{\mu}, \boldsymbol{\sigma}^2) \cong \frac{1}{N_{MCS}} \sum_{j=1}^{N_{MCS}} I_{\Omega_R}(G(\mathbf{x}^{(j)})) \\
 & \quad I_{\Omega_R}(G(\mathbf{x})) = \begin{cases} 1, & G(\mathbf{x}) \leq 0 \\ 0, & \text{otherwise} \end{cases} \\
 & \quad \Phi(p_{2i-1}) = F_{\sigma_i^2}(\sigma_i^2 | * \mathbf{x}_i) \\
 & \quad \Phi(p_{2i}) = F_{\mu_i}(\mu_i | \sigma_i^2, * \mathbf{x}_i) \quad \text{for } i = 1, 2, \dots, N
 \end{aligned}$$

In (5), $\text{Re}(G(\mathbf{X})|\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$ is the reliability for the given input distribution parameters, and $\text{CL}_{\text{Target}}$ is the target confidence to be satisfied in CBDO. The indicator function for a reliable domain, $I_{\Omega_R}(G(\mathbf{x}^{(j)}))$ is to judge whether the j th sampling point is in the reliable region or not. N_{MCS} means the number of samples for MCS, and N is the number of input random variables including both design variables and parameters. The random variables are assumed to be independent in this study, but it is easily able to be expanded to correlated random variables. It can be seen that two input parameter distributions for the i th random variable denoted as $F_{\mu_i}(\mu_i | \sigma_i^2, * \mathbf{x}_i)$ and $F_{\sigma_i^2}(\sigma_i^2 | * \mathbf{x}_i)$ are transformed to the standard normal distribution (Rosenblatt 1952). Thus, the number of random variables in P-space becomes $2N$ since a two-parameter distribution is assumed for the input random variables. Details regarding (5) and the concept of RMA are found in the literature (Jung et al. 2019b). In addition, sensitivity analysis for reliability in P-space has been developed using the first-order score function. Therefore, the optimization in (5) to find the input parameter realizations having the minimum reliability in P-space can be performed employing any gradient-based MPP-search algorithm.

2.4 Confidence-based design optimization

Instead of directly estimating the confidence of reliability at each design point, RMA judges whether the probabilistic constraint is satisfied or violated following the concept of PMA. Therefore, RMA facilitates efficient CBDO by checking the only reliability at MPP obtained from MCS. Therefore, CBDO is formulated as

$$\begin{aligned}
 & \text{given } * \mathbf{x} \\
 & \underset{\mathbf{d}}{\text{minimize}} \quad \text{cost}(\mathbf{d}) \\
 & \text{subject to } \text{Re}(G_j(\mathbf{X}|\boldsymbol{\psi}_j^*)) - \text{Re}_j^{\text{Target}} \geq 0 \quad \text{for } j = 1, 2, \dots, nc & (6)
 \end{aligned}$$

where $\boldsymbol{\psi}_j^*$ is the MPP of the j th constraint in P-space obtained from (5); $\text{Re}_j^{\text{Target}}$ is the user-specified target confidence for the j th constraint; \mathbf{d} is the mean vector of random design variables among random vector \mathbf{X} ; nc is the number of constraints. The

sensitivity analysis with respect to design point (i.e., the mean vector of random design variables) is obtained from the gradient at the MPP. To avoid repetitive reliability computations for numerical sensitivity analysis, a gradient vector of reliability at MPP using the first-order score function is required whose derivation is given in the literature (Lee et al. 2010; Jung et al. 2019b).

3 Bi-objective confidence-based design optimization

Epistemic uncertainty caused by insufficient input data yields uncertainty of performance reliability. Different reliabilities may be obtained under different realizations of uncertain input models. Since CBDO takes uncertainty of input distribution parameters into account, a more conservative optimum would be obtained compared with an RBDO optimum, which leads to a loss in the objective function of the system defined as an operating cost. That is, the operating cost at the RBDO optimum without input model uncertainty is always less than that of the CBDO optimum for compensating the input model uncertainty. Therefore, reduction of the epistemic uncertainty by increasing the sample size for input variables saves the operating cost of the system, while it increases the development cost. If so, *how to determine sample size for input variables in RBDO to minimize the overall cost?*

To answer the question above, we present a new bi-objective CBDO accounting for both the operating cost and the development cost simultaneously during the design optimization. There are two ways to satisfy confidence constraints in CBDO: (1) shifting the design vector toward the feasible region, which increases the operating cost, and (2) adding more input samples to reduce the epistemic uncertainty, which increases the development cost. Once the development cost for testing samples of each variable and relative weights between two costs are quantified and aggregated, a balanced optimum can be estimated through the proposed bi-objective CBDO where a mean vector and the sample size for input variables are included in design vector to be optimized.

The purpose of the proposed method is to provide estimated optimal sample size under given conditions before actually acquiring samples. In other words, the optimal sample size for each input variable is estimated through initial sample estimates since the true parameters cannot be assessed without a large number of samples. Therefore, if the proposed bi-objective CBDO provides the optimal amount of additional data, and new samples are obtained, designers have to conduct CBDO again with new input data to validate the results.

3.1 Uncertainty on input distribution type

Uncertainties on input distribution parameters should be decoupled with uncertainties on input distribution types to employ the framework of RMA for CBDO. Discrete uncertainty on the input distribution type is difficult to handle compared with those of input distribution parameters. Since the proposed research assumes that the input distribution types are known, this section briefly explains several ways to select appropriate input distribution types based on given information.

1. Model identification method

Model identification methods could be practical when multiple types of distribution candidates are given (Kang et al. 2016). For instance, various goodness-of-fit (GOF) tests can verify the suitability of a candidate distribution such as the Kolmogorov-Smirnov (K-S) test. On the other hand, the model selection method is capable of ranking multiple candidate distributions based on a specific criterion, so that the best distribution type fitting the given data can be selected. Akaike information criterion (AIC) and Bayesian information criterion (BIC) are the typical criteria to measure the fitness (Akaike 1974).

2. Johnson distribution

The Johnson distribution is a four-parameter distribution that has flexibility for a wide range of different distributions and used to resolve the difficulty caused by uncertainty in distribution types (Johnson 1949). Since the Johnson distribution includes four parameters, higher moments such as skewness and kurtosis need to be incorporated in the framework of RMA which increases dimension from $2N$ to $4N$ in P-space where N is the number of random variable with insufficient samples.

3. Kernel density estimation

Kernel density estimation (KDE) is an alternative for a parametric distribution when there is no prior knowledge on distribution type (Silverman 2018). As used in previous works on CBDO, an explicit PDF is obtained from given sample data. It is not necessary to determine specific distribution type, and its accuracy gradually increases as the sample size of input variable increases. In addition, when input variables follow a multimodal distribution, KDE can be the best option to handle the input distribution type. Readers can refer to the literature (Jung et al. 2019b) for detailed explanations on exploiting KDE in CBDO.

It is difficult to address that one method is superior since its effectiveness varies according to a given condition, and the thorough comparison between three methods is beyond the purpose of this research. In this paper, the Johnson distribution will be used for comparison since the model identification

method and KDE have been already validated in the previous study (Kang et al. 2016; Jung et al. 2019b). For the comparison, it is assumed that the first two moments of the Johnson distribution will have uncertainties given by (1) and (2), and skewness and kurtosis will be deterministically estimated.

3.2 Formulation of bi-objective confidence-based design optimization

The proposed bi-objective CBDO is formulated as

$$\begin{aligned} & \underset{\mathbf{d}, \mathbf{ND}}{\text{minimize}} \quad w_{\mathbf{d}} \times \text{cost}_{\text{operating}}(\mathbf{d}) + w_{\mathbf{ND}} \times \text{cost}_{\text{development}}(\mathbf{ND}) \\ & \text{subject to} \quad \Pr[\Pr[G(\mathbf{X}) \leq 0] \geq \text{Re}_{\text{Target}}] \geq \text{CL}_{\text{Target}} \\ & \text{where } X_i \sim \begin{cases} \zeta_i(\mu_i(d_i, ND_i, \sigma_i^2), \sigma_i^2(ND_i, s_i)) & \text{for } i = 1, \dots, N_d \\ \zeta_i(\mu_i(*\bar{x}_i, ND_i, \sigma_i^2), \sigma_i^2(ND_i, s_i)) & \text{for } i = N_d + 1, \dots, N_d + N_p \end{cases} \end{aligned} \tag{7}$$

where $w_{\mathbf{d}}$ and $w_{\mathbf{ND}}$ are weights for operating and development cost, respectively; $\mathbf{d} = \{d_1, d_2, \dots, d_{N_d}\}^T$ is the mean vector of random design variables and $\mathbf{ND} = \{ND_1, ND_2, \dots, ND_{N_d+N_p}\}^T$ is the vector including the sample size for input variables including random design variables and parameters; \mathbf{X} includes random design variables and random parameters; ND_i , s_i^2 , $*\bar{x}_i$, and d_i are the sample size, sample variance, sample mean, and design vector for the i th random variable, respectively; N_d is the number of random design variables and N_p is the number of random parameters. In (7), it is assumed that all input variables have input model uncertainty. Variabilities of the mean and variance are shown in (1) and (2), and the distribution type is denoted as ζ .

There are two major differences between conventional CBDO and the proposed bi-objective CBDO:

1. The sample size for input variables denoted as \mathbf{ND} is the design variable in the proposed bi-objective CBDO. Even if the number of data has to be an integer, it will be treated as a continuous positive variable to employ a gradient-based optimization algorithm and sensitivity analysis. That is, not only the mean vector but also the sample size for input variables can change the confidence of reliability. It should be noted that the sample size for random parameters is also included in \mathbf{ND} .
2. The development cost, which is a function of \mathbf{ND} , is included in the objective function as can be seen in (7). Figure 1 shows four examples of functional relationship between the development cost and the sample size for input variables in case of a one-dimensional problem: linear, logarithmic, high-order polynomial, and exponential. As the sample size increases, all functions seem to converge to a point; however, that is only for better visualization.

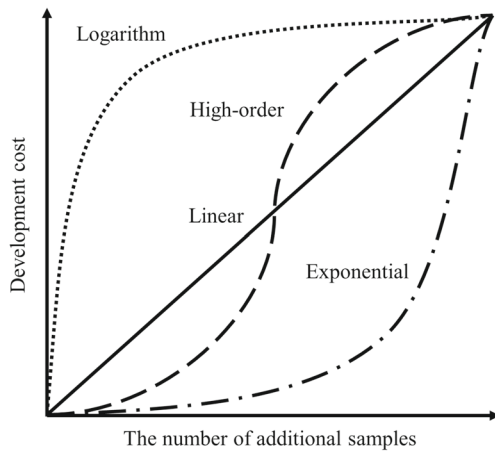


Fig. 1 Four examples of relationship between development cost and number of data

Therefore, to perform the proposed bi-objective CBDO in (7), relative weights which are w_d and w_{ND} between the operating cost and development cost need to be quantified. In addition, functional forms of the operating cost primarily related to design performances and the development cost as a function of the sample size for input variables need to be established. The formulations of operating cost and development cost obviously depend on what kind of applications it will be applied to, and how to quantify each cost is beyond the purpose of the proposed research.

3.3 Sensitivity analysis of confidence with respect to design variables

Sensitivity analysis is an essential process in gradient-based optimizers to provide an accurate and efficient direction for searching the optimum. Since the framework of RMA is adopted, the sensitivity analysis with respect to design vector is identical to the gradient vector of reliability with respect to mean vector and the sample size at the MPP (Jung et al. 2019b). The stochastic sensitivity analysis for reliability with respect to mean vector can be performed analytically using the first-order score functions (Lee et al. 2011). Therefore, sensitivity analysis for confidence constraints with respect to i th design variable is to obtain the gradient of reliability at MPP written as

$$\begin{aligned} \left. \frac{\partial Re}{\partial d_i} \right|_{\psi=\psi^*} &= \left. \frac{\partial Re}{\partial \mu_i} \right|_{\psi=\psi^*} \\ &= \int_{\Omega_R} (G(\mathbf{x})) \frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \psi^*)}{\partial \mu_i} f_{\mathbf{x}}(\mathbf{x}; \psi^*) d\mathbf{x} \end{aligned} \tag{8}$$

where $\frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \psi^*)}{\partial \mu_i}$ is the first-order score function with respect to the mean of i -th random variable and ψ^* is the MPP in P-space which represents the specific realizations of input

distribution parameters obtained from (5). The first-order score functions for parametric distributions have been explicitly derived in the literature (Lee et al. 2011). However, the sensitivity of reliability with respect to the sample size for each input variable is also necessary.

In addition to (8), the proposed optimization demands additional sensitivity analysis with respect to ND formulated as

$$\begin{aligned} \frac{\partial Re}{\partial ND_i} &= \int_{\Omega_R} (G(\mathbf{x})) \frac{d \ln f_{\mathbf{x}}(\mathbf{x}; \psi)}{d ND_i} f_{\mathbf{x}}(\mathbf{x}; \psi) d\mathbf{x} \\ &= \int_{\Omega_R} (G(\mathbf{x})) \left(\frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \psi)}{\partial \mu_i} \frac{d \mu_i}{d ND_i} + \frac{d \ln f_{\mathbf{x}}(\mathbf{x}; \psi)}{d \sigma_i^2} \frac{\partial \sigma_i^2}{\partial ND_i} \right) f_{\mathbf{x}}(\mathbf{x}; \psi) d\mathbf{x} \\ &= \int_{\Omega_R} (G(\mathbf{x})) \left[\frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \psi)}{\partial \mu_i} \left(\frac{\partial \mu_i}{\partial ND_i} + \frac{d \mu_i}{d \sigma_i^2} \frac{\partial \sigma_i^2}{\partial ND_i} \right) \right. \\ &\quad \left. + \left(\frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \psi)}{\partial \sigma_i^2} + \frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \psi)}{\partial \mu_i} \frac{d \mu_i}{d \sigma_i^2} \right) \frac{\partial \sigma_i^2}{\partial ND_i} \right] f_{\mathbf{x}}(\mathbf{x}; \psi) d\mathbf{x} \end{aligned} \tag{9}$$

where ψ includes mean and variance of input variables since two-parameter distribution is assumed in this study.

To further simplify (9), $\frac{dx_i}{dp_i}$ is obtained by taking a derivative of CDF with respect to p_i which is either mean or variance as (Cho et al. 2017)

$$\frac{dx_i}{dp_i} = -\frac{1}{f_{X_i}(x_i; a_i, b_i)} \frac{\partial}{\partial p_i} F_{X_i}(x_i; a_i, b_i) \tag{10}$$

where a_i and b_i are distribution parameters expressed by mean and variance. Using (10), $\frac{\partial \mu_i}{\partial ND_i}$ in (9) is derived as

$$\begin{aligned} \frac{\partial \mu_i}{\partial ND_i} &= \frac{\partial \mu_i}{\partial \sigma_i^2} \frac{\partial \sigma_i^2}{\partial ND_i} \\ &= -\frac{(-ND_i(\mu_i - \bar{x}_i)/\sigma_i^2)}{f_{\mu}(\mu_i; \bar{x}_i, \sqrt{\sigma_i^2/ND_i})} \times \phi\left(\frac{\mu_i - \bar{x}_i}{\sqrt{\sigma_i^2/ND_i}}\right) \times \left(-\frac{1}{2} \sqrt{\sigma_i^2/ND_i^3}\right) \\ &= -\frac{1}{2} \left(\frac{\mu_i - \bar{x}_i}{ND_i}\right) \end{aligned} \tag{11}$$

where $\sigma_{\mu_i}^2$ is the variance of μ_i expressed as $\sqrt{\frac{\sigma_i^2}{ND_i}}$ distinguished from σ_i^2 , and f_{μ} is PDF of mean described in (2). Similarly,

$$\begin{aligned} \frac{\partial \sigma_i^2}{\partial ND_i} &= \frac{\partial \sigma_i^2}{\partial \alpha} \frac{\partial \alpha}{\partial ND_i} + \frac{\partial \sigma_i^2}{\partial \beta} \frac{\partial \beta}{\partial ND_i} \\ &= -\frac{1}{2} \frac{1}{f_{\sigma^2}(\sigma_i^2; \alpha, \beta)} \times \left[\frac{(\ln(\beta/\sigma_i^2) - \psi_0(\alpha)) \Gamma(\alpha, \beta/\sigma_i^2) + T(3, \alpha, \beta/\sigma_i^2)}{\Gamma(\alpha)} \right. \\ &\quad \left. + \left(-\frac{e^{-\beta/\sigma_i^2} (\beta/\sigma_i^2)^{\alpha}}{\beta \Gamma(\alpha)} \right) \sigma_i^2 \right] \end{aligned} \tag{12}$$

where $T(\bullet, \bullet, \bullet)$ is a Meijer G-function; $\Gamma(s, x) = \int_x^{\infty} t^{s-1} e^{-t} dt$ is an upper incomplete gamma function; $\alpha = \frac{ND_i - 1}{2}$ and β

$= \frac{(ND_i-1)s_i^2}{2}$ are the shape and scale parameter of an inverse-gamma distribution of σ_i^2 , respectively. f_{σ^2} is PDF of variance described in (1). Since μ_i and σ_i^2 are correlated as can be seen from (2), $\frac{d\mu_i}{d\sigma_i^2}$ is obtained by taking a derivative on (2) as

$$\begin{aligned} \frac{d\mu_i}{d\sigma_i^2} &= -\frac{(-ND_i(\mu_i - \bar{x}_i)/\sigma_i^2)}{f_{\mu}(\mu_i, \bar{x}_i, \sqrt{\sigma_i^2/ND_i})} \times \phi\left(\frac{\mu_i - \bar{x}_i}{\sqrt{\sigma_i^2/ND_i}}\right) \times \frac{1}{2\sigma_i\sqrt{ND_i}} \\ &= \frac{1}{2} \left(\frac{\mu_i - \bar{x}_i}{\sigma_i^2}\right) \end{aligned} \tag{13}$$

Finally, the sensitivity in (9) is capable of being calculated through (11) to (13). Note that \bar{x}_i would be replaced with the i th entry of the design vector d_i when the i th random variable is a design variable, not a random parameter.

Through the sensitivity analysis in this section, the optimizer can explore next design candidates. The sensitivity analysis with respect to mean vector and the sample size in this section should be distinguished from the sensitivity analysis during MPP-search in P-space. This process is performed after the MPP is found as (5).

3.4 Overall procedures

As shown in Fig. 2, the overall procedure of the proposed bi-objective CBDO is very similar to PMA of RBDO except for the Rosenblatt transformation to P-space and the sensitivity analysis with respect to the sample size for input variables. The identification of input distribution type in Fig. 2 means determination of distribution types among parametric distributions, Johnson distribution, and KDE. In the bi-objective CBDO, MPP in P-space could be found at a current design point as (5) by MPP-search algorithm, and gradient-based optimizers such as sequential quadratic programming (SQP) provides appropriate search direction and step size for searching the optimal design vector as (7).

It should be noted that there are two main assumptions made for the bi-objective CBDO: (1) the confidence of reliability is treated in MPP-based approach as conventional CBDO, so that the linearization of reliability with respect to input distribution parameters is employed, and (2) the sample estimates to quantify the input parameter distributions are invariant during the optimization since there is no additional knowledge to change it. In other words, the variabilities of input distribution parameters are only affected by the mean vector \mathbf{d} and the sample size \mathbf{ND} since the input dataset is not actually updated during the bi-objective CBDO even though \mathbf{ND} is changed.

4 Numerical studies: 2D mathematical example

A mathematical 2D optimization problem widely used in the previous RBDO studies is analyzed in various ways. Firstly, the proposed stochastic sensitivity analysis with respect to the sample size for input variables in Sect. 3.3 is validated. Secondly, the 2D bi-objective CBDO is tested with various weights for operating and development costs, and the discrepancies with true optimum are calculated to capture the error due to the bias of sample estimates. Thirdly, the Johnson distribution is used to grasp the error compared with the results when the true distribution type is identified. Fourthly, various types of development cost functions are tested for comparison. Finally, repeated tests with various initial input datasets are performed.

The bi-objective CBDO for the 2D example is formulated as

$$\begin{aligned} \underset{\mathbf{d}, \mathbf{ND}}{\text{minimize}} \quad & w \times \left(-\frac{(d_1 + d_2 - 10)^2}{30} - \frac{(d_1 - d_2 + 10)^2}{120} \right) \tag{14} \\ & + \text{cost}_{\text{development}}(ND_1, ND_2) \end{aligned}$$

subject to $\Pr[\Pr[G_j(\mathbf{X}) \leq 0] \geq 97.72\%] \geq 90.00\%$ for $j = 1, 2, 3$

where $G_1(\mathbf{X}) = 1 - \frac{X_1^2 X_2}{20}$,

$$\begin{aligned} G_2(\mathbf{X}) &= -1 + (0.9063X_1 + 0.4226X_2 - 6)^2 \\ &\quad + (0.9063X_1 + 0.4226X_2 - 6)^3 \\ &\quad - 0.6(0.9063X_1 + 0.4226X_2 - 6)^4 \\ &\quad - (-0.4226X_1 + 0.9063X_2), \end{aligned}$$

$$G_3(\mathbf{X}) = 1 - \frac{80}{X_1^2 + 8X_2 + 5},$$

$$X_i \sim N(\mu_i(d_i, s_i^2, ND_i), \sigma_i^2(s_i^2, ND_i)),$$

$$10 \leq ND_i \leq 100 \quad \text{for } i = 1, 2$$

$$w = \frac{w_{\mathbf{d}}}{w_{\mathbf{ND}}}$$

where target confidence and target reliability are set to 90.00% and 97.72%, respectively. The initial number of data is set to 10 for both random variables. w is the weight ratio of two costs, and $\text{cost}_{\text{development}}(ND_1, ND_2)$ is the development cost. The operating cost is assumed to be the same as the objective function of the original optimization problem. Lower bounds of \mathbf{ND} are set to 10 which is the initial number of data, and upper bounds are set to 100. The initial samples are drawn from a normal distribution with the true variance of 0.5². Note that initial design vector is RBDO optimum obtained from deterministic sample estimates to enhance the convergence.

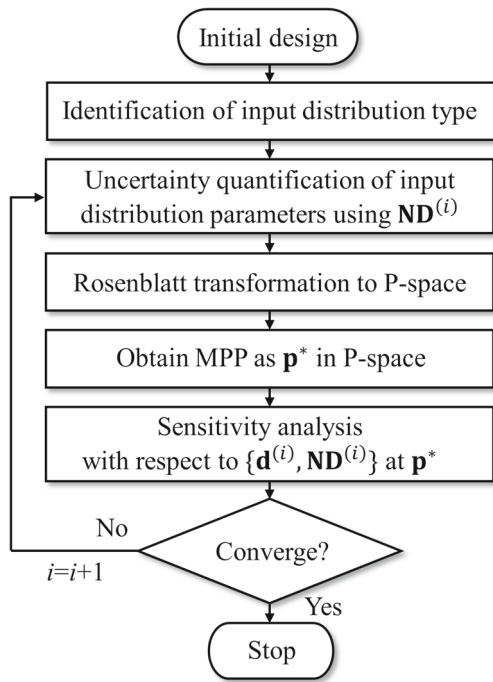


Fig. 2 Flowchart of the proposed bi-objective CBDO

4.1 Validation of sensitivity analysis

Sensitivity analysis of reliability with respect to **ND** is validated in this section by comparing analytically calculated values with results of the finite difference method (FDM). Since sensitivity with respect to mean vector is already validated in the previous works (Lee et al. 2010), only the results for the sample size are shown in this section. The highly nonlinear second constraint in (14) is used for the comparison. Therefore, when the vector of the sample size is $\mathbf{ND} = \{10, 10\}^T$ at the lower bound, the MPP, $\mathbf{p}^* = \{1.1853, 0.4671, 0.1276, -0.0551\}^T$, is obtained through the hybrid mean value (HMV) method (Youn et al. 2003) in P-space using the given sample mean and variance. The number of samples to compute the reliability using MCS is 10^8 . Table 1 shows sensitivity analysis results obtained from FDM and the proposed method with two different **ND** cases. To alleviate sampling uncertainty due to repetitive random samplings, the random seeds for MCS are controlled in this test. It can be

Table 1 Results of sensitivity analysis of reliability with respect to the number of data

Analysis method	10 samples		15 samples	
	$\frac{\partial Re}{\partial ND_1}$	$\frac{\partial Re}{\partial ND_2}$	$\frac{\partial Re}{\partial ND_1}$	$\frac{\partial Re}{\partial ND_2}$
FDM (1.0%)	-0.0029	-1.714E-4	-9.338E-4	-5.920e-5
FDM (0.5%)	-0.0029	-1.750E-4	-9.520E-4	-6.426E-5
FDM (0.1%)	-0.0029	-1.720E-4	-9.566E-4	-6.400E-5
Proposed method	-0.0030	-1.706E-4	-9.577E-4	-6.019E-5

Table 2 Results of bi-objective CBDO with various weight ratios

Weight ratio	Bi-objective CBDO optimum			
	μ_{X_1}	μ_{X_2}	ND_1	ND_2
100	4.5907	2.6149	10.3577	10.5671
200	4.6237	2.4863	13.3243	14.6549
300	4.6362	2.4249	15.7140	18.1443
400	4.6431	2.3873	17.8801	21.1178
500	4.6475	2.3581	19.7201	24.2515
1000	4.6580	2.2893	28.2633	35.6679
2000	4.6630	2.2349	41.5480	54.0487
5000	4.6771	2.1776	70.0856	91.9751

concluded from Table 1 that the proposed sensitivity analysis is very accurate compared with FDM.

4.2 Results of various weights for two costs

Effect of the weight ratio in the bi-objective CBDO is shown in this section. The development function is set to $ND_1 + ND_2$, in which the cost is linear to each number of data. Increasing the weight ratio indicates that the operating cost becomes higher than the development cost. Thus, the optimizer would try to reduce the operating cost by adding more data rather than shifting the mean vector. The test results of various weight ratios are listed in Table 2.

Table 2 shows that the optimal sample size increases as the weight ratio increases, representing that large weight on operating cost forces to add more samples to satisfy the confidence constraints by reducing uncertainty of input distribution models. Similarly, the optimal sample size decreases as the weight ratio decreases meaning that the optimization tries to find an optimum by sacrificing the operating cost since adding more samples is much more expensive. In case of $w = 500$, the bi-objective CBDO recommends to add 10 more input samples for X_1 and 14 more for X_2 to reduce the total cost.

Table 3 shows results of the bi-objective CBDO when sample estimates are inaccurate compared with the true bi-objective CBDO optimum where its sample variance is equal to the true variance (i.e., population variance) of 0.25 for both

Table 3 Validation of bi-objective CBDO optima under inaccurate sample variance

Sample variance	Optimum = $\{\mu_{X_1}, \mu_{X_2}, ND_1, ND_2\}^T$	Total cost
0.20	$\{4.674, 2.095, 36.466, 51.846\}^T$	- 3244.5
0.22	$\{4.673, 2.156, 38.536, 52.820\}^T$	- 3190.0
0.24	$\{4.672, 2.215, 43.312, 52.461\}^T$	- 3136.5
0.25 (true)	$\{4.666, 2.246, 42.153, 54.138\}^T$	- 3110.1
0.26	$\{4.664, 2.276, 43.167, 54.152\}^T$	- 3084.5
0.28	$\{4.661, 2.338, 48.551, 50.500\}^T$	- 3031.7
0.30	$\{4.651, 2.397, 50.171, 52.744\}^T$	- 2980.2

random variables. The weight ratio is set to 2000. As the initial sample size increases, the sample variance will become closer to the true variance. It can be seen that discrepancy between estimated and true total costs gradually reduces as the sample variance is closer to the true variance. Consequently, Table 3 shows how much error quality of initial sample can cause to estimate the optimal sample size. For instance, if we have the initial dataset having the sample variance as 0.22, the estimated optimal sample size is $\{39, 53\}^T$. However, the true optimum is $\{42, 54\}^T$ which makes the total cost minimum. The discrepancy between two optima indicates the error due to bias of initial sample estimates.

4.3 Results of uncertainty on input distribution type

This section shows feasibility of the Johnson distribution instead of the parametric distribution selected from prior knowledge or model identification methods. In this test, the identical dataset used in the previous section. The weight ratio is set to 1000 with linear development cost, and the initial sample size for two random variables is 10. Note that each trial has a different dataset so that different optima are obtained since the sample variance, skewness, and kurtosis would be all different. Table 4 compares two results of the bi-objective CBDO with normal distributions and the Johnson distribution as an input, respectively. The optimal sample size using population variance is $\mathbf{ND} = \{28.7602, 35.6763\}^T$, and the mean of 5

Table 4 Bi-objective CBDO results of two different input distribution types

Trial number	Normal distribution				Johnson distribution			
	μ_{X_1}	μ_{X_2}	ND_1	ND_2	μ_{X_1}	μ_{X_2}	ND_1	ND_2
1	4.6148	2.4770	37.2448	35.5813	4.4758	2.2982	26.5322	29.2628
2	4.6455	2.3985	31.0973	37.0581	4.5946	2.3249	23.4575	33.1698
3	4.6657	2.3394	27.1869	38.9949	4.7407	2.4566	22.9170	35.4182
4	4.6884	2.2569	22.6936	38.4188	4.5820	2.1131	25.6007	30.8697
5	4.6495	2.4094	30.3166	37.7880	4.6764	2.3846	26.4505	33.4234
Mean	-	-	29.7020	33.5588	-	-	24.9915	31.8287

Table 5 Results of bi-objective CBDO with various weight ratios under logarithm development cost

Weight ratio	Bi-objective CBDO optimum			
	μ_{X_1}	μ_{X_2}	ND_1	ND_2
100	4.5994	2.5953	10.4875	10.9738
200	4.6439	2.3869	17.0697	21.1062
300	4.6540	2.2933	25.2740	36.3807
400	4.6626	2.2367	35.8474	55.8690
500	4.6680	2.1991	49.2196	80.1483
1000	4.6845	2.1641	99.8540	100.0000
2000	4.6844	2.1641	100.0000	100.0000
5000	4.6844	2.1641	100.0000	100.0000

trials is shown in the last row of Table 4. It is shown that the Johnson distribution can be successfully implemented in the proposed framework even though there is an error because the true distribution is a normal distribution, and the biased sample skewness and kurtosis are used. However, deviation of each optimum seems to be smaller for the Johnson distribution because higher-order moments are considered even though they have no uncertainty. In addition, the Johnson distribution can widely cover various types of distributions which may be more appropriate in the real world.

4.4 Results of various development costs

In real engineering problems, the behavior of the development cost would vary depending on applications. As shown in Fig. 1, it behaves like a simple linear function or nonlinear functions. Therefore, to verify the effectiveness of the proposed bi-objective CBDO in various conditions, two kinds of nonlinear development cost functions with opposite properties are additionally tested in this section.

First, a logarithm function given by

$$10\log(ND_1 \times ND_2) \tag{15}$$

Table 6 Results of bi-objective CBDO varying weight ratio under exponential development cost

Weight ratio	Bi-objective CBDO optimum			
	μ_{X_1}	μ_{X_2}	ND_1	ND_2
100	4.6364	2.4624	14.1546	15.3662
200	4.6487	2.4066	16.8988	18.6470
300	4.6529	2.3800	18.7169	20.7489
400	4.6543	2.3640	20.0977	22.3195
500	4.6563	2.3522	21.1892	23.5811
1000	4.6621	2.3195	24.8491	27.7140
2000	4.6670	2.2933	28.8465	32.1357
5000	4.6702	2.2663	34.6163	38.3892

is used for the development cost and test results with various weight ratios are listed in Table 5. Since the logarithm development function in (15) has gradually decreasing slope, the optimal sample size increases rapidly as the weight for operating cost increases.

Secondly, the development cost is given as an exponential function written as

$$\exp\left(\frac{ND_1}{10}\right) + \exp\left(\frac{ND_2}{10}\right) \tag{16}$$

The results are listed in Table 6 which indicates that larger sample size is avoided even for large weight ratio due to rapid increases in the development cost. The results of these different trends for the two contrasting functions, logarithm and exponential, support the feasibility of the proposed method.

4.5 Repeated tests with various initial samples

Previous test results are obtained using the same initial samples for fair comparisons. In this section, iteratively generated

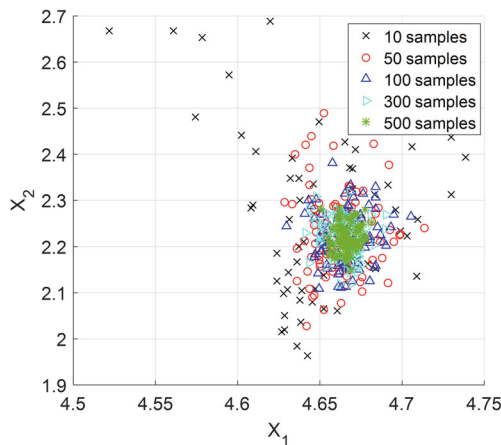


Fig. 3 Optima of 100 repeated multi-objective CBDOs

Table 7 Variance of optima for each number of samples

No. of samples	Variance of optima			
	μ_{X_1}	μ_{X_2}	ND_1	ND_2
10	1.688E-3	3.802E-2	606.6	602.5
50	2.948E-4	8.056E-3	227.9	200.1
100	1.876E-4	3.468E-3	162.4	157.4
150	1.872E-4	4.690E-3	138.7	145.6
200	9.886E-5	2.035E-3	56.1	110.6
300	9.465E-5	1.648E-3	77.3	78.6
400	6.467E-5	1.241E-3	47.3	114.6
500	3.659E-5	8.200E-4	24.1	93.0
600	3.504E-5	6.789E-4	27.1	54.7

input datasets are repeatedly tested to demonstrate robustness of the proposed method. Other parameters and conditions such as weight ratio, development cost, and true input statistical models are invariant during the repeated tests. The different initial dataset in the bi-objective CBDO means that its sample estimates are different, thereby resulting in the different optima. With more data, sample variances are gradually closer to the true variance, but the sample size in the design variable is fixed as 10 in tests. In other words, the actual sample size for calculating the sample variance increases for validation since the purpose of repeated tests is to capture the effect on the bias of sample variance. However, it does not mean that the initial sample size in the design vector increases.

In Fig. 3, optima of the repeated tests with different sample variances are illustrated where 100 trials are performed for each case. It can be seen in Table 7 that variability of optima decreases as more samples are used to estimate the sample variance meaning that the optimal mean vectors gradually converge to the true optimum. The number of data in Fig. 3 and Table 7 is only for calculation of the sample variance and has no relation with the sample size in the design vector. Note that the optimum often goes to the boundary of the design space since the confidence of reliability may have multiple local optima.

5 Engineering example: multi-scale composite frame optimization

To validate feasibility of the proposed method, multi-scale composite frame optimization is used in this section. Deterministic multi-scale design optimization of the composite frames for the minimum structural compliance with manufacturing constraints (Yan et al. 2017), maximum fundamental frequency design with continuous fiber winding angles (Duan et al. 2018), and a two-step optimization scheme for forcing convexity of fiber winding angles in the composite

Table 8 Material properties of the uni-directional carbon fiber-reinforced epoxy

Material property	E_{11}	$E_{22} = E_{33}$	G_{12}	G_{13}	G_{23}	ν_{12}	ν_{13}	ν_{23}	ρ
Value	143 GPa	10 GPa	6 GPa	5 GPa	3 GPa	0.3	0.2	0.52	1800 kg/m ³

frames (Duan et al. 2019) have been studied. Similarly with the previous studies considering practical engineering applications, it is assumed in this example that each composite tube has the same number of layers, (i.e., $N_{lay} = 20$) with the same thickness in the initial design of 0.1 mm for the sake of simplicity. The fiber candidate material is carbon fiber-reinforced epoxy with orthotropic properties as listed in Table 8. Loading/boundary conditions and geometry of the composite frame are shown in Fig. 4. There are four deterministic design variables, radius of each beam, and three random parameters under insufficient data that are longitudinal modulus, magnitude, and direction of the load. True mean and variance of each random parameter are listed in Table 9.

5.1 Formulation of bi-objective confidence-based design optimization for 4-beam composite frame

The bi-objective CBDO for 4-beam composite frame structure is formulated as

$$\begin{aligned}
 &\text{given } \bar{p}_i, s_i^2 && (17) \\
 &\underset{\mathbf{r}, \mathbf{ND}}{\text{minimize}} && w \times V(\mathbf{r}) + \text{cost}_{\text{Development}}(\mathbf{ND}) \\
 &\text{subject to } \Pr[\Pr[C(\mathbf{r}, \mathbf{P}) \leq C_{\text{Target}}] \geq \text{Re}_{\text{Target}}] \geq \text{CL}_{\text{Target}} \\
 &&& P_i \sim N\left(\mu_i(\bar{p}_i, s_i^2, ND_i), \sigma_i^2(s_i^2, ND_i)\right), \\
 &&& 10 \leq ND_i \leq 100 \text{ for } i = 1, 2, 3 \\
 &&& 0.05 \leq r_i \leq 1 \text{ for } j = 1, 2, 3, 4
 \end{aligned}$$

where $V(\mathbf{r})$ is the total volume; $C(\mathbf{r}, \mathbf{P})$ is the compliance as a function of radius vector \mathbf{r} and three random parameters

denoted as \mathbf{P} in Table 9; C_{Target} , $\text{Re}_{\text{Target}}$, and $\text{CL}_{\text{Target}}$ are the target compliance, target reliability, and target confidence, respectively; \bar{p}_i and s_i^2 are the sample mean and variance of the i th random parameter in Table 9 obtained from the initial samples; ND_i is the sample size for the i th random parameter. The target compliance and reliability are set to 0.7 and 95%, respectively. Various target confidences and weight ratios between two costs are utilized for validation. Linear development cost is used for the bi-objective CBDO. The initial radius is 0.25 mm for all beams, and the initial sample size for three random parameters is set to 10. In this example, Kriging model for compliance is utilized to improve computational efficiency for reliability estimation. To generate the surrogate models, 300 samples in the design domain by the Latin hypercube sampling are used.

5.2 Results of bi-objective confidence-based design optimization

Results of the bi-objective CBDO under various weights are listed in Table 10. Volume of the structure decreases as the weight ratio increases since large weight means that reducing volume is relatively more valuable than the development cost for increasing sample size. Therefore, the optimal sample size gradually increases as the weight ratio increases. The estimated optimal radii of 4 beams are exact only when the sample variance of the initial sample is exact. It is evident that the RBDO optimum has the minimal volume since it has no epistemic uncertainty. On the other hand, increasing the target confidence means more conservative design is achieved, so that the optimal volume and sample size increase as listed in

Fig. 4 Configuration of 4-beam composite frame structure

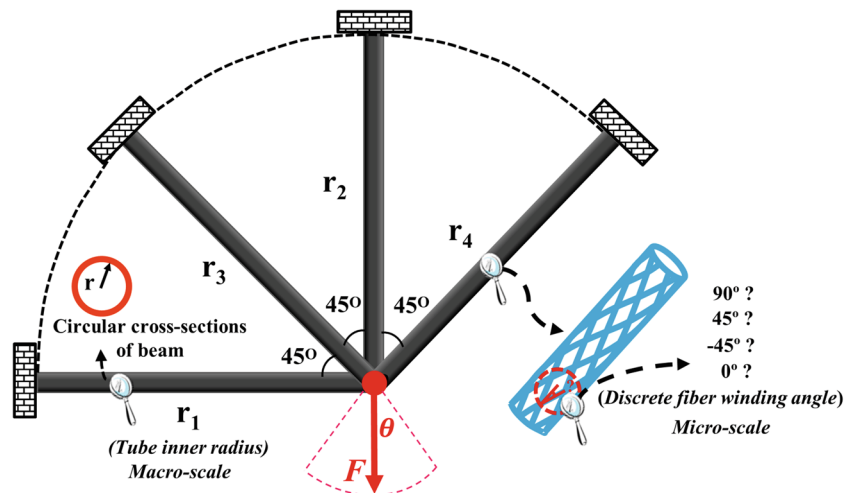


Table 9 Statistical representations of random parameters

Random parameter	Symbol	Mean	Standard deviation	Distribution type
Longitudinal modulus (Pa)	$E_{11} \equiv P_1$	1.43E11	1.43E10	Normal
Load magnitude (N)	$F \equiv P_2$	- 1.00E4	1.00E3	Normal
Load direction (°)	$\theta \equiv P_3$	0	10	Normal

Table 10 Optima of bi-objective CBDO with various weights between two costs

Weight	Bi-objective CBDO optimum							Volume
	r_1	r_2	r_3	r_4	ND_1	ND_2	ND_3	
500	0.6460	0.2816	0.0500	0.0500	20.0326	23.1935	10.0000	0.0260
1000	0.5887	0.3312	0.0500	0.0500	22.8590	29.0776	10.0000	0.0258
2000	0.6346	0.2676	0.0503	0.0500	34.9868	43.8620	10.0000	0.0254
RBDO	0.6226	0.2737	0.0500	0.0500	-	-	-	0.0252

Table 11. Both costs are raised since the target confidence to be satisfied is increased, leading to much conservative design. In all tests, ND_3 which is the sample size for loading direction goes to the lower bound implying that the loading direction is relatively insignificant on compliance compared with material properties and load magnitude.

6 Conclusion

Bi-objective CBDO accounting for both operating and development costs simultaneously is proposed for practical application of CBDO. The overall process of the bi-objective CBDO is developed based on the CBDO framework to efficiently handle the confidence of reliability that is derived from epistemic uncertainty. The confidence of reliability increases by collecting more input samples which increases the development cost as well as shifting the design vector toward feasible domain which increases the operating cost. Thus, the objective function includes both costs, and the sample size for input variables is handled as the design variable in the bi-objective CBDO which enable to facilitate decision making for the designers on how to allocate engineering efforts. The

estimated optimal sample size is affected by the relative weights between two costs and the specification of a given system such as the process to acquire input samples, represented as an explicit expression for a development cost. Since the sample size changes during the optimization, the stochastic sensitivity analysis of confidence with respect to the sample size for input variables is developed to avoid repetitive reliability computations. Although the optimal sample size is obtained based on the assumption that the sample estimates are invariant during optimization, the estimated optimum can guarantee accuracy as the initial sample size increases. Various numerical tests and one engineering application have successfully supported the effectiveness of the proposed method. In consequence, the proposed bi-objective CBDO answers to the question: *How to determine sample size for input variables in RBDO to minimize the overall cost?*

On the other hand, there are many types of epistemic uncertainties besides the uncertainty of input statistical models, and each epistemic uncertainty can be reduced through additional information such as experiments to validate the simulation model. Therefore, it is crucial to consider the cost for gathering information which enables to reduce the epistemic uncertainty in practical applications. In future work, the multi-

Table 11 Optima of bi-objective CBDO with various target confidences

CL_{Target}	Bi-objective CBDO optimum							Volume
	r_1	r_2	r_3	r_4	ND_1	ND_2	ND_3	
80%	0.6115	0.2918	0.0500	0.0500	17.8362	21.0700	10.0000	0.0254
85%	0.6147	0.2962	0.0500	0.0500	19.7775	24.0683	10.0000	0.0256
90%	0.5887	0.3312	0.0500	0.0500	22.8590	29.0776	10.0000	0.0258
95%	0.6144	0.3167	0.0500	0.0500	27.4510	31.8341	10.0000	0.0261
RBDO	0.6226	0.2737	0.0500	0.0500	-	-	-	0.0252

objective CBDO including model uncertainty to distribute the optimal number of experiments, simulations, and collecting input data will be investigated.

7 Replication of results

Matlab codes for the mathematical examples in Sect. 4 are uploaded on <https://github.com/Yongsu-Jung/Bi-objective-CBDO.git>. Overall concepts and algorithms can be validated through the mathematical example.

Funding information This research was supported by the development of thermoelectric power generation system and business model utilizing non-use heat of industry funded by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade Industry & Energy (MOTIE) of the Republic of Korea (No.20172010000830).

Compliance with ethical standards

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

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723
- Bae S, Kim NH, Jang SG (2018) Reliability-based design optimization under sampling uncertainty: shifting design versus shaping uncertainty. *Struct Multidiscip Optim* 57(5):1845–1855
- Bae S, Kim NH, Jang SG (2019) System reliability-based design optimization under tradeoff between reduction of sampling uncertainty and design shift. *J Mech Des* 141(4):041403
- Cho H, Choi KK, Gaul NJ, Lee I, Lamb D, Gorsich D (2016) Conservative reliability-based design optimization method with insufficient input data. *Struct Multidiscip Optim* 54(6):1609–1630
- Cho H, Choi KK, Lamb D (2017) Sensitivity developments for RBDO with dependent input variable and varying input standard deviation. *J Mech Des* 139(7):071402
- Der Kiureghian A, Ditlevsen O (2009) Aleatory or epistemic? Does it matter? *Struct Saf* 31(2):105–112
- Duan Z, Yan J, Lee I, Wang J, Yu T (2018) Integrated design optimization of composite frames and materials for maximum fundamental frequency with continuous fiber winding angles. *Acta Mech Sinica* 34(6):1084–1094
- Duan Z, Yan J, Lee I, Lund E, Wang J (2019) A two-step optimization scheme based on equivalent stiffness parameters for forcing convexity of fiber winding angle in composite frames. *Struct Multidiscip Optim* 59(6):2111–2129
- Dubourg V, Sudret B (2014) Meta-model-based importance sampling for reliability sensitivity analysis. *Struct Saf* 49:27–36
- Gelman A, Stern HS, Carlin JB, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian data analysis. 3rd edn. Chapman and Hall/CRC
- Gunawan S, Papalambros PY (2006) A Bayesian approach to reliability-based optimization with incomplete information. *J Mech Des* 128(4):909–918
- Hao P, Wang Y, Liu C, Wang B, Wu H (2017) A novel non-probabilistic reliability-based design optimization algorithm using enhanced chaos control method. *Comput Methods Appl Mech Eng* 318:572–593
- Ito M, Kim NH, Kogiso N (2018) Conservative reliability index for epistemic uncertainty in reliability-based design optimization. *Struct Multidiscip Optim* 57(5):1919–1935
- Jiang Z, Chen W, Fu Y, Yang RJ (2013) Reliability-based design optimization with model bias and data uncertainty. *SAE Int J Mater Manuf* 6(3):502–516
- Johnson NL (1949) Systems of frequency curves generated by methods of translation. *Biometrika* 36(1/2):149–176
- Jung Y, Cho H, Lee I (2019a) MPP-based approximated DRM (ADRM) using simplified bivariate approximation with linear regression. *Struct Multidiscip Optim* 59(5):1761–1773
- Jung Y, Cho H, Lee I (2019b) Reliability measure approach for confidence-based design optimization under insufficient input data. *Struct Multidiscip Optim*. <https://doi.org/10.1007/s00158-019-02299-3>
- Kang YJ, Lim OK, Noh Y (2016) Sequential statistical modeling method for distribution type identification. *Struct Multidiscip Optim* 54(6):1587–1607
- Kang SB, Park JW, Lee I (2017) Accuracy improvement of the most probable point-based dimension reduction method using the hessian matrix. *Int J Numer Methods Eng* 111(3):203–217
- Kang K, Qin C, Lee B, Lee I (2019) Modified screening-based Kriging method with cross validation and application to engineering design. *Appl Math Model* 70:626–642
- Lee I, Choi KK, Gorsich D (2010) Sensitivity analyses of FORM-based and DRM-based performance measure approach (PMA) for reliability-based design optimization (RBDO). *Int J Numer Methods Eng* 82(1):26–46
- Lee I, Choi KK, Noh Y, Zhao L, Gorsich D (2011) Sampling-based stochastic sensitivity analysis using score functions for RBDO problems with correlated random variables. *J Mech Des* 133(2):021003
- Lee I, Noh Y, Yoo D (2012) A novel second-order reliability method (SORM) using noncentral or generalized chi-squared distributions. *J Mech Des* 134(10):100912
- Lee I, Choi KK, Noh Y, Lamb D (2013) Comparison study between probabilistic and possibilistic methods for problems under a lack of correlated input statistical information. *Struct Multidiscip Optim* 47(2):175–189
- McDonald M, Zaman K, Mahadevan S (2013) Probabilistic analysis with sparse data. *AIAA J* 51(2):281–290
- Meng Z, Li G, Wang BP, Hao P (2015) A hybrid chaos control approach of the performance measure functions for reliability-based design optimization. *Comput Struct* 146:32–43
- Moon MY, Choi KK, Cho H, Gaul N, Lamb D, Gorsich D (2017) Reliability-based design optimization using confidence-based model validation for insufficient experimental data. *J Mech Des* 139(3):031404
- Moon MY, Cho H, Choi KK, Gaul N, Lamb D, Gorsich D (2018) Confidence-based reliability assessment considering limited numbers of both input and output test data. *Struct Multidiscip Optim* 57(5):2027–2043
- Moon MY, Choi KK, Gaul N, Lamb D (2019) Treating epistemic uncertainty using bootstrapping selection of input distribution model for confidence-based reliability assessment. *J Mech Des* 141(3):031402
- Moustapha M, Sudret B (2019) Surrogate-assisted reliability-based design optimization: a survey and a unified modular framework. *Struct Multidiscip Optim*. <https://doi.org/10.1007/s00158-019-02290-y>
- Nannapaneni S, Mahadevan S (2016) Reliability analysis under epistemic uncertainty. *Reliab Eng Syst Saf* 155:9–20
- Noh Y, Choi KK, Lee I, Gorsich D, Lamb D (2011) Reliability-based design optimization with confidence level under input model uncertainty due to limited test data. *Struct Multidiscip Optim* 43(4):443–458
- Pan H, Xi Z, Yang RJ (2016) Model uncertainty approximation using a copula-based approach for reliability based design optimization. *Struct Multidiscip Optim* 54(6):1543–1556

- Peng X, Li J, Jiang S (2017) Unified uncertainty representation and quantification based on insufficient input data. *Struct Multidiscip Optim* 56(6):1305–1317
- Picheny V, Kim NH, Haftka RT (2010) Application of bootstrap method in conservative estimation of reliability with limited samples. *Struct Multidiscip Optim* 41(2):205–217
- Rosenblatt M (1952) Remarks on a multivariate transformation. *Ann Math Stat* 23(3):470–472
- Sankararaman S, Mahadevan S (2011) Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data. *Reliab Eng Syst Saf* 96(7):814–824
- Sankararaman S, McLemore K, Mahadevan S, Bradford SC, Peterson LD (2013) Test resource allocation in hierarchical systems using Bayesian networks. *AIAA J* 51(3):537–550
- Silverman BW (2018) *Density estimation for statistics and data analysis*. Routledge
- Tu J, Choi KK, Park YH (1999) A new study on reliability-based design optimization. *J Mech Des* 121(4):557–564
- Xi Z (2019) Model-based reliability analysis with both model uncertainty and parameter uncertainty. *J Mech Des* 141(5):051404
- Yan J, Duan Z, Lund E, Wang J (2017) Concurrent multi-scale design optimization of composite frames with manufacturing constraints. *Struct Multidiscip Optim* 56(3):519–533
- Yoo D, Lee I (2014) Sampling-based approach for design optimization in the presence of interval variables. *Struct Multidiscip Optim* 49(2):253–266
- Youn BD, Wang P (2008) Bayesian reliability-based design optimization using eigenvector dimension reduction (EDR) method. *Struct Multidiscip Optim* 36(2):107–123
- Youn BD, Choi KK, Park YH (2003) Hybrid analysis method for reliability-based design optimization. *J Mech Des* 125(2):221–232

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