



Reliability measure approach for confidence-based design optimization under insufficient input data

Yongsu Jung¹ · Hyunkyoo Cho² · Ikjin Lee¹

Received: 18 December 2018 / Revised: 1 May 2019 / Accepted: 2 May 2019 / Published online: 11 June 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

In most of the reliability-based design optimization (RBDO) researches, accurate input statistical model has been assumed to concentrate on the variability of random variables; however, only a limited number of data are available to quantify the input statistical model in many practical engineering applications. In other words, irreducible variability and uncertainty due to lack of knowledge exist simultaneously in random design variables, which may result in uncertainty of reliability. Therefore, the uncertainty induced by insufficient data has to be accounted for RBDO to guarantee the confidence of reliability. Using the Bayesian approach, the uncertainty of input distributions is successfully propagated to a cumulative distribution function (CDF) of reliability under reasonable assumptions, but it requires a number of function evaluations in double-loop Monte Carlo simulation (MCS). To tackle this challenge, the reliability measure approach (RMA) in confidence-based design optimization (CBDO) is proposed to handle the uncertainty of reliability following the idea of performance measure approach (PMA) in RBDO. Input distribution parameters are transformed to random variables following the standard normal distribution for the most probable point (MPP) search based on the proposed stochastic sensitivity analysis of reliability. Therefore, the reliability is approximated at MPP with respect to input distribution parameters. The proposed CBDO can treat confidence constraints employing the reliability value at the target confidence level that is approximated by MPP in standard normal space. In conclusion, CBDO can be performed in a probabilistic space of input distribution parameters corresponding to the conventional U-space in RBDO to yield the probability (confidence) that reliability is larger than the target reliability. The proposed method can significantly reduce the number of function evaluations by eliminating outer-loop MCS while maintaining acceptable accuracy. Numerical examples are used to demonstrate the effectiveness of the developed sensitivity analysis and RMA to estimate the confidence of reliability in CBDO.

Keywords Reliability-based design optimization (RBDO) · Performance measure approach (PMA) · Epistemic uncertainty · Kernel density estimation (KDE) · Conservative RBDO (CRBDO)

1 Introduction

Reliability has been utilized to measure the safety of the system in the presence of various uncertainties that can be classified as inherent variability (i.e., aleatory uncertainty) and reducible uncertainty due to a lack of knowledge (i.e.,

epistemic uncertainty). These two uncertainties are propagated to the uncertainty of the system performances that may lead to unexpected failure of the system, so that accurate reliability analysis with exhaustive representation and quantification of the uncertainties is indispensable for design optimization.

There have been numerous researches on reliability-based design optimization (RBDO) to obtain a reliable optimum design employing probabilistic quantification of uncertain performance. It can be typically categorized into sampling approaches (Dubourg et al. 2011; Lee et al. 2011a; Zhao et al. 2013; Wang and Wang 2014) and most probable point (MPP)-based approaches (Tu et al. 1999; Youn et al. 2003; Li et al. 2015; Choi et al. 2018). These extensive RBDO researches have generally assumed that statistical information of input random variables is fully accessible. However,

Responsible Editor: Nam Ho Kim

✉ 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

obtaining true input statistical information from a set of finite samples is difficult due to the limitation of resource for testing samples to quantify the variability. Therefore, inaccurate input model is a much practical condition in real engineering problems. Epistemic uncertainty in simulation-based design optimization can be comprehensively defined including simulation model uncertainty due to the uncertainty of model parameters and model bias (Der Kiureghian and Ditlevsen 2009; Liang and Mahadevan 2011; Nannapaneni and Mahadevan 2016; Moon et al. 2017; Peng et al. 2017; Bae et al. 2017, 2018); hence, the uncertainty of input statistical models will be mainly treated in this paper.

There have been a few researches to quantify the uncertainty of input random variables induced by insufficient input dataset. In the early stage, possibility-based approaches (Du et al. 2006a, b; Lee et al. 2013) and interval uncertainty approaches (Rao and Cao 2002; Sankararaman and Mahadevan 2011; Yoo and Lee 2014; Muscolino et al. 2016) have been adopted, but the Bayesian approach is a more direct and intuitive way to use insufficient input data statistically. Gunawan and Papalambros (2006) assumed that the variability of reliability could be exhibited as the beta distribution and performed RBDO while maximizing the confidence of reliability. Youn and Wang (2008) proposed a similar approach focusing on the extreme case of the beta distribution to make a sufficiency, in which Bayesian reliability has to be smaller than exact reliability. On the other hand, the uncertainty of the input random variable can be handled by adjusting the standard deviation and correlation coefficient (Noh et al. 2011a, b). In addition, likelihood-based approaches have been proposed that the distribution of the parameters can be quantified through the maximum likelihood estimation (MLE) and Gaussian process interpolation (Sankararaman and Mahadevan 2011, 2013a, b). Instead of Bayes' theorem, the researches utilizing the bootstrap method without the assumption that data follow certain distribution have been proposed (Picheny et al. 2010; Noh et al. 2011b). Ito et al. (2018) propose a conservative reliability index (CRI) decomposed into the aleatory part and the epistemic part for Gaussian input distribution. Meng et al. (2015b) and Hao et al. (2017) developed non-probabilistic RBDO to treat lack of input data by employing a convex model to describe uncertain-but-bounded parameters rather than the stochastic model. Recently, Cho et al. (2016a) and Moon et al. (2018) proposed conservative RBDO (CRBDO) to fully model the probability of reliability in the presence of the uncertainty of input distribution types and parameters. Employing Bayes' theorem and Monte Carlo simulation (MCS) (Rubinstein and Kroese 2016), accurate confidence of reliability under given dataset can be obtained. However, estimating a cumulative distribution function (CDF) of reliability is very computationally demanding.

Therefore, the motivation of our research is to reduce the excessive amount of function evaluations to obtain the confidence of reliability. It is noted that the confidence of reliability has to be assessed for design optimization rather than the whole CDF of reliability. In this paper, we propose a reliability measure approach (RMA) and confidence-based design optimization (CBDO) which have constraints for the confidence of reliability. The reliability will be treated as a function of input distribution types and parameters. Due to the discreteness of input distribution type, it is divided into two approaches depending on prior knowledge of input distribution types, and kernel density estimation (KDE) (Silverman 2018) is used to quantify the variability of a random variable when its distribution type is totally unknown. Once the input distribution types are given or identified through any model identification method (Kang et al. 2016), the probabilistic domain that consists of input distribution parameters can be defined. The process has been inspired from performance measure approach (PMA) (Tu et al. 1999), but random design variables, performance functions, and reliability are replaced with input distribution parameters, reliability, and confidence of reliability, respectively. In other words, the input distribution parameter space is regarded as the upper-level space of input random variables where RMA is performed. The proposed method handles the confidence of reliability by searching MPP in the probabilistic space denoted as P-space transformed from input distribution parameter space. Consequently, the double-loop MCS can be reduced to a nested RMA-MCS loop leading to a significant reduction of function evaluations for MCS. Moreover, sensitivity analysis in P-space is derived through the first-order score functions that are an essential concept to facilitate stochastic sensitivity analysis in sampling-based RBDO (Lee et al. 2011b) and explicit expression of uncertainties.

The paper is organized as follows. PMA for RBDO and confidence of reliability in CRBDO are explained in Section 2. In Section 3, general RMA for CBDO and stochastic sensitivity analysis is formulated, and two approaches depending on the prior knowledge of input distribution types are given in Section 4. Then, the feasibility and effectiveness of the developed sensitivity analysis, RMA, and CBDO are verified in Section 5. The conclusion and future research will be given in Section 6.

2 Overview of reliability analysis and confidence of reliability

2.1 Performance measure approach for reliability-based design optimization

Reliability analysis is to obtain the probability of failure which is the probability that a performance function is

larger than zero since $G(\mathbf{X}) > 0$ means failure in this paper. The probability of failure is in general computed through the multi-dimensional integration as

$$P_F \equiv \Pr[G(\mathbf{X}) > 0] = \int_{G(\mathbf{x}) > 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \tag{1}$$

where $G(\mathbf{X})$ is the performance function, and \mathbf{X} is the vector of random design variables following the joint probability density function (PDF) of $f_{\mathbf{X}}(\mathbf{x})$.

Reliability-based design optimization (RBDO) with probabilistic constraints is formulated to

$$\underset{\mathbf{d}}{\text{minimize}} \quad \text{cost}(\mathbf{d}) \tag{2}$$

$$\text{subject to } \Pr[G_i(\mathbf{X}) > 0] \leq P_{F_i}^{\text{Target}}, \quad i = 1, \dots, nc$$

$$\mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U$$

where \mathbf{X} is the vector of random variables, \mathbf{d} is the design vector defined as the mean of \mathbf{X} , $P_{F_i}^{\text{Target}}$ is the target probability of failure for the i th constraint, and nc is the number of probabilistic constraints. Exact estimation of the probabilistic constraints in (2) defined by (1) is very difficult since it requires a large number of simulations. Therefore, performance measure approach (PMA) using MPP has been proposed (Tu et al. 1999; Youn et al. 2003, 2005; Yang and Yi 2009; Meng et al. 2015a; Keshtegar and Lee 2016). In PMA for RBDO, the MPP has the largest performance function value on the hypersphere with a radius of the given target reliability index β_t . If the performance function value at MPP is smaller than zero, the probabilistic constraint of PMA is satisfied. Thus, MPP in PMA can be obtained by solving the following optimization as

$$\underset{\mathbf{u}}{\text{maximize}} \quad G(\mathbf{x}(\mathbf{u})) \tag{3}$$

$$\text{subject to } \|\mathbf{u}\| = \beta_t$$

$$\text{where } \mathbf{u} = \Phi^{-1}(F_{\mathbf{X}}(\mathbf{x}))$$

where \mathbf{u} is the random variable vector following the standard normal distribution transformed from original X -space by the Rosenblatt transformation (Rosenblatt 1952) where cumulative distribution function (CDF) of \mathbf{X} is $F_{\mathbf{X}}(\mathbf{x})$. MPP, the optimum of (3), is denoted as \mathbf{u}^* . Hence, the probabilistic constraint in (2) can be rewritten using PMA as

$$g(\mathbf{u}^*) \leq 0 \tag{4}$$

which is identical with $G(\mathbf{x}(\mathbf{u}^*)) \leq 0$ as a function at MPP in the original domain.

2.2 Confidence of reliability with insufficient input data

Confidence of reliability is defined as the probability that given target reliability is satisfied at a design point. Uncertainty of reliability is originated from input model uncertainty, distinguished from input variability, which means that insufficient input data induce uncertain input statistical models such as distribution types and parameters. Using Bayes' theorem, joint PDF of the reliability under given dataset $^*\mathbf{x}$ can be obtained as (Cho et al. 2016a; Moon et al. 2018)

$$f(\text{Re}, \zeta, \boldsymbol{\psi} | ^*\mathbf{x}) = f(\text{Re} | \zeta, \boldsymbol{\psi}, ^*\mathbf{x}) P(\zeta | \boldsymbol{\psi}, ^*\mathbf{x}) f(\boldsymbol{\psi} | ^*\mathbf{x}) \tag{5}$$

where ζ and $\boldsymbol{\psi}$ are input distribution types and parameters, respectively. Since the distribution types are not continuous variables, the probability mass function (PMF) obtained from Bayes' theorem is used in (5). Therefore, the confidence level of the given target reliability is obtained using CDF of the reliability as (Cho et al. 2016a)

$$\text{CL}(\text{Re} | ^*\mathbf{x}) \equiv 1 - F_{\text{Re}}(\text{Re} | ^*\mathbf{x}) = 1 - \int_0^{\text{Re}} \int_{\Omega_{\zeta}} \int_{\Omega_{\boldsymbol{\psi}}} f(\phi, \zeta, \boldsymbol{\psi} | ^*\mathbf{x}) d\boldsymbol{\psi} d\zeta d\phi \tag{6}$$

where ϕ is a variable for reliability bounded from 0 to 1.

The main difficulty to evaluate (6) is its large computational demands since double-loop MCS with a large number of samples has to be used: inner-loop MCS for reliability estimation and outer-loop MCS for estimation of confidence of reliability caused by the uncertainty of input distribution types and parameters. Using the double-loop MCS, CDF of the reliability in (6) is expressed as (Cho et al. 2016a)

$$F_{\text{Re}}(\text{Re} | ^*\mathbf{x}) \cong \frac{1}{NMCS_{\mathbf{Z}} NMCS_{\boldsymbol{\Psi}}} \sum_{n=1}^{NMCS_{\boldsymbol{\Psi}}} \sum_{m=1}^{NMCS_{\mathbf{Z}}} I_{[0, \text{Re}]}(\text{Re}(\zeta^{(m)}, \boldsymbol{\psi}^{(n)})) \tag{7}$$

where $I_{[0, \text{Re}]}(\phi)$ is an indicator function in which the value is 1 if ϕ is between 0 and Re , and 0 otherwise. It should be noted that the number of samples for the double-loop MCS is $NMCS_{\text{Re}} \times NMCS_{\mathbf{Z}} \times NMCS_{\boldsymbol{\Psi}}$ where $NMCS_{\text{Re}}$, $NMCS_{\mathbf{Z}}$, and $NMCS_{\boldsymbol{\Psi}}$ are the number of samples for estimation of $\text{Re}(\zeta^{(m)}, \boldsymbol{\psi}^{(n)})$, input distribution types, and input distribution parameters, respectively. Reliability is a function of input distribution types and parameters using indicator function due to the insufficient input dataset. So, the reliability is estimated through the sampling method as (Rubinstein and Kroese 2016)

$$\text{Re}(\zeta, \boldsymbol{\psi}) = \int_{\mathbb{R}^N} I(G(\mathbf{x})) f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\psi}, \zeta) d\mathbf{x} \tag{8}$$

where $\boldsymbol{\psi}$ is the input distribution parameter vector, ζ is

the input distribution types, and $I(G(\mathbf{x}))$ is the indicator function defined as

$$I(G(\mathbf{x})) = \begin{cases} 1, & G(\mathbf{x}) \leq 0 \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

2.3 Conservative reliability-based design optimization

General formulation of CRBDO considering the confidence of reliability as constraints can be formulated as

$$\begin{aligned} & \text{given } *x \\ & \text{minimize } \underset{\mathbf{d}}{\text{cost}(\mathbf{d})} \\ & \text{subject to } \Pr\left[\Pr[G_i(\mathbf{X}) \leq 0] \geq \text{Re}_i^{\text{Target}}\right] \geq \text{CL}_i^{\text{Target}} \text{ for } i = 1, 2, \dots, nc \end{aligned} \quad (10)$$

which uses (7) to estimate the confidence of reliability at a design point under given dataset $*x$. As can be seen from (10), CRBDO yields an optimum considering epistemic uncertainty caused by insufficient input data but requires a huge amount of function evaluations. Uncertainty quantification of input distribution parameters and types used to evaluate the confidence of reliability in (10) will be explained in the next section.

3 Confidence-based design optimization using a reliability measure approach

In this section, confidence-based design optimization (CBDO) using a reliability measure approach (RMA) is proposed to resolve the efficiency issue of CRBDO. Decomposition of input dataset can be referred to Section 2.2 of the literature (Cho et al. 2016a). An underlying condition in CBDO is that there are random variables in which statistical information is not given and should be estimated using only a few samples. Accordingly, the confidence of reliability will be mainly treated as a probabilistic constraint instead of reliability used in conventional RBDO since the reliability of a system has also uncertainty propagated from the uncertainty on the input statistical model.

There are two sources of uncertainty in the statistical model of input random variable: input distribution types and parameters. Since the input distribution type is not a continuous random variable, the proposed method takes variability of input distribution parameters into consideration in the probabilistic domain to obtain confidence of reliability under two circumstances: 1) input distribution types are known, but input distribution parameters are unknown and 2) both input distribution types and parameters are unknown. In this study, uncertainties of input distribution parameters are quantified

based on the normality condition, and unknown distribution types are described by kernel density estimation (KDE).

Under the aforementioned configurations, the confidence of reliability defined as a probability that the reliability of a performance function is larger than the user-specified target reliability needs to be computed in CBDO. However, it is very difficult to be calculated even if a surrogate model is used because double-loop MCS is necessary for repetitive reliability computations with respect to uncertain parameters. Therefore, RMA is proposed to efficiently estimate the confidence of reliability in parameter space by utilizing the framework of PMA. Reliability and input distribution parameters correspond to a performance function and random variables in PMA, respectively. Therefore, RMA finds MPP for the reliability with the target confidence in the input distribution parameter space and thus requires several MCSs only for searching MPP, unlike CRBDO. After finding the MPP, the confidence of reliability is obtained using a linear approximation of the reliability function in parameter space. Consequently, RMA can significantly reduce the computational burden in CRBDO since it estimates the confidence of reliability in an analytic way rather than sampling methods.

The uncertainty quantification of input distribution model is presented in Section 3.1. The formulation of the proposed RMA in P-space and stochastic sensitivity analysis for reliability based on uncertainty quantifications of input distribution parameters are explained in Section 3.2. Two cases for known and unknown input distribution types are discussed in Section 4.

3.1 Uncertainty of input distribution model

Variabilities of input distribution parameters that arise due to insufficient input data have to be explicitly quantified to define probabilistic space for reliability estimation. Under the normality assumption described in the literature (Gelman et al. 2013; Cho et al. 2016a), variabilities of mean and variance are expressed as normal and inverse-gamma distributions, respectively. The normality assumption does not mean that the true distribution type is a normal distribution, but it is the intermediate assumption for uncertainty quantifications of input distribution parameters without causing a significant error. Uncertainties of variance and mean of the i th random variable can be described as

$$\sigma_i^2 | *x_i \sim \text{IG}\left(\frac{ND-1}{2}, \frac{(ND-1)s_i^2}{2}\right) \quad (11)$$

and

$$\mu_i | \sigma_i, *x_i \sim \mathcal{N}\left(*\bar{x}_i, \sqrt{\frac{\sigma_i^2}{ND}}\right), \quad (12)$$

respectively, where $*\mathbf{x}_i$ is given input data for the i th random variable; ND is the number of input data; s_i^2 is the sample variance of the i th random variable; and $*\bar{\mathbf{x}}_i$ is the sample mean of the i th random variable. It is noted that the normal distribution of the mean in (12) is conditional to the variance in (11). The sample mean in (12) will change during design optimization since it is used as a design vector \mathbf{d} in (10). In other words, the dataset is used to identify the initial design and sample variance which is assumed to be constant during the optimization, and design movement can be achieved by shifting the data. Obviously, the sample mean of random parameter is invariant during the optimization

The PDF of reliability under given input distribution parameters and types becomes a Dirac delta function expressed by

$$f(\text{Re}|\boldsymbol{\psi}, \zeta, *\mathbf{x}) = \delta(\text{Re}-\text{Re}(\boldsymbol{\psi}, \zeta)) \tag{13}$$

where $\boldsymbol{\psi}$ is the input distribution parameters, and ζ is the input distribution types that can be either specific parametric distribution or nonparametric KDE depending on prior knowledge as will be explained in Section 4. Consequently, when $\boldsymbol{\psi}$ and ζ are given, the reliability is considered as a deterministic value obtained from MCS, so that PDF of reliability has an infinity value only at the deterministic reliability, $\text{Re}(\boldsymbol{\psi}, \zeta)$.

3.2 Reliability measure approach

In this section, RMA to explore MPP in the standard normal parameter space (P-space) transformed from input distribution parameter space is explained. In RMA, MPP search is performed in P-space to find the smallest reliability on the given hypersphere with a radius β_{CL} derived from the target confidence level, $\text{CL}_{\text{Target}}$. Note that the smallest reliability should be explored unlike the maximization problem in (3). Figure 1

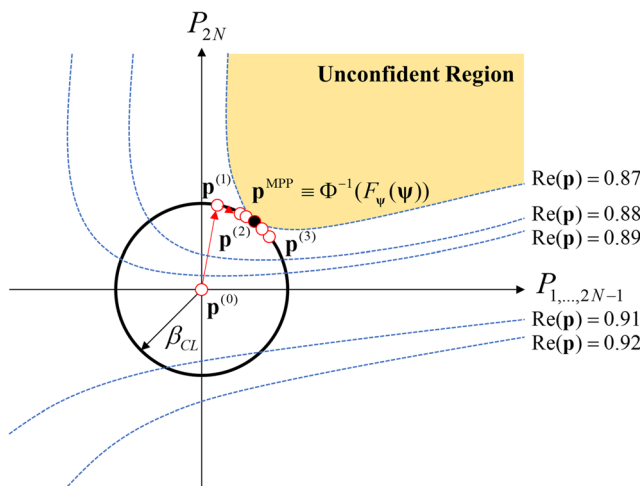


Fig. 1 Concept of RMA in P-space

illustrates the concept of RMA in P-space and design history for searching MPP where N is the number of random variables with insufficient samples. The initial point is set to origin in the standard normal P-space in general, and optimizer finds the next MPP candidate point based on the gradient of reliability in P-space which can be obtained from sensitivity analysis of reliability with respect to each parameter employing the first-order score functions. Once the MPP is identified as black filled circle in Fig. 1, the reliability at the MPP calculated from transformed input distribution parameters denoted as $\boldsymbol{\psi}$ including mean and variance is used to judge the violation of the confidence constraint. In Fig. 1, the reliability at the MPP is estimated as 0.87 that has to be larger than the target reliability. If the target confidence is set to 0.87, RMA for CBDO will decide that the current design point satisfies the confidence constraint even though the yellow region is the true unconfident region where reliability is less than the target reliability. Details of formulation and sensitivity analysis will be explained in Sections 3.2.1, 3.2.2, and 3.2.3.

3.2.1 Formulation of RMA

The optimization to find MPP in P-space can be formulated to

$$\underset{\mathbf{p}}{\text{minimize}} \quad \text{Re}(G(\mathbf{X})|\boldsymbol{\mu}, \boldsymbol{\sigma}^2) \tag{14}$$

$$\text{subject to } \|\mathbf{p}\| = \Phi^{-1}(\text{CL}_{\text{Target}})$$

$$\text{where } \text{Re}(G(\mathbf{X})|\boldsymbol{\mu}, \boldsymbol{\sigma}^2) = \int_{I_{\Omega_R}(G(\mathbf{x}))} f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\sigma}^2) d\mathbf{x} \cong \frac{1}{N_{MCS}} \sum_{j=1}^{N_{MCS}} I_{\Omega_R}(G(\mathbf{x}^{(j)}))$$

$$I_{\Omega_R}(G(\mathbf{x})) = \begin{cases} 1, & G(\mathbf{x}) \leq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\Phi(p_{2i-1}) = F_{\sigma_i^2}(\sigma_i^2 | *\mathbf{x}_i)$$

$$\Phi(p_{2i}) = F_{\mu_i}(\mu_i | \sigma_i^2, *\mathbf{x}_i) \quad \text{for } i = 1, 2, \dots, N$$

In Eq. (14) $\text{Re}(G(\mathbf{X})|\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$ is reliability for the given performance function $G(\mathbf{X})$ and realizations of input distribution parameters. It means that uncertainty of performance function propagated from a random vector \mathbf{X} is conditional to input distribution parameters. $\text{CL}_{\text{Target}}$ is the target confidence level 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 reliability estimation, and N is the number of random variables which are assumed to be independent in this study. It can be seen that two input distribution parameters for the i th random variable following $F_{\mu_i}(\mu_i | \sigma_i^2, *\mathbf{x}_i)$ and $F_{\sigma_i^2}(\sigma_i^2 | *\mathbf{x}_i)$ are transformed to P-space following the standard normal distribution. The MPP that has minimal reliability as a function of input distribution parameters in P-space would be explored, and the number of random variables in P-space becomes $2N$ since a two-parameter distribution is assumed for the input random variables.

3.2.2 Stochastic sensitivity analysis in RMA

Sensitivity analysis of reliability with respect to mean and variance is essential for efficient and accurate gradient-based optimization. The stochastic sensitivity analysis has been discussed in the previous studies to develop the sampling-based RBDO (Lee et al. 2011a, b; Cho et al. 2016b). The sensitivity analysis with respect to the i th mean can be described using the first-order score function and chain rule as

$$\begin{aligned} \frac{\partial \text{Re}}{\partial P_{\mu_i}} &= \frac{\partial \text{Re}}{\partial \mu_i} \frac{\partial \mu_i}{\partial P_{\mu_i}} \\ &= \int I_{\Omega_R}(G(\mathbf{x})) \frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2)}{\partial \mu_i} f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2) d\mathbf{x} \times \frac{\partial \mu_i}{\partial P_{\mu_i}} \end{aligned} \tag{15}$$

where the PDF of mean is given in (12). The derivative of mean with respect to a corresponding P_{μ_i} can be obtained through the following relationship as

$$\frac{\partial \mu_i}{\partial P_{\mu_i}} = \frac{\partial \Phi(P_{\mu_i})}{\partial P_{\mu_i}} / \frac{\partial F_{\mu_i}(\mu_i)}{\partial \mu_i} = \frac{\phi(P_{\mu_i})}{f_{\mu_i}(\mu_i)} \tag{16}$$

where $F_{\mu_i}(\mu_i) = \Phi(P_{\mu_i})$ and P_{μ_i} is a random variable following the standard normal distribution transformed from the mean of i -th random variable denoted as μ_i . ϕ and Φ are PDF and CDF of the standard normal distribution, respectively. On the other hand, the sensitivity analysis with respect to the variance of i -th random variable, which seems to be more complicated since the mean is conditional to the variance, is derived as

$$\begin{aligned} \frac{\partial \text{Re}}{\partial P_{\sigma_i^2}} &= \frac{\partial \text{Re}}{\partial \sigma_i^2} \frac{\partial \sigma_i^2}{\partial P_{\sigma_i^2}} \\ &= \int I_{\Omega_R}(G(\mathbf{x})) \frac{d \ln f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2)}{d \sigma_i^2} f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2) d\mathbf{x} \times \frac{\partial \sigma_i^2}{\partial P_{\sigma_i^2}} \\ &= \int I_{\Omega_R}(G(\mathbf{x})) \left(\frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2)}{\partial \sigma_i^2} + \frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2)}{\partial \mu_i} \frac{d \mu_i}{d \sigma_i^2} \right) \\ &\quad f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2) d\mathbf{x} \times \frac{\phi(P_{\sigma_i^2})}{f_{\sigma_i^2}(\sigma_i^2)} \end{aligned} \tag{17}$$

where $P_{\sigma_i^2}$ is a random variable following the standard normal distribution corresponding to σ_i^2 . The PDF of variance is given in (11). The derivatives of the random variable with respect to distribution parameters can be shown in the literature using the relationship of the Rosenblatt transformation written as (Cho et al. 2017)

$$\frac{dx_i}{dp_i} = -\frac{1}{f_{X_i}(x_i; a_i, b_i)} \frac{\partial}{\partial p_i} F_{X_i}(x_i; a_i, b_i) \tag{18}$$

where p_i corresponds to either mean or variance of i -th random variable, and a_i and b_i are general distribution parameters

as functions of mean and variance. Therefore, the derivative of mean with respect to variance in (17) can be derived as

$$\begin{aligned} \frac{d \mu_i}{d \sigma_i^2} &= -\frac{1}{f_{\mu_i}(\mu_i; \bar{x}_i, \sqrt{\sigma_i^2/ND})} \times \frac{-(\mu_i - \bar{x}_i)}{2\sigma_i^3/\sqrt{ND}} \times \phi\left(\frac{\mu_i - \bar{x}_i}{\sqrt{\sigma_i^2/ND}}\right) \\ &= \frac{\mu_i - \bar{x}_i}{2\sigma_i^2} \end{aligned} \tag{19}$$

where \bar{x}_i is the sample mean of data which equivalent to design point of i th random variable. In addition, the first-order score functions of various parametric distribution, $\frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2)}{\partial \mu_i}$ and $\frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2)}{\partial \sigma_i^2}$ used in (15–17), are derived analytically in the literature (Lee et al. 2011a, b; Cho et al. 2016b).

Searching MPP with any algorithms of PMA demands the gradient vector at the current point, and thus the gradient vector of reliability with respect to input distribution parameters is necessary. Employing developed sensitivity analysis in this section, searching MPP in P-space can be performed without any additional MCSs to obtain the gradient vector. Therefore, only one MCS can provide the gradient vector to any optimizer in searching MPP.

3.2.3 Sensitivity analysis in CBDO

Like conventional PMA, sensitivity analysis with respect to a design variable vector—mean vector of random variables—in CBDO can be described as (Lee et al. 2010)

$$\begin{aligned} \left. \frac{\partial \text{Re}}{\partial d_i} \right|_{\mathbf{p}=\mathbf{p}^*} &= \left. \frac{\partial \text{Re}}{\partial \mu_i} \right|_{\mathbf{p}=\mathbf{p}^*} \\ &= \int I_{\Omega_R}(G(\mathbf{x})) \frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2)}{\partial \mu_i} f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2) d\mathbf{x} \end{aligned} \tag{20}$$

Equation (20) can be easily obtained since $\frac{\partial \ln f_{\mathbf{x}}(\mathbf{x}; \mu_i)}{\partial \mu_i}$ is already used in (15) and (17). Unlike sensitivity analysis in searching MPP described in Section 3.2.2, sensitivity analysis of reliability with respect to mean vector is used to find the optimal design point rather than MPP. In other words, CBDO has a double-loop algorithm composed of MPP search as an inner loop and optimal design search as an outer loop. Sensitivity analyses in Sections 3.2.2 and 3.2.3 are for the inner loop and outer loop, respectively. Note that (20) is equivalent to (15), except that the probabilistic space has to be transformed to P-space.

Using RMA, MPP in P-space is obtained, and reliability at the MPP is used as an index of violation under the given confidence of a reliability constraint. Since the function approximation method is the same as the conventional MPP-based reliability analysis, the second-order reliability method (SORM) (Lee et al. 2012; Park and Lee 2018), and MPP-

based dimension reduction method (DRM) (Rahman and Xu 2004; Jung et al. 2019) can be utilized for more accurate confidence estimation.

4 Two approaches for input distribution type uncertainty

As mentioned earlier, there are two sources of uncertainties in input statistical models: input distribution parameters and distribution types. In Section 3, only uncertainties of input distribution parameters are treated for RMA and their sensitivity analysis. In this section, two approaches for input distribution type uncertainties—known input distribution type and unknown input distribution type—are explained for estimation of reliability and its sensitivity for CBDO.

4.1 Known input distribution type

Knowing input distribution types means that a designer has set the distribution types based on any prior knowledge or has figured out the most likely distribution type. For instance, model selection methods can determine the best distribution for the given data among multiple candidate distributions, and acceptance or rejection of the hypothesis with respect to the adequacy of given distribution type can be determined through goodness-of-fit (GOF) tests. Thus, the designer is able to use the model identification methods before the proposed CBDO as a preprocessing if there is no particular clue about the distribution type. Since the identification of a distribution type is beyond the purpose of this research, readers can refer to recent research on distribution type identification (Kang et al. 2016). The reason for separating uncertainty of a distribution type and its parameters is that it is difficult to include discrete variables for MPP-based approaches, and even if it is possible, it requires additional assumption and can become too complicated since two uncertainties are correlated.

When input distribution types are known as parametric distributions, the input joint PDF can be expressed as

$$f_{\mathbf{x}}(\mathbf{x}; \Psi) = \prod_{i=1}^N f_{X_i}(x_i; \Psi) \tag{21}$$

Since input random variables are assumed to be independent in this study. The PDF of the *i*th random variable can be described as (Cho et al. 2016b)

$$f_{X_i}(x_i; \Psi) = f_{X_i}(x_i; \mu_i, \sigma_i^2) \tag{22}$$

Assuming that the PDF is a two-parameter distribution with parameters of $\mu_i(a_i, b_i)$ and $\sigma_i(a_i, b_i)$ as a function of two general distribution parameters. Under this configuration, RMA and its sensitivity analysis can share the sampling-based RBDO framework, and the relationship between mean and variance and

general distribution parameters can be found in the literature (Lee et al. 2011a, b; Cho et al. 2016b). Note that any two-distribution parameters need to be expressed as mean and variance and vice versa to utilize the framework of RMA. Furthermore, the first-order score functions with respect to mean and variance for various parametric distributions can be found.

Once input distribution types are decided, epistemic uncertainty induced by an insufficient number of data is merely quantified as uncertainties of input distribution parameters. The reliability of target performance is a function of input distribution parameters so that the proposed MPP-search in P-space is available. The sensitivity analysis of reliability in (15) and (17) can be performed since $f_{\mathbf{x}}(\mathbf{x}; \mu_i, \sigma_i^2)$ is given as an explicit function of mean and variance. However, when input distribution types are still unknown, PDFs should be defined using KDE prior to RMA, which will be presented in the next section.

4.2 Unknown input distribution type

4.2.1 Uncertainty quantification using kernel density estimation

When input distribution types are unknown, it is necessary to describe $f_{\mathbf{x}}(\mathbf{x}; \Psi)$ without using any parametric distribution. In this research, kernel density estimation (KDE) is implemented for the purpose. KDE under given input data is described as (Silverman 2018)

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{X_i - x}{h}\right) \tag{23}$$

where *n* is the number of samples; X_1, X_2, \dots, X_n are independent and identically distributed random samples; $k(u)$ is a kernel function satisfying $\int_{-\infty}^{\infty} k(u)du = 1$; *h* is a smoothing parameter of KDE. In this paper, the standard normal density function is employed for the kernel function.

It should be noted that the shape of the KDE expression in (23) does not change during design optimization. Only mean, which is the sample mean initially, changes in the optimization process. Thus, during the design optimization, a shifted KDE of the independent *j*th random design variable with *n* samples is used as an input distribution given by

$$\hat{f}_{X_j}(x; {}^* \mathbf{x}_j, \mu_j) = \frac{1}{nh_j} \sum_{i=1}^n k\left(\frac{s_i({}^* \mathbf{x}_j, \mu_j) - x}{h_j}\right) \tag{24}$$

where $s_i({}^* \mathbf{x}_j, \mu_j) = {}^* x_{j,i} - {}^* \bar{x}_j + \mu_j$ is the *i*th shifted sample of the *j*th random variable, and μ_j is the given mean of the *j*th random variable. On the other hand, the uncertainty of variance cannot be directly reflected in (24) since the variance of KDE is always larger than the sample variance

(Hansen 2009). Therefore, using the same normality condition used in Section 3.1, the bandwidth in (24) can be determined using the normal reference rule as (Silverman 2018)

$$h_j = \left(\frac{4\sigma_j^5}{3n} \right)^{\frac{1}{5}} \approx 1.06\sigma_j n^{-\frac{1}{5}} \tag{25}$$

where σ_j^2 is the given variance of the j th random variable. The bandwidth of normal reference rule is to choose the optimal bandwidth that minimizes the mean integrated squared error when underlying density function is a normal distribution. Despite the normality assumption, (24) can express a plausible PDF for any input distribution. Through (24) and (25), KDE can reflect the given mean and variance to compute the reliability in P-space. Note that given mean and variance as μ_j and σ_j^2 indicate the input distribution parameters to specify the uncertain input distribution since mean and variance are changed during MPP search in P-space, and KDE has to reflect to difference mean and variance following the current MPP candidate.

4.2.2 Stochastic sensitivity analysis for kernel density estimation

To utilize the stochastic sensitivity analysis in Section 3.2.2 for unknown input distribution cases, the first-order score function for KDE needs to be obtained. Using (24) and (25), the first-order score function with respect to mean for KDE of the j th independent random variable is derived as

$$\begin{aligned} \frac{\partial \ln \hat{f}_{X_j}(x; h)}{\partial \mu_j} &= \frac{\partial \ln \hat{f}_{X_j}}{\partial x_{j,1}} \frac{\partial^* x_{j,1}}{\partial \mu_j} + \dots + \frac{\partial \ln \hat{f}_{X_j}}{\partial x_{j,n}} \frac{\partial^* x_{j,n}}{\partial \mu_j} \\ &= \sum_{i=1}^n \frac{x^*_{j,i}}{\sqrt{2\pi n h_j^3} \hat{f}_{X_j}(x)} e^{-\frac{1}{2} \left(\frac{x^*_{j,i}}{h_j} \right)^2} \end{aligned} \tag{26}$$

where $x^*_{j,i}$ is i th sample of j th random variable, $\hat{f}_{X_j}(x) = \frac{1}{nh_j} \sum_{i=1}^n \phi\left(\frac{x^*_{j,i}}{h_j}\right)$, and $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ which is the standard normal density function. Similarly, the first-order score function with respect to the variance for KDE of the j th independent random variable is derived as

$$\begin{aligned} \frac{\partial \ln \hat{f}_{X_j}(x; h)}{\partial \sigma_j^2} &= \frac{\partial \ln \hat{f}_{X_j}}{\partial h_j} \frac{\partial h_j}{\partial \sigma_j^2} \\ &= -\frac{1}{h_j} + \frac{\sum_{i=1}^n (x^*_{j,i})^2 \phi\left(\frac{x^*_{j,i}}{h_j}\right)}{h_j^3 \sum_{i=1}^n \phi\left(\frac{x^*_{j,i}}{h_j}\right)} \frac{(1.06n^{-1/5})^2}{2h_j} \end{aligned} \tag{27}$$

Hence, (26) and (27) are substituted in (15–17) and (20) for the stochastic sensitivity analysis for CBDO when there is no prior knowledge on input distribution types.

4.3 Flowchart and practical issues

Figure 2 shows the flowchart of the proposed CBDO with RMA. Starting at the initial design point, variabilities of mean and variance are quantified as a normal and inverse-gamma distribution using the given input dataset. Before MPP search, input distribution types have to be specified according to which method is used. If the distribution types are parametrically identified, the parametric distributions are directly used for random sampling with respect to realizations of mean and variance. Otherwise, KDE is used to describe input distributions based on the given input dataset. To check violation of confidence constraints, uncertain input distribution parameters are transformed to P-space to find MPP using the stochastic sensitivity analysis for

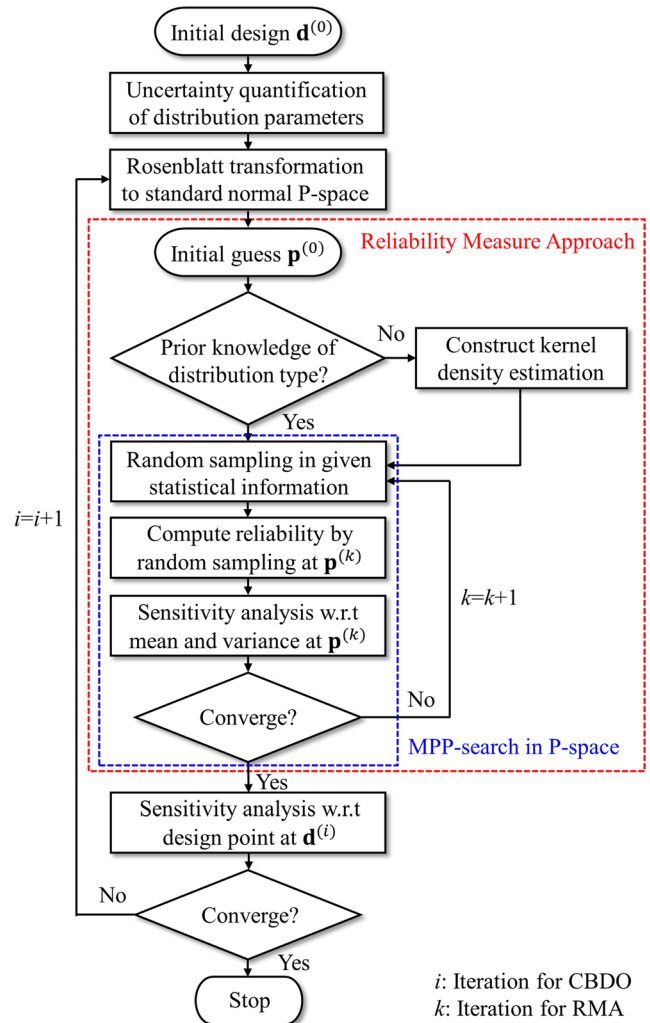


Fig. 2 Flowchart for the proposed CBDO with RMA

Table 1 Sensitivity analysis results using known parametric distribution

| | Normal | Lognormal | Weibull | Gumbel |
|---------------------|-----------------|-----------------|-----------------|-----------------|
| FDM - mean | 0.0649 | 0.0313 | 0.0663 | 0.0289 |
| Proposed - mean | 0.0655 (99.09%) | 0.0315 (99.37%) | 0.0659 (99.40%) | 0.0277 (95.85%) |
| FDM - variance | 0.0723 | 0.0240 | 0.0248 | 0.0128 |
| Proposed - variance | 0.0731 (98.90%) | 0.0238 (99.17%) | 0.0250 (99.20%) | 0.0118 (92.19%) |

reliability. After finding the MPP, the current design point is moved to the next point using any nonlinear optimizer (e.g., sequential quadratic programming).

There are two remarks compared to common MPP searches used in PMA. Firstly, since reliability is always bounded between 0 and 1, sensitivity may become zero in case of totally violated or safe constraints. Therefore, it is recommended to check reliability values before MPP search in P-space to prevent numerical instability due to zero sensitivity. Secondly, characteristics of reliability with respect to mean and variance may result in multiple MPPs. This happens when a performance function has multiple failure regions in the probabilistic domain. It can occur in conventional reliability analysis as well, but may be more frequent in RMA since reliability is a much more complex function with respect to input distribution parameters. The issue in the second remark will be discussed in further works.

5 Numerical examples

Two mathematical examples and one engineering example are introduced in this section for demonstration of the proposed CBDO with RMA. The proposed sensitivity analysis with respect to each input distribution parameter is validated, and MPP search in P-space is visualized to show its feasibility by checking contours of reliability functions for the 1-D mathematical function. In the 2-D mathematical example, the results of proposed CBDO under various circumstances and its effectiveness have been analyzed. Finally, the 11-D side impact example is used to test the applicability of the proposed CBDO to practical and complex engineering problems.

Table 2 Sensitivity analysis results using KDE

| | Normal | Lognormal | Weibull | Gumbel |
|---------------------|-----------------|-----------------|-----------------|-----------------|
| FDM - mean | 0.0621 | 0.0368 | 0.0327 | 0.0152 |
| Proposed - mean | 0.0605 (97.43%) | 0.0359 (97.56%) | 0.0314 (96.03%) | 0.0149 (98.03%) |
| FDM - variance | 0.0954 | 0.0353 | 0.0222 | 0.0103 |
| Proposed - variance | 0.0937 (98.22%) | 0.0355 (99.44%) | 0.0225 (98.67%) | 0.0106 (98.06%) |

5.1 One-dimensional mathematical function for RMA

To demonstrate the proposed RMA and its sensitivity analysis, 1-D mathematical performance function given by

$$G(X) = -0.35 + X \sin(X) - 0.5X^2 \quad (28)$$

is introduced. Even though the performance function has only one random design variable, its probabilistic space of input distribution parameters is expressed in 2-D space since a two-parameter distribution will be used for the random variable X . There are 10 input data for X randomly drawn from normal, lognormal, Weibull, and Gumbel distributions where the true variance is set to 0.5^2 , respectively.

5.1.1 Verification of sensitivity analysis in RMA

The proposed stochastic sensitivity analysis with respect to mean and variance in P-space is verified in this section by comparing with the finite difference method (FDM). Although FDM could be used for sensitivity analysis, it requires additional function evaluations, and proper perturbation is unknown in general due to truncation and round-off error in computation. Therefore, analytical sensitivity analysis is critical in design optimization in terms of both efficiency and accuracy. The FDM results are merely used as a reference to validate the derivation on analytical sensitivity analysis. For the test, a design point of $\mathbf{p}^* = \{1.1897, 1.1357\}^T$ which is the MPP in P-space with given dataset drawn from the normal distribution is utilized. Sensitivities of reliability with respect to distribution parameters for parametric distributions and nonparametric distribution given in (15) and (17) and FDM sensitivities with 0.5% perturbation are listed in Tables 1 and 2. Agreements between the proposed and FDM sensitivities are shown in the parenthesis. From the tables, it is shown that

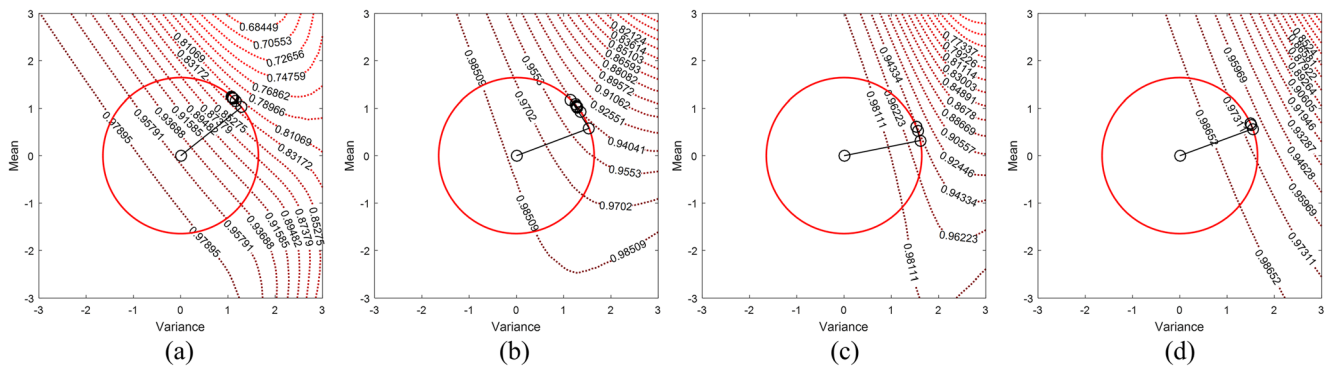


Fig. 3 Contour of reliability and MPP-search history in P-space with samples drawn from normal (a), lognormal (b), Weibull (c), and Gumbel distribution (d)

accurate sensitivities in P-space are obtained even when a distribution of a random variable is expressed using KDE.

5.1.2 Results of RMA

In this section, the proposed RMA is explained using the same 1-D performance function. Figure 3 shows the contours of reliability and hyperspheres for 95% target confidence in P-space assuming input distributions are known. Figure 3 also shows MPP search history which is obtained by hybrid mean value method (HMV) (Youn et al. 2003), and all the reliability values at converged points are the smallest on the target confidence circle. Hence, it can be seen that the MPP search has been successfully performed based on accurate sensitivity information. Each axis in Fig. 3 represents input distribution parameters transformed to P-space.

In case of no prior knowledge on input distribution types, contours of reliability are shown in Fig. 4 where the dataset of the previous case is reused for KDE. Even though the same dataset is used, the variability of random design variables changes resulting in different reliability contours. This is because only 10 samples are used for input statistical modeling.

Results of the proposed RMA for 95% target confidence level are listed in Table 3. For validation, MCS using original parametric distributions is performed at MPP obtained using RMA and agreements between two results are written in the parentheses. Table 3 shows that RMA is more accurate when input distributions are known and KDE shows slightly higher error especially in non-normal cases. It can be seen from Fig. 3 that limit-state reliability functions near MPP seem to be linear so that the RMA results using parametric distributions are accurate. On the other hand, in case of unknown distribution as shown in Fig. 4, it should be addressed that KDE cannot accurately approximate the original input distribution of the random variable with 10 samples only. Therefore, the larger errors in KDE are mainly caused by inexact estimations of the original PDF.

Table 4 shows RMA results when the number of input data is 30. Accuracy is generally improved using more data in two ways. First, increasing the number of data will reduce the variability of each input distribution parameter. Therefore, the limit-state reliability function near MPP in P-space becomes more linear, thereby reducing error due to the

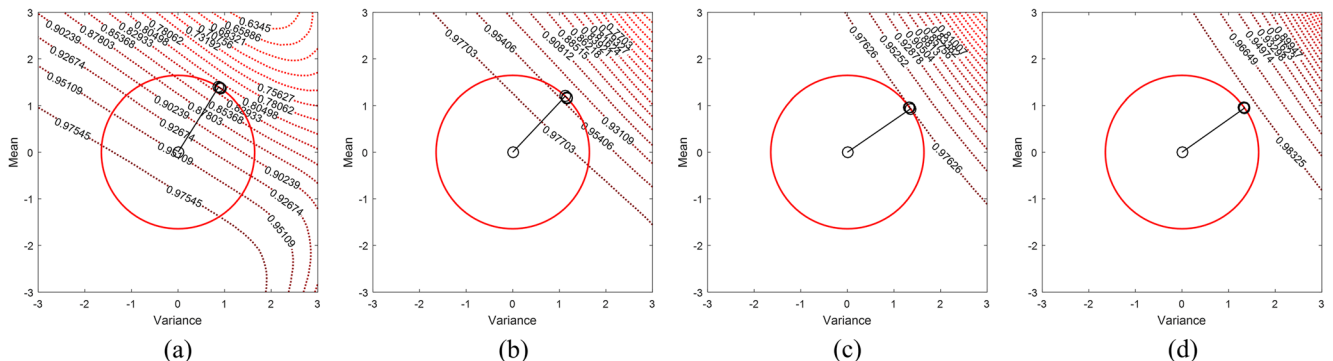


Fig. 4 Contour of reliability and MPP-search history using KDE in P-space with samples drawn from normal (a), lognormal (b), Weibull (c), and Gumbel distribution (d)

Table 3 MCS results at MPP obtained using RMA with a 95% confidence level using 10 data

| True distribution | Known distribution type (%) | KDE (%) |
|-------------------|-----------------------------|---------------|
| Normal | 95.19 (0.20) | 96.01 (0.96) |
| Lognormal | 95.60 (0.63) | 93.52 (1.55) |
| Weibull | 95.25 (0.24) | 85.96 (9.51) |
| Gumbel | 95.05 (0.05) | 72.64 (21.24) |

*Values in parenthesis are discrepancy error with target confidence as 95%

linearization. Second, when input distribution types are unknown, increasing the number of data can reduce error due to inaccurate estimation of input distributions by KDE as well. Therefore, increasing the number of data improves the accuracy of RMA with KDE more as shown in Table 4.

5.2 Two-dimensional confidence-based design optimization

Efficiency and accuracy of the proposed CBDO with RMA are verified through a 2-D mathematical example formulated as

$$\begin{aligned}
 \min \quad & -\frac{(d_1 + d_2 - 10)^2}{30} - \frac{(d_1 - d_2 + 10)^2}{120} \\
 \text{s.t.} \quad & \Pr[\Pr[G_j(\mathbf{X}(\mathbf{d})) \leq 0] \geq 0.9772] \geq 0.9000, j = 1, 2, 3 \\
 & 0 \leq d_i \leq 10, i = 1, 2 \\
 \text{where} \quad & G_1(\mathbf{X}) = 1 - \frac{X_1^2 X_2}{20}, \\
 & G_2(\mathbf{X}) = -1 + (0.9063X_1 + 0.4226X_2 - 6)^2 \\
 & \quad + (0.9063X_1 + 0.4226X_2 - 6)^3 \\
 & \quad - 0.6(0.9063X_1 + 0.4226X_2 - 6)^4 \\
 & \quad - (-0.4226X_1 + 0.9063X_2), \\
 & G_3(\mathbf{X}) = 1 - \frac{80}{X_1^2 + 8X_2 + 5}
 \end{aligned} \tag{29}$$

where the target reliability and target confidence of reliability are set to 97.72% and 90.00%, respectively. In other words, the optimization in Eq. (29) explores a

Table 4 MCS results at MPP obtained using RMA with a 95% confidence level using 30 data

| True distribution | Known distribution type (%) | KDE (%) |
|-------------------|-----------------------------|--------------|
| Normal | 94.74 (0.27) | 92.88 (2.23) |
| Lognormal | 94.98 (0.02) | 98.28 (3.45) |
| Weibull | 95.12 (0.13) | 95.24 (0.25) |
| Gumbel | 94.95 (0.05) | 85.91 (9.51) |

*Values in parenthesis are discrepancy error with target confidence as 95%

conservative optimum satisfying the confidence constraints with 90.00 % target confidence and 97.79 % target reliability. The number of MCS samples for reliability computation is 10^6 , and the number of MCS samples for input distribution parameters in double-loop CRBDO is 10^4 .

The design history of CBDO is illustrated in Fig. 5 when 10 input data are given. For efficiency, the sampling-based RBDO is performed from the initial point, and then CBDO is carried out from the RBDO optimum which is supposed to be near the CBDO optimum compared to arbitrary initial points. It is reasonable that the CBDO optimum is farther from each limit-state function than the RBDO optimum to compensate the input uncertainties. Note that the sampling-based RBDO is performed only with the inherent randomness of input variables using sample mean and variance.

Tables 5, 6, and 7 show design optimization results obtained using the proposed CBDO for known and unknown input distribution cases and conventional double-loop CRBDO. The confidence in each table means exact confidence for each constraint obtained using MCS with true input distributions. When input distribution types are known, the CBDO optimum gradually approaches the double-loop CRBDO optimum as the number of data increases. On the other hand, in case of unknown input distribution type, the CBDO optimum becomes conservative especially for the second constraint as the number of data increases. This is because the reliability for the second constraint obtained using KDE is overestimated due to its nonlinearity and inaccurate density estimation. The number of MCSs in the last columns of Tables 5, 6, and 7 means the number of reliability calculations using MCS. Since the third constraint is perfectly safe during the design optimization, only one MCS is required to check feasibility in each iteration of CBDO. However, in the case of CRBDO, MCS with 10^4 samples for input distribution parameters is performed for a constraint in each iteration. Thus, the total number of MCSs in CRBDO becomes the number of design iterations \times the number of constraints \times the number of MCS samples for input distribution parameters. As a result, the total number of MCSs is reduced from 180,000 in CRBDO to 44 and 36 in CBDO for a case of 100 input samples meaning that efficiency of CBDO is significantly improved maintaining the acceptable level of accuracy compared to CRBDO. Notably, confidence at the optimum is slightly overestimated or underestimated because the proposed method is an MPP-based approximation method inducing approximation error. In this paper, we cannot estimate nonlinearity of reliability with respect to input distribution parameters, but the accuracy will be definitely

Fig. 5 Design history of the proposed CBDO with 10 input data

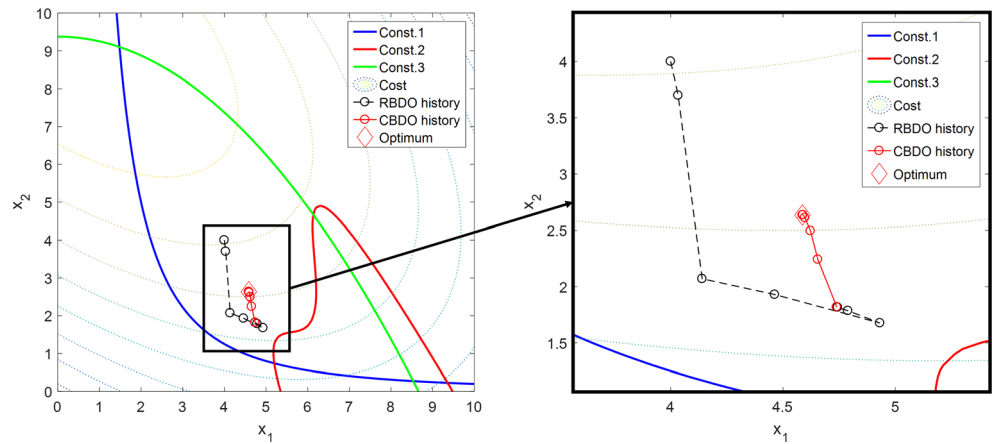


Table 5 Results of CBDO when input distribution types are known

| ND | Optimum | Objective | Confidence ⁽¹⁾ (%) | Confidence ⁽²⁾ (%) | Confidence ⁽³⁾ (%) | # of MCSs |
|-----|-------------------------------|-----------|-------------------------------|-------------------------------|-------------------------------|---|
| 10 | {4.5871, 2.6353} ^T | -1.447 | 84.25 | 87.60 | 99.70 | 46 ⁽¹⁾ + 38 ⁽²⁾ + 17 ⁽³⁾ |
| 30 | {4.6461, 2.2038} ^T | -1.621 | 87.32 | 89.72 | 100.00 | 19 ⁽¹⁾ + 15 ⁽²⁾ + 4 ⁽³⁾ |
| 50 | {4.6516, 2.2416} ^T | -1.605 | 87.77 | 89.86 | 100.00 | 20 ⁽¹⁾ + 19 ⁽²⁾ + 5 ⁽³⁾ |
| 100 | {4.6607, 2.1605} ^T | -1.638 | 88.69 | 90.01 | 100.00 | 20 ⁽¹⁾ + 19 ⁽²⁾ + 5 ⁽³⁾ |

*(1), (2), and (3) mean the first, second, and third constraint, respectively

Table 6 Results of CBDO when input distribution types are KDE

| ND | Optimum | Objective | Confidence ⁽¹⁾ (%) | Confidence ⁽²⁾ (%) | Confidence ⁽³⁾ (%) | # of MCSs |
|-----|-------------------------------|-----------|-------------------------------|-------------------------------|-------------------------------|--|
| 10 | {4.5804, 2.5786} ^T | -1.469 | 81.54 | 86.95 | 99.82 | 40 ⁽¹⁾ + 43 ⁽²⁾ + 5 ⁽³⁾ |
| 30 | {4.5938, 2.2273} ^T | -1.611 | 85.95 | 94.27 | 100.00 | 21 ⁽¹⁾ + 20 ⁽²⁾ + 4 ⁽³⁾ |
| 50 | {4.6187, 2.2539} ^T | -1.600 | 86.90 | 93.70 | 100.00 | 18 ⁽¹⁾ + 19 ⁽²⁾ + 4 ⁽³⁾ |
| 100 | {4.6255, 2.1857} ^T | -1.629 | 88.97 | 96.45 | 100.00 | 16 ⁽¹⁾ + 16 ⁽²⁾ + 4 ⁽³⁾ |

*(1), (2), and (3) mean the first, second, and third constraint, respectively

improved using higher-order MPP-based approximation methods.

On the other hand, true reliability, reliability at mean value (i.e., sample variance), and reliability at the target confidence level can be compared to see how much

uncertainty should be allowed in Table 8. The differences between reliabilities can be indices to quantify the impact of epistemic uncertainty. In Table 8, all results are taken from the CBDO optimum when distribution types are known. Certainly, reliability at MPP has to be the same

Table 7 Results of double-loop CRBDO

| ND | Optimum | Objective | Confidence ⁽¹⁾ (%) | Confidence ⁽²⁾ (%) | Confidence ⁽³⁾ (%) | # of MCSs |
|-----|-------------------------------|-----------|-------------------------------|-------------------------------|-------------------------------|-------------------------|
| 10 | {4.5840, 2.8009} ^T | -1.384 | 90.00 | 90.00 | 100.00 | 4 × 3 × 10 ⁴ |
| 30 | {4.6578, 2.2283} ^T | -1.610 | 90.00 | 90.00 | 100.00 | 5 × 3 × 10 ⁴ |
| 50 | {4.6590, 2.2599} ^T | -1.599 | 90.00 | 90.00 | 100.00 | 6 × 3 × 10 ⁴ |
| 100 | {4.6659, 2.1667} ^T | -1.636 | 90.00 | 90.00 | 100.00 | 6 × 3 × 10 ⁴ |

Table 8 Reliability under various input distribution parameters with known distribution types

| | Reliability at MPP | | | Reliability at the mean value | | | True reliability | | |
|----------|-----------------------|-----------------------|-----------------------|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 10 data | 0.9772 ⁽¹⁾ | 0.9772 ⁽²⁾ | 1.0000 ⁽³⁾ | 0.9982 ⁽¹⁾ | 0.9982 ⁽²⁾ | 1.0000 ⁽³⁾ | 0.9976 ⁽¹⁾ | 0.9979 ⁽²⁾ |
| 30 data | 0.9772 ⁽¹⁾ | 0.9772 ⁽²⁾ | 1.0000 ⁽³⁾ | 0.9925 ⁽¹⁾ | 0.9915 ⁽²⁾ | 1.0000 ⁽³⁾ | 0.9868 ⁽¹⁾ | 0.9880 ⁽²⁾ | 1.0000 ⁽³⁾ |
| 50 data | 0.9772 ⁽¹⁾ | 0.9772 ⁽²⁾ | 1.0000 ⁽³⁾ | 0.9898 ⁽¹⁾ | 0.9888 ⁽²⁾ | 1.0000 ⁽³⁾ | 0.9890 ⁽¹⁾ | 0.9892 ⁽²⁾ | 1.0000 ⁽³⁾ |
| 100 data | 0.9772 ⁽¹⁾ | 0.9772 ⁽²⁾ | 1.0000 ⁽³⁾ | 0.9870 ⁽¹⁾ | 0.9855 ⁽²⁾ | 1.0000 ⁽³⁾ | 0.9845 ⁽¹⁾ | 0.9855 ⁽²⁾ | 1.0000 ⁽³⁾ |

*(1), (2), and (3) mean the first, second, and third constraint, respectively

Table 9 Properties of random input (9 variables and 2 parameters) for side impact crashworthiness problem

| Random input | Type | STD* | Lower bound | Initial design | Upper bound |
|-------------------------------|--------|--------------|-------------|----------------|-------------|
| X_1 : B-pillar inner | Normal | 0.05 | 0.500 | 1.000 | 1.500 |
| X_2 : B-pillar reinforce | Normal | 0.05 | 0.450 | 1.000 | 1.350 |
| X_3 : Floor side inner | Normal | 0.05 | 0.500 | 1.000 | 1.500 |
| X_4 : Cross member | Normal | 0.05 | 0.500 | 1.000 | 1.500 |
| X_5 : Door beam | Normal | 0.05 | 0.875 | 2.000 | 2.625 |
| X_6 : Door belt line | Normal | 0.05 | 0.400 | 1.000 | 1.200 |
| X_7 : Roof rail | Normal | 0.05 | 0.400 | 0.300 | 1.200 |
| X_8 : Mat. B-pillar inner | Normal | 0.006 | 0.192 | 0.300 | 0.345 |
| X_9 : Mat. Floor side inner | Normal | 0.006 | 0.192 | 0.300 | 0.345 |
| X_{10} : Barrier height | Normal | 10.00 (true) | – | 0.000 | – |
| X_{11} : Barrier hitting | Normal | 10.00 (true) | – | 0.000 | – |

*STD stands for standard deviation

Table 10 Optima under various numbers of data when input distribution types are known

| ND | Optimum | Objective |
|------|---|-----------|
| 10 | {1.3289, 1.3500, 1.0428, 1.5000, 1.4765, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 33.45 |
| 30 | {0.8396, 1.3500, 0.6658, 1.5000, 1.5087, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 31.82 |
| 50 | {0.6057, 1.3500, 0.5867, 1.5000, 1.4418, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 30.97 |
| 100 | {0.6082, 1.3500, 0.6100, 1.5000, 1.4239, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 30.96 |
| RBDO | {0.5110, 1.3500, 0.5000, 1.4710, 1.4270, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 30.25 |

Table 11 Confidences at optima when input distribution types are known

| ND | Confidence ⁽¹⁾ (%) | Confidence ⁽²⁾ (%) | Confidence ⁽³⁾ (%) |
|-----|-------------------------------|-------------------------------|-------------------------------|
| 10 | 95.25 | 90.43 | 95.26 |
| 30 | 94.91 | 92.90 | 95.20 |
| 50 | 94.76 | 93.90 | 95.31 |
| 100 | 94.50 | 93.39 | 93.96 |

*(1), (2), and (3) mean the first, second, and third constraint, respectively

with the target reliability of 97.72% for the first and second constraints which are active. Reliability at mean value indicates the reliability when variances of random variables are set to sample variances. It should be gradually closer to the true reliability in which the exact variances of random variables are known. Unlike overall tendency of results, it seems that the reliability at mean value is very similar to the true reliability when the number of input data is 10 since the sample variance of drawn data from a population is coincidentally very close to the true variance.

Table 12 Optima under various numbers of data when input distribution types are KDE

| ND | Optimum | Objective |
|------|---|-----------|
| 10 | {1.1494, 1.3500, 0.9113, 1.5000, 1.4842, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 32.58 |
| 30 | {0.8730, 1.3500, 0.6880, 1.5000, 1.5028, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 31.89 |
| 50 | {0.6167, 1.3500, 0.5974, 1.5000, 1.4391, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 31.00 |
| 100 | {0.6278, 1.3500, 0.6430, 1.5000, 1.4124, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 31.04 |
| RBDO | {0.5110, 1.3500, 0.5000, 1.4710, 1.4270, 1.2000, 0.4000, 0.3450, 0.1920} ^T | 30.25 |

5.3 Eleven-dimensional side impact problem

In this section, the proposed CBDO is applied to the vehicle side impact crashworthiness problem with nine random variables and two random parameters (Youn et al. 2004). It is assumed that statistical information of nine random variables is fully known and two random parameters have uncertainty induced by insufficient input with various numbers of data. Design optimization for the vehicle side impact problem is formulated as

$$\begin{aligned} \min \quad & \text{cost}(\mathbf{d}) = 1.98 + 4.9d_1 + 6.67d_2 + 6.98d_3 + 4.01d_4 \\ & + 1.78d_5 + 2.73d_7 \\ \text{s.t.} \quad & \Pr[G_j(\mathbf{X}(\mathbf{d})) \leq 0] \geq 0.95, j = 1, 2, 3 \\ & \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U \end{aligned} \quad (30)$$

$$\text{where } G_1(\mathbf{X}) = 14.36 - 9.9X_2 - 12.9X_1X_8 + 0.1107X_3X_{10}$$

$$G_2(\mathbf{X}) = 0.72 - 0.5X_4 - 0.19X_2X_3 - 0.0122X_4X_{10} + 0.009325X_6X_{10} + 0.000191X_{11}^2$$

$$G_3(\mathbf{X}) = 1.35 - 0.489X_3X_7 - 0.843X_5X_6 + 0.0432X_9X_{10} - 0.0556X_9X_{11} - 0.000786X_{11}^2$$

where the objective is the vehicle weight, both the target confidence of reliability and target reliability are set to 95%, and the statistical information of random variables is listed in Table 9. The original side impact problem contains 10 constraints among which only three are active, and they are included in (30). Note that X_{10} and X_{11} are random parameters which are invariant during the optimization, and samples are randomly drawn from the normal distribution as shown in Table 9.

CBDO results with various numbers of input data are listed in Tables 10 and 11. In addition, CBDO results using KDE are listed in Tables 12 and 13. Overall results show that as the

Table 13 Confidences at optima when input distribution types are KDE

| ND | Confidence ⁽¹⁾ (%) | Confidence ⁽²⁾ (%) | Confidence ⁽³⁾ (%) |
|-----|----------------------------------|----------------------------------|----------------------------------|
| 10 | 92.89 | 84.92 | 93.37 |
| 30 | 96.82 | 94.55 | 94.83 |
| 50 | 96.19 | 95.19 | 94.99 |
| 100 | 96.02 | 97.94 | 90.00 |

* (1), (2), and (3) mean the first, second, and third constraint, respectively

number of data increases, the CBDO optimum gradually approaches the RBDO optimum where true input distributions of two parameters are used. It is natural that an insufficient number of samples may provide large uncertainty in distribution parameters, and thus the optima become more conservative to compensate for the lack of knowledge on input statistical models. As discussed in the previous example, the accuracy of KDE is slightly worse than known distribution cases due to inaccurate input distribution estimation and propagation of variance to the smoothing parameter in KDE.

6 Conclusion

This paper proposes RMA of CBDO that can be defined as an MPP-based approximation method in P-space to eliminate outer-loop MCS of CRBDO to efficiently obtain the confidence of reliability. RMA to deal with the uncertainty of input distributions searches MPP in P-space and approximates the reliability based on MPP, resulting in significant improvement in efficiency with an acceptable level of accuracy. For the MPP search in P-space, two approaches using input distribution parameters quantified through the normality assumption are proposed according to the uncertainty of input distribution types. In addition, stochastic sensitivity analysis for reliability with respect to input distribution parameters is derived employing the first-order score function to reduce the number of MCSs and provides accurate search direction to an optimizer. Finally, CBDO with RMA accounting for the confidence of reliability is proposed adopting the framework of RBDO using PMA when the uncertainty of input statistical models exists. As a result, a conservative optimum in the presence of the uncertainty of input statistical models can be obtained using much-reduced computations compared to the previous CRBDO. Moreover, other existing MPP-based approximation methods can be directly combined with the proposed CBDO for more accurate results.

In future work, reweighting samples of MCS to estimate the reliability will be investigated to further reduce the number of MCSs. Samples to compute the reliability at a design point in P-space can be reused in the next point during MPP search using the scheme of the importance sampling. It can be effective in RMA in CBDO because input distribution parameters

vary locally in P-space. In addition, the cost to add input samples will be treated as development cost to find the balanced optimum between system cost and development cost satisfying confidence constraints.

7 Replication of Results

Matlab codes for the mathematical examples in Section 5 are uploaded on https://github.com/Yongsu-Jung/SMO_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

- Bae S, Kim NH, Park C, Kim Z (2017) Confidence interval of Bayesian network and global sensitivity analysis. *AIAA J*:3916–3924
- 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
- Cho H, Choi KK, Gaul NJ, Lee I, Lamb D, Gorsich D (2016a) Conservative reliability-based design optimization method with insufficient input data. *Struct Multidiscip Optim* 54(6):1609–1630
- Cho H, Choi KK, Lee I, Lamb D (2016b) Design sensitivity method for sampling-based RBDO with varying standard deviation. *J Mech Des* 138(1):011405
- 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
- Choi SH, Lee G, Lee I (2018) Adaptive single-loop reliability-based design optimization and post optimization using constraint boundary sampling. *J Mech Sci Technol* 32(7):3249–3262
- Der Kiureghian A, Ditlevsen O (2009) Aleatory or epistemic? Does it matter? *Struct Saf* 31(2):105–112
- Du L, Choi KK, Youn BD (2006a) Inverse possibility analysis method for possibility-based design optimization. *AIAA J* 44(11):2682–2690
- Du L, Choi KK, Youn BD, Gorsich D (2006b) Possibility-based design optimization method for design problems with both statistical and fuzzy input data. *J Mech Des* 128(4):928–935
- Dubourg V, Sudret B, Bourinet JM (2011) Reliability-based design optimization using kriging surrogates and subset simulation. *Struct Multidiscip Optim* 44(5):673–690
- 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
- Hansen BE (2009) Lecture notes on nonparametrics. Lecture notes
- 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
- Jung Y, Cho H, Lee I (2019) MPP-based approximated DRM (ADRM) using simplified bivariate approximation with linear regression. *Struct Multidiscip Optim* 59(5):1761–1773
- Kang YJ, Lim OK, Noh Y (2016) Sequential statistical modeling method for distribution type identification. *Struct Multidiscip Optim* 54(6):1587–1607
- Keshtegar B, Lee I (2016) Relaxed performance measure approach for reliability-based design optimization. *Struct Multidiscip Optim* 54(6):1439–1454
- 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, Zhao L (2011a) Sampling-based RBDO using the stochastic sensitivity analysis and dynamic Kriging method. *Struct Multidiscip Optim* 44(3):299–317
- Lee I, Choi KK, Noh Y, Zhao L, Gorsich D (2011b) 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
- Li G, Meng Z, Hu H (2015) An adaptive hybrid approach for reliability-based design optimization. *Struct Multidiscip Optim* 51(5):1051–1065
- Liang B, Mahadevan S (2011) Error and uncertainty quantification and sensitivity analysis in mechanics computational models. *Int J Uncertain Quantif* 1(2)
- Meng Z, Li G, Wang BP, Hao P (2015a) A hybrid chaos control approach of the performance measure functions for reliability-based design optimization. *Comput Struct* 146:32–43
- Meng Z, Hao P, Li G, Wang B, Zhang K (2015b) Non-probabilistic reliability-based design optimization of stiffened shells under buckling constraint. *Thin-Walled Struct* 94:325–333
- 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
- Muscolino G, Santoro R, Sofi A (2016) Reliability analysis of structures with interval uncertainties under stationary stochastic excitations. *Comput Methods Appl Mech Eng* 300:47–69
- 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 (2011a) Reliability-based design optimization with confidence level under input model uncertainty due to limited test data. *Struct Multidiscip Optim* 43(4):443–458
- Noh Y, Choi KK, Lee I, Gorsich D, Lamb D (2011b) Reliability-based design optimization with confidence level for non-Gaussian distributions using bootstrap method. *J Mech Des* 133(9):091001

- Park JW, Lee I (2018) A study on computational efficiency improvement of novel SORM using the convolution integration. *J Mech Des* 140(2):024501
- 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
- Rahman S, Xu H (2004) A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics. *Probab Eng Mech* 19(4):393–408
- Rao SS, Cao L (2002) Optimum design of mechanical systems involving interval parameters. *J Mech Des* 124(3):465–472
- Rosenblatt M (1952) Remarks on a multivariate transformation. *Ann Math Stat* 23(3):470–472
- Rubinstein RY, Kroese DP (2016) *Simulation and the Monte Carlo method* (Vol. 10). John Wiley & Sons
- 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, Mahadevan S (2013a) Distribution type uncertainty due to sparse and imprecise data. *Mech Syst Signal Process* 37(1–2):182–198
- Sankararaman S, Mahadevan S (2013b) Separating the contributions of variability and parameter uncertainty in probability distributions. *Reliab Eng Syst Saf* 112:187–199
- 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
- Wang Z, Wang P (2014) A maximum confidence enhancement based sequential sampling scheme for simulation-based design. *J Mech Des* 136(2):021006
- Yang D, Yi P (2009) Chaos control of performance measure approach for evaluation of probabilistic constraints. *Struct Multidiscip Optim* 38(1):83
- 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
- Youn BD, Choi KK, Yang RJ, Gu L (2004) Reliability-based design optimization for crashworthiness of vehicle side impact. *Struct Multidiscip Optim* 26(3–4):272–283
- Youn BD, Choi KK, Du L (2005) Enriched performance measure approach for reliability-based design optimization. *AIAA J* 43(4):874–884
- Zhao L, Choi KK, Lee I, Gorsich D (2013) Conservative surrogate model using weighted Kriging variance for sampling-based RBDO. *J Mech Des* 135(9):091003

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