RESEARCH PAPER

CrossMark

Reliability-based design optimization under sampling uncertainty: shifting design versus shaping uncertainty

Sangjune Bae[1] · Nam H. Kim[1] · Seung-gyo Jang[2]

Abstract
The objective of this paper is a tradeoff between changing design and controlling sampling uncertainty in reliability-based design optimization (RBDO). The former is referred to as 'living with uncertainty', while the latter is called 'shaping uncertainty'. In RBDO, a conservative estimate of the failure probability is defined using the mean and the upper confidence limit, which are obtained from samples and from the normality assumption. Then, the sensitivity of the conservative probability of failure is derived with respect to design variables as well as the number of samples. It is shown that the proposed sensitivity is much more accurate than that of the finite difference method and close to the analytical sensitivity. A simple RBDO example showed that once the design variables reach near the optimum point, the number of samples is adjusted to satisfy the conservative reliability constraints. This example showed that not only shifting design but also shaping uncertainty plays a critical role in the optimization process.

Keywords Reliability-based design optimization · Uncertainty quantification · Sampling uncertainty · Conservative · Sensitivity

1 Introduction

The probability of failure is usually utilized in a design process to assure the safety of a component or a system. Notwithstanding the importance of the probability of failure, its accuracy has always been in a question. That is to say, there is uncertainty in the probability calculation (Cadini and Gioletta 2016; Howard 1988; Mosleh and Bier 1996), especially when the level of probability of failure is extremely low.

There are two categories of uncertainty involved in a system: aleatory uncertainty and epistemic uncertainty (Park et al. 2014; Hofer et al. 2002). The former is inherent variability or natural randomness that is irreducible. The aleatory

uncertainty is usually represented using a probability theory (Li et al. 2013). This randomness is reflected in the input variables in a form of probability distributions. Epistemic uncertainty, on the other hand, comes from a lack of knowledge. If additional information is provided, then the uncertainty can be reduced. However, modeling epistemic uncertainty is not a trivial task. The sources of both uncertainties can be found throughout the entire design process. For instance, the material property can be taken as an example of the aleatory uncertainty: Young's modulus and Poisson's ratio are different from a sample to another even if the samples are made from the same material. Likewise, the input design variable can be thought to have variability, which is modeled as a random variable in RBDO (Choi et al. 2005).

As well as for the aleatory uncertainty, there have been numerous trials to systematically consider the epistemic uncertainty which can be divided into the following types – surrogate model error, model form error, sampling error, numerical error, and other unknown error (Liang and Mahadevan 2011) – in a design process. For example, fuzzy set theory, possibility theory, interval analysis, and evidence theory are used to mathematically represent the epistemic uncertainty (Helton and Oberkampf 2004; Youn et al. 2006; Agarwal et al. 2004). Sometimes, the limit of the prediction interval of a surrogate model is used as a design criterion (Zhuang and Pan 2012).

✉ Nam H. Kim
   nkim@ufl.edu

   Sangjune Bae
   sjune.bae@ufl.edu

   Seung-gyo Jang
   Jsg4580@add.re.kr

[1] Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA

[2] Agency for Defense Development, Daejeon, Republic of Korea

Recently, Jiang et al. tried to improve the modeling capability in a multidisciplinary system by efficiently allocating samples by using the prediction interval (Jiang et al. 2016).

In the field of reliability analysis, the epistemic uncertainty has been a major concern as well. For example, Nannapaneni and Mahadevan utilized a novel FORM-based approach and Monte Carlo simulation (Nannapaneni and Mahadevan 2016), and Martinez et al. applied the belief function theory to a system under epistemic uncertainty to estimate the reliability (Martinez et al. 2015). However, most works focus on how to consider the effect of epistemic uncertainty, but not controlling it. Since epistemic uncertainty is related to the lack of knowledge or information, it is possible to control the epistemic uncertainty by improving a model or increasing the number of samples. Therefore, it would be necessary to include the control of epistemic uncertainty in the design process. For example, the building-block process in the aircraft industry is developed to identify design errors, which are caused by model form errors (epistemic uncertainty), and fix them in the early design stage; i.e., reducing epistemic uncertainty. So far, the industry relies on trial-and-error to reduce epistemic uncertainty.

In this paper, as a first attempt to explicitly control the epistemic uncertainty, a tradeoff between changing design variables and controlling sampling uncertainty is presented in the framework of RBDO. The conservative estimate of the failure probability is used as a constraint in sampling-based RBDO. The constraint can be satisfied by either (a) changing design, which in turn, will shift the mean of the probability of failure or (b) shaping uncertainty—adding more samples can reduce the uncertainty in the distribution. The former may increase the system weight and thus increase the operating cost, while the latter may increase the development cost as it requires more samples and tests. With a proper cost model, it is possible that the optimization algorithm can tradeoff between the two.

The paper is composed of six sections. Section 2 shows two different ways of dealing with uncertainty in the design process. Section 3 explains how the epistemic uncertainty from sampling is reflected in the sensitivity analysis. Section 4 exhibits the accuracy of the formulation by comparing the sensitivity obtained from two different methods, and Section 5 demonstrates how the derived sensitivity could be utilized in RBDO using a 2-d mathematical problem, followed by conclusions in Section 6.

2 Living with uncertainty versus shaping uncertainty

Design optimization helps to minimize the objective function while satisfying all constraints. The process is well utilized in industry to design a structure that satisfies constraints, which can be a function of external force, vibration, or heat transfer (Choi and Kim 2004). However, when the optimized structure is built, not all of them show expected performance. That is, some of them do not satisfy the prescribed constraints. This is because of the uncertainty involved in the loading conditions applied to the structure, manufacturing process, material properties, and computational errors that the optimization process did not consider. For example, a structure may undergo a larger force than the force used in the analysis or there could be heterogeneity in the material property that is hard to express with a mathematical model. Sometimes, the manufacturing tolerance is too large to make an optimized shape. In conventional optimization, the effect of uncertainty is indirectly considered by adopting the concept of safety factor or safety margin.

In order to consider uncertainty directly, reliability-based design optimization (RBDO) is developed to find an optimum with constraints on the probability of failure. RBDO takes randomness of inputs into account and searches for the optimum (Tu and Choi 1999). The idea has been thoroughly investigated with different approaches. One of them is to use the sampling method, in which the probability of failure is evaluated through Monte-Carlo sampling, possibly with a surrogate model (Lee et al. 2011).

A challenge in sampling-based reliability calculation is that the sampling uncertainty also affects the probability calculation, which in turn, results in the distribution of the probability (see Fig. 1). If the epistemic uncertainty can be quantified, then it can be included in the design process by introducing a conservative probability of failure. It is possible to make the design strategically conservative to consider such uncertainty.

Fig. 1 Options to satisfy reliability constraint with conservative probability of failure. (**a**) living with uncertainty, (**b**) shaping uncertainty
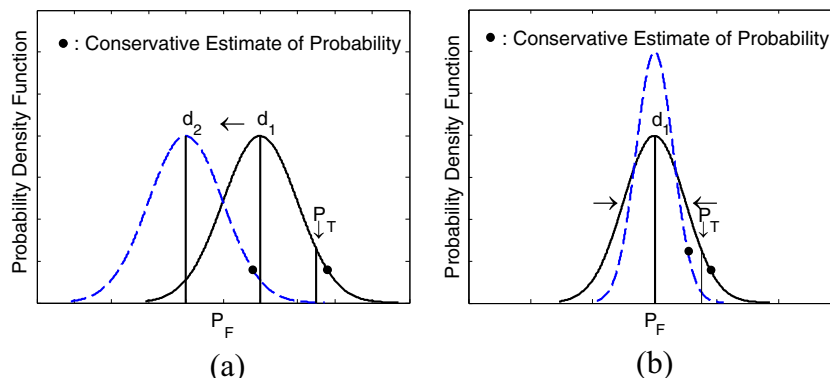
In detail, the target probability of failure in a probabilistic constraint can be achieved with the upper confidence limit of the probability of failure. If the desired confidence level is $1 - \alpha$ and the probability of failure follows a normal distribution, then the reliability constraint with the conservative estimate can be written as

$$P_{F,cons} = P_F + z_{1-\alpha}\sigma_{P_F} \leq P_T \tag{1}$$

where P_F and σ_{P_F} are, respectively, the mean and standard deviation of the probability of failure estimated from samples, $z_{1-\alpha}$ is the z-score, and P_T is the target probability of failure.

When the conservative probability of failure violates the reliability constraint, the mean probability of failure can be reduced, which can be achieved by changing design toward a conservative direction. Although the mean probability can be reduced by a conservative design, it still needs to quantify the effect of epistemic uncertainty σ_{P_F} on the probability of failure to shift the mean properly. In this paper, this approach is called 'living with uncertainty' as shown in Fig. 1a. This is the basic approach in conventional RBDO — designing the system to satisfy reliability with all given uncertainties. However, since there are many sources of epistemic uncertainty, the design may end up too conservative if all uncertainties are included. Therefore, it is often necessary to reduce epistemic uncertainty by using a high fidelity model or having more tests. In this paper, we only consider sampling uncertainty as for the source of epistemic uncertainty. That is, the uncertainty in the probability of failure calculation depends on the number of samples. If more samples are used, the uncertainty in the probability of failure will be reduced, but the computation will be more costly. However, the population mean of the probability of failure remains unchanged, and therefore, only the distribution of the probability of failure is narrowed. Note that the estimate of the mean probability can change, but the population mean does not change. In this paper, this approach is called 'shaping uncertainty' as shown in Fig. 1b.

In the figure, the distribution of the probability of failure at the original design is given as the solid line. At the design d_1, the conservative estimate $P_{F, cons}$ which is represented by the black dot in the figure does not satisfy the reliability constraint; i.e., it is larger than the target probability of failure. In order to make the conservative estimate to satisfy the constraint, either the standard deviation of the distribution should be reduced or the mean should be shifted by changing the design. When the design is changed, the conservative estimate can satisfy the constraint as the dot moves to the left as in Fig. 1a. If more samples are provided so that the uncertainty is reduced, $P_{F, cons}$ can also satisfy the constraint, although the design does not change, as shown in Fig. 1b. In other words, not only changing the design variables but also reducing the uncertainty can be used to satisfy reliability constraint in RBDO. The conventional RBDO only looks for changing design to satisfy the reliability constraint.

3 Design sensitivity analysis under epistemic uncertainty

As explained in the previous section, epistemic uncertainty yields the distribution of the probability of failure. In order to compensate for this uncertainty, a conservative probability of failure was introduced by using both the mean and standard deviation of this distribution. When design changes, it is possible that either the mean of the distribution is shifted or the standard deviation of the distribution changes, both of which can change the conservative probability of failure. It is also possible that the standard deviation of the distribution can change when the number of samples changes. In this section, the design sensitivity of the conservative probability of failure with respect to either design variables or the number of samples is developed.

There are many methods available for calculating the probability of failure, such as the first- or second-order reliability methods (Haldar and Mahadevan 2000), surrogate-based methods (Bichon et al. 2011), and sampling-based methods (Picheny et al. 2010; Bae et al. 2017). Different methods have their own advantages but it is the out of the scope of this paper to discuss them. In this paper, it is assumed that the probability of failure is calculated using a sampling-based method, more specifically, Monte Carlo Simulation (MCS).

Let us assume that N samples are available for calculating the probability of failure. In MCS, the probability of failure is calculated as

$$P_F(\mathbf{d}) = \frac{1}{N}\sum_{i=1}^{N} I_F(y_i) \tag{2}$$

where I_F is the indicator function, which becomes 1 if y_i is in the failure region; otherwise 0:

$$I_F(y_i) = \begin{cases} 0 & \text{if } y_i \notin \Omega_F \\ 1 & \text{if } y_i \in \Omega_F \end{cases} \tag{3}$$

The sum of the indicator function, $NP_F(\mathbf{d})$, approximately follows a normal distribution $N(NP_F, NP_F(1 - P_F))$ when the normality condition is satisfied, such that $NP_F > 10$ and $N(1 - P_F) > 10$ (Fraser 1958). Therefore the probability of failure follows a normal distribution, $P_F \sim N(P_F, P_F(1-P_F)/N)$. The unique characteristic of sampling-based methods is that the calculated probability of failure has sampling uncertainty. That is, P_F is random because a different set of samples may yield different values. Note that $NP_F(\mathbf{d})$ follows a normal distribution regardless of the distribution type or correlation of input random variables. When MCS is used to calculate P_F, its standard deviation can be estimated as

$$\sigma_{P_F} = \sqrt{\frac{P_F(1-P_F)}{N}} \tag{4}$$

where N is the number of samples used to calculate P_F. An important observation in (4) is that the uncertainty is

independent of the number of input random variables. Rather, it is a function of the level of probability and the number of samples. Note that the variance of P_F is inversely proportional to the number of samples used. Therefore, the variance of the probability of failure can be controlled by changing the number of samples. Equation (4) is also an estimate, which means that uncertainty exists in (4) because P_F is an estimate of the true failure probability. However, the uncertainty of standard deviation is not considered in this research because it is small when compared to the uncertainty in P_F. The expression in (4) is valid when all samples are independent and identically distributed, which is the case of MCS. Therefore, it is unnecessary that the underlying distribution is Gaussian.

Using (2) and (4), the ratio between P_F and σ_{P_F} is shown when the number of samples N is changed from 1000 to 10,000 in Fig. 2 at different probabilities. Even if the absolute magnitude of the standard deviation decreases as P_F decreases, its relative magnitude compared to P_F increases. That is to say, the uncertainty in probability can be much larger than the probability itself. Therefore, this uncertainty plays a key role when the probability is small. Further, the relative uncertainty decreases as more samples are provided, which confirms that introducing more samples is indeed a way to reduce the uncertainty.

To compensate for sampling epistemic uncertainty, (1) can be used for reliability analysis. To calculate the sensitivity of (1) with respect to an input design variable, the mean probability of failure is considered first. Using (2), the design sensitivity can be calculated using Leibniz's rule as (Lee et al. 2011)

$$\frac{\partial P_F(\mathbf{d})}{\partial d_i} = \frac{1}{N} \sum_{j=1}^{N} I_F(\mathbf{x}_j) s_{d_i}(\mathbf{x}_j; \mathbf{d}) \tag{5}$$

In (5), $s_{d_i}(\mathbf{x}; \mathbf{d})$ is the partial derivative of the log-likelihood function with respect to its argument, which is called a score function. If the probability is determined based on a finite number of samples, then the sampling uncertainty as in (4)
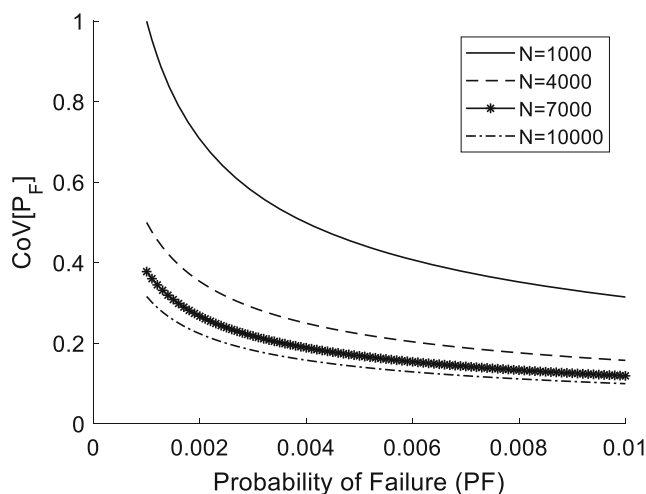

Fig. 2 Relative importance of uncertainty in probability of failure

is induced. Using the result, the design sensitivity of (1) can be derived using the chain rule of differentiation as

$$\frac{\partial P_{F,cons}}{\partial d_i} = \frac{\partial P_F}{\partial d_i} \left(1 + z_{1-\alpha} \frac{1 - 2P_F}{2N\sigma_{P_F}} \right) \tag{6}$$

As mentioned before, (6) is the sensitivity of the conservative estimate of the probability of failure. The sensitivity consists of two parts: the effect of design perturbation on the probability and that on the uncertainty of the probability. If the uncertainty of the probability is not considered, then the latter part of (6) vanishes, making it same as (5).

Fig. 1b suggests another way of design. That is, utilizing the number of samples to make the conservative estimate satisfy the reliability constraint. The design sensitivity of (1) with respect to the number of samples can be derived as

$$\frac{\partial P_{F,cons}}{\partial N} = -z_{1-\alpha} \frac{\sigma_{P_F}}{2N} \tag{7}$$

Equation (7) shows that changing the number of samples to the probability calculation does not change the mean probability but only the uncertainty of the probability. Also, the sign of (7) is always negative, which suggests that additional samples always reduces the uncertainty.

4 Accuracy of design sensitivity analysis under epistemic uncertainty

To demonstrate the accuracy of the sensitivity derived in (6) and (7), a simple linear function with a known input distribution is considered. Let $y(x) = x$ and $x \sim N(d, 1^2)$, where d is the current design point. Here the failure region is defined as $\Omega_F = \{y | y < y_{th}\}$ with $y_{th} = -2.3263$. A conservative estimate is chosen with the confidence level to 95%; that is, $z_{95\%} = 1.645$. When the current design point is $d = 0$, the true probability of failure is equal to 1%. The target probability of failure that the conservative estimate must satisfy is 1.3%. The number of samples used in this problem is 1000, and therefore, the conservative estimate of the probability is $P_F + z_{95\%}\sigma_{P_F} = 1.52\%$. Before carrying out the sensitivity analysis, it is verified if the target probability can be achieved by the proposed methods as shown in Fig. 3. To satisfy the target probability, either the design must be shifted by 0.0691 as in Fig. 3a, or the number of samples must be increased by 1970 as in Fig. 3b. Note that the mean probability of failure changes when the design is shifted, but it remains the same when more samples are added. Also, the variance of the distribution changes when the design is shifted because the variance is also a function of the probability.

Fig. 3 Verification of options to satisfy reliability constraint using conservative estimate. (**a**) by shifting design, (**b**) by increasing the number of samples

(a)　　　　(b)
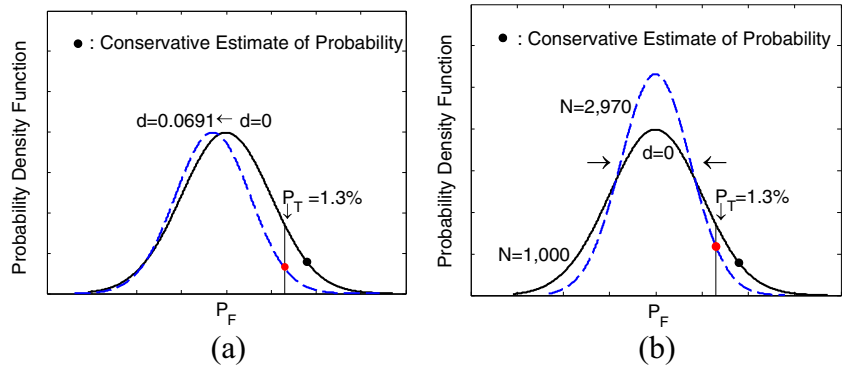
To show the accuracy of the sensitivity result, the score function is calculated first. Since The input random variable follows a normal distribution with the current design point $d = 0$, the score function becomes

$$\frac{\partial \ln f_x(\mathbf{x}; d, \sigma)}{\partial d} = x \qquad (8)$$

Using the chain rule of differentiation, the sensitivity of the conservative estimate in (6) is calculated as

$$\frac{\partial P_{F,cons}}{\partial d} \simeq \left(\frac{1}{N} \sum_{m=1}^{N} I(x_m) x_m \right) \left(1 + z_{0.95} \frac{1 - 2P_F}{2\sqrt{N P_F(1 - P_F)}} \right) \qquad (9)$$

The sensitivity in (9) is compared with that of the finite difference method (FDM), which is based on the perturbation of design as

$$\frac{\partial P_{F,cons}^{FDM}}{\partial d} \simeq \frac{P_{F,cons}(d + \Delta d) - P_{F,cons}(d)}{\Delta d} \qquad (10)$$

Figure 4 shows the estimated distribution of the probability of failure with 10,000 repetitions. As expected, the probability of failure is normally distributed and the parameters are estimated as $\mu = 0.01$, $\sigma = 3.118 \times 10^{-4}$, while the
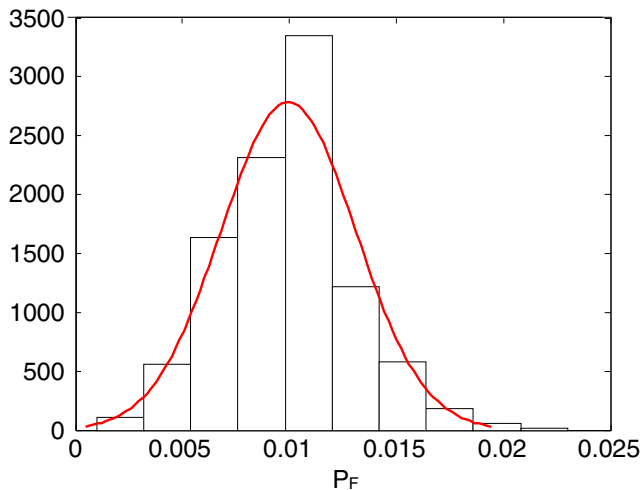
Fig. 4 Distribution of probability of failure

population mean and variance of the probability of failure are calculated as $\mu_{\text{true}} = 0.01$, $\sigma_{\text{true}} = 3.146 \times 10^{-4}$.

Figures 5 , 6 and Table 1 compare the distribution of the sensitivity using the two different methods with 10,000 repetitions. The mean of sensitivity obtained from (9) is equal to -0.0348, while the mean obtained by FDM varies from -0.0322 to -0.0148 depending on Δd. Considering that the exact value is -0.0348, the proposed method can provide an accurate sensitivity of the conservative probability of failure. The standard deviation of FDM is the smallest when $\Delta d = 1$, which is approximately 2 times smaller than that from the proposed method. However, the accuracy of the sensitivity is inferior to the proposed method. Moreover, the sensitivity obtained using FDM can be larger than 0, which contradicts the fact that the sampling uncertainty is only reduced when additional samples are provided. Thus, it would be difficult, if not impossible, to calculate the sensitivity using FDM when only a small number of samples are available. In fact, it is highly recommended to use the sensitivity derived in (9) instead of the FDM since the problems in this research arise from a limited number of samples. Note that the distribution of sensitivity using (9) converges to the true value as the number of samples is increased in Fig. 6.

For this simple analytical example, it is possible to calculate the analytical sensitivity of the conservative estimate, which is given in Table 1. In (6), only the calculation of $\partial P_F / \partial d$ requires samples, which can be obtained through numerical integration in this example as

$$\frac{\partial P_F(d)}{\partial d} = \int_{-\infty}^{y_{th}} \frac{1}{\sqrt{2\pi}\sigma} \frac{(x - d)}{\sigma^2} \exp\left[-\frac{(x - d)^2}{2\sigma^2} \right] dx \qquad (11)$$

Therefore, the analytical sensitivity in Table 1 is calculated by substituting (11) in (6), which is shown in Table 1. Note that since the analytical sensitivity does not use samples, its value is deterministic. The sensitivity obtained using the proposed method still has uncertainty in itself. However, the analytical solution is seldom available in reality because of the nonlinearity in the limit state function that defines the failure region. Therefore it is suggested to

Fig. 5 Distribution of design
sensitivity using finite difference
method
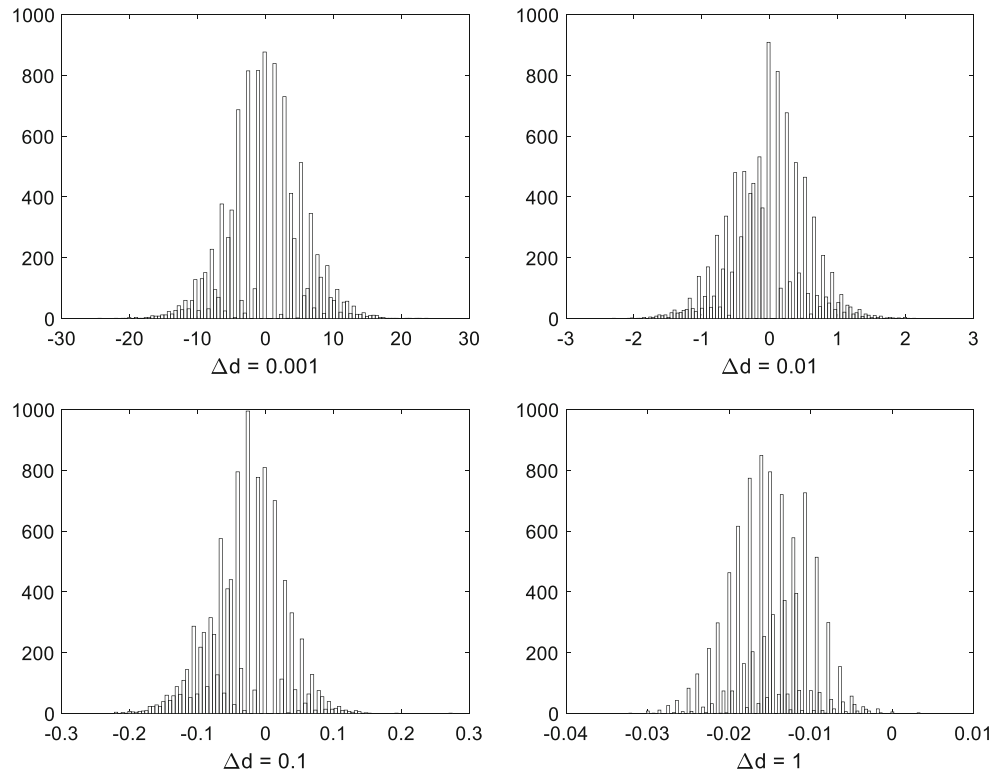
utilize the proposed method which is still more accurate and
precise than FDM.

In addition to the design variable, the conservative proba-
bility of failure also depends on the number of samples. The

sensitivity with respect to the number of samples is much
smaller than that of the design variables because adding or
removing one sample has a negligible effect on the probability
estimation. Therefore, a scaled variable, $n = N/1,000$, is

Fig. 6 Distribution of design
sensitivity using proposed
method for different number of
samples
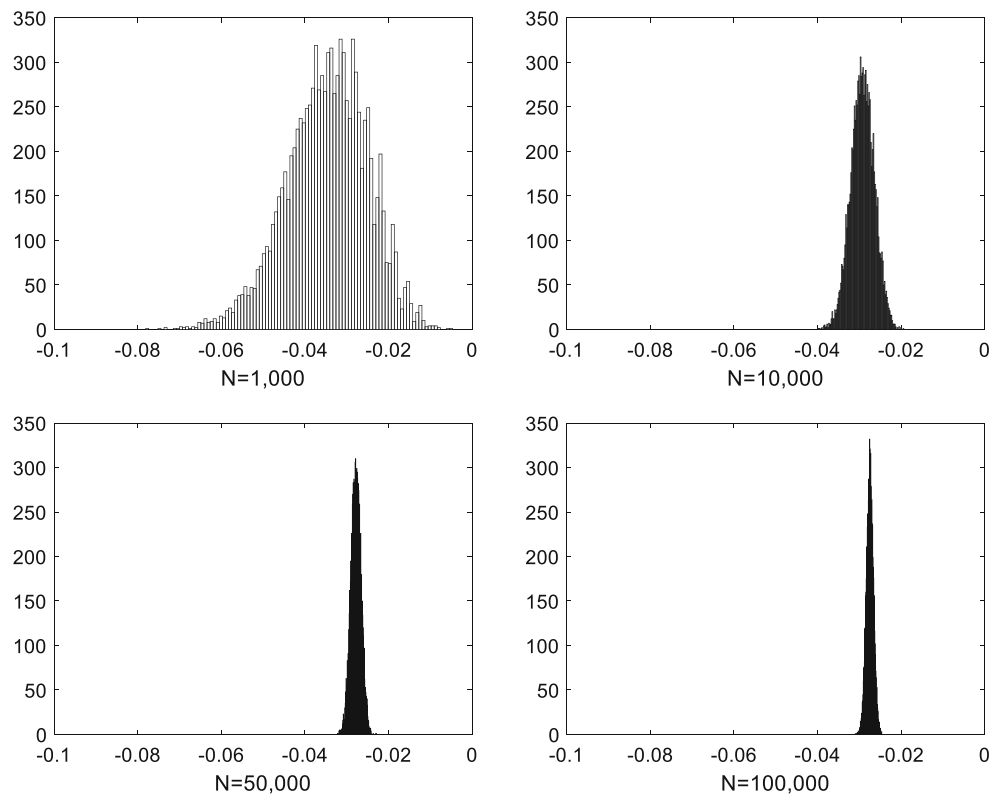
Table 1 Mean and standard deviation of sensitivity of conservative probability of failure with respect to design variable

	Analytical sensitivity	Sensitivity using (10)	Sensitivity using FDM			
			$\Delta d = 0.001$	$\Delta d = 0.01$	$\Delta d = 0.1$	$\Delta d = 1$
Mean	−0.0348	−0.0348	−0.0255	−0.0311	−0.0322	−0.0148
Standard deviation	–	0.01	5.8945	0.3345	0.0556	0.0045

introduced in optimization. Then, the sensitivity expression in (7) can be converted to

$$\frac{\partial P_{F,cons}}{\partial n} = z_{1-\alpha} \frac{\partial \sigma_{P_F}}{\partial N} \frac{\partial N}{\partial n} \simeq -500 z_{1-\alpha} \frac{\sigma_{P_F}}{N} \quad (12)$$

Table 2 shows the sensitivity results from all three methods based on 10,000 repetitions. The sensitivity using (12) with the population mean of $P_F = 0.01$, is compared with those obtained by FDM and (12) with random sampling. The process is repeated 10,000 times to compare the accuracy of the proposed method. For the convergence study of FDM, Δn is varied from 0.001 to 1. As shown in the table, the derived sensitivity is more accurate than the finite difference method and has a small standard deviation as well while FDM failed to estimate the sensitivity accurately.

5 Design optimization of 2-D mathematical problem

To illustrate the effect of sampling uncertainty on optimum design, a 2-D mathematical problem is formulated. Information regarding the random variables is provided in Table 3, where the design variable is the mean of a random variable. The objective function consists of two parts: the operation cost and the design cost. The operation cost is an expected expense of a design during the operation, and therefore, it is a function of the design variables. On the other hand, the design cost is the expense of design process itself, and thus, this is a function of the number of samples. There are three constraints, among which $G_2(\mathbf{X})$ is highly nonlinear. The lower and upper bounds of the number of samples, N, are 1000 and 10,000, respectively. A scaled variable $n = N/10,000$ is used to normalize the cost function and the constraints. P_F is evaluated for each constraint, and the combined effect of more than two constraints is ignored in this paper. For the effect of dependency between multiple constraints, readers are referred to the work of Park et al. (Park et al. 2015). The design cost is modeled as a monotonically increasing function of N, however, it is modeled as a mildly convex function to show that RBDO can make tradeoff between the two costs in this problem.

$$
\begin{aligned}
\min f(\mathbf{d}) &= \text{operation cost} + \text{design cost} \\
&= \left[-\frac{(d_1 + d_2 - 10)^2}{30} - \frac{(d_1 - d_2 + 10)^2}{120} + 10 \right] + \left[\frac{0.01}{n} + 0.2n \right] \\
S.T. \quad &P[G_i(\mathbf{x}) > 0] + z_{97.5\%} \sigma_{P_i} \leq P_{T,i}, \ \ i = 1{\sim}3 \\
&z_{97.5\%} = 1.96 \\
&\mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U, \mathbf{d} \in \mathbf{R}^2 \text{ and } \mathbf{X} \in \mathbf{R}^2 \\
&0.1 \leq n \leq 1
\end{aligned}
$$

$$(13)$$

Table 2 Mean and standard deviation of sensitivity of conservative probability of failure with respect to number of samples

	Sensitivity using (12) with population mean of P_F	Sensitivity using (12)	Sensitivity using FDM			
			$\Delta n = 0.001$	$\Delta n = 0.01$	$\Delta n = 0.1$	$\Delta n = 1$
Mean	−2.013E-03	−1.987E-03	−5.554E-03	−2.704E-03	−1.803E-03	−1.137E-03
Standard deviation	–	3.191E-04	9.155E-01	8.988E-02	8.665E-03	7.219E-04

Table 3 Distribution of random variables

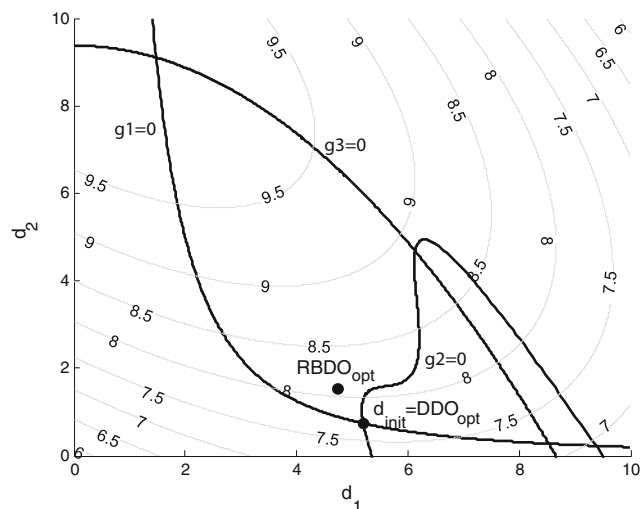| Random Variables	Distribution Type	Mean	Standard Deviation	Lower Limit of mean	Upper Limit of mean
X_1	Normal	d_1	0.3	0	10
X_2	Normal	d_2	0.3	0	10

Fig. 7 Reliability-based design optimum using 1 million samples

Equation (14) shows the three constraint functions applied in this problem.

$$
\begin{aligned}
G_1(\mathbf{X}) &= 1 - \frac{X_1^2 X_2}{20} \\
G_2(\mathbf{X}) &= -1 + (Y-6)^2 + (Y-6)^3 - 0.6(Y-6)^4 + Z \\
G_3(\mathbf{X}) &= 1 - \frac{80}{X_1^2 + 8X_2 + 5}
\end{aligned}
\tag{14}
$$

where Y and Z are intermediate variables correlated as

$$
\begin{Bmatrix} Y \\ Z \end{Bmatrix} = \begin{bmatrix} 0.9063 & 0.4266 \\ 0.4266 & -0.9063 \end{bmatrix} \begin{Bmatrix} X_1 \\ X_2 \end{Bmatrix}
\tag{15}
$$

Figure 7 shows the contour of the objective function and the limit states of the three constraints. The initial design point is set to the deterministic design optimum (DDO) which is $(d_1, d_2) = (5.1969, 0.7405)$. $N = 7500$ samples are initially used to evaluate the probabilities of failure and their sensitivities. The operation cost at the DDO is 7.7083. To ensure that the proposed sensitivity finds an optimum correctly, first, the reliability-based design optimum is found using 1 million samples, which is assumed large enough to ignore the

sampling uncertainty. Because the number of samples is fixed, the design cost becomes constant.

With aleatory uncertainty alone, the optimum design from the conventional RBDO is located at $(d_1, d_2) = (4.7324, 1.5544)$ and the operation cost is 8.0919, which is shown in Fig. 7. The optimum design has been shifted inside the feasible region, and the cost is increased by 0.3836 from that of DDO. The probability constraints of G_1 and G_2 are active, and thus, the probability of failure for both constraints is $P_F(G_1) = P_F(G_2) = 2.275\%$.

Now the RBDO searches for the optimum design considering the sampling uncertainty with $N = 750$. The optimization is repeated 100 times with different sets of samples to show the behavior of the RBDO using a conservative estimate of P_F. Figure 8 summarizes the optimization results.

All trials found an optimum fairly close to the RBDO optimum. In the magnifying window, the black dot is the optimum design with aleatory uncertainty only, while red dots are 100 optimum designs considering sampling uncertainty. As expected, the optimum designs are shifted further inside the feasible region and increased the operation cost to compensate for the sampling uncertainty as shown in Fig. 9. The box plot of the d_1 and d_2 location at the optimum is provided in Fig. 10.

The mean of the number of samples required to reach the optimum is around 3000. This is not enough number to precisely calculate the probability of failure and therefore the probability distribution still exists. However, this number is obtained by the RBDO considering both the operation cost and the design cost. It is always advisory to increase the number of samples to reach the exact optimum, but doing so costs too much design expense. Therefore, the RBDO tries to find a balance between the two costs. If the design cost is too high compared to the operation cost, then the RBDO minimizes the design cost. In contrast, the RBDO maximizes the design cost when it is trivial. The definition of non-trivial design cost may vary in context, but in this problem, about 1% of the total cost is design cost.
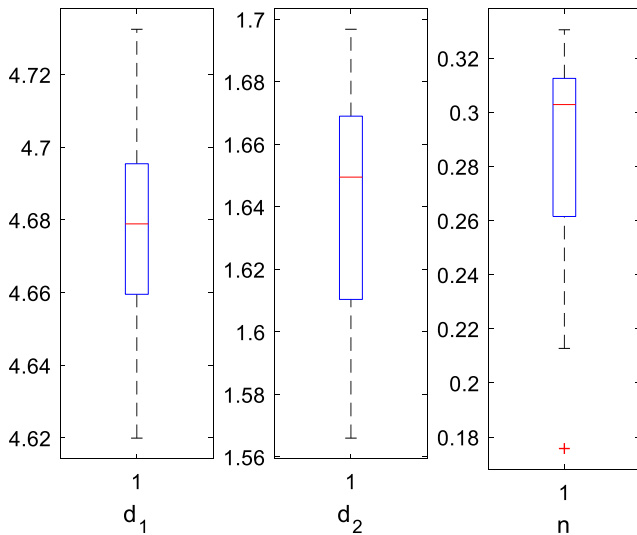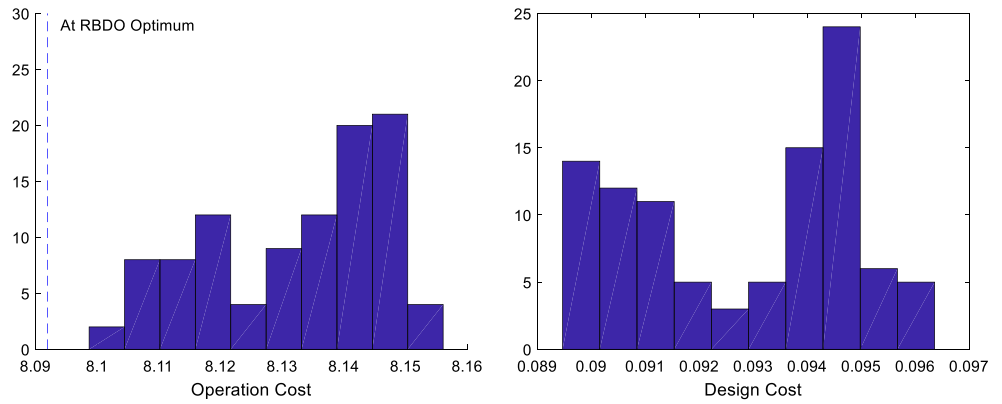
Fig. 10 shows the d_1 and d_2 locations and the number of samples required to calculate the probability of failure at the optimum point for 100 repetition of the RBDO. The RBDO

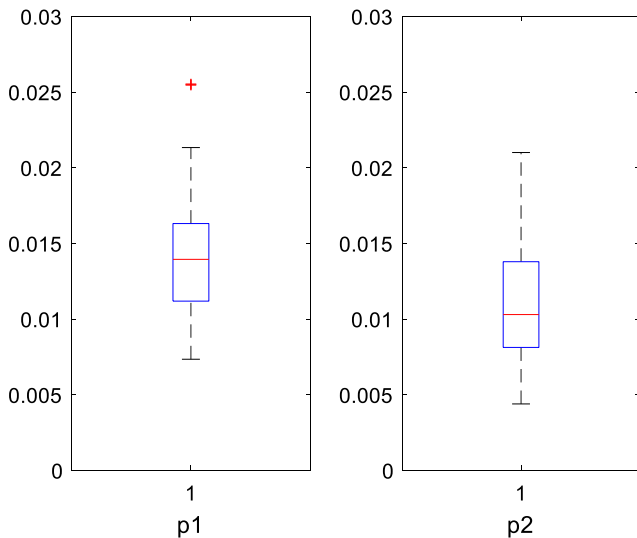**Fig. 8** Reliability-based design optimization under sampling uncertainty
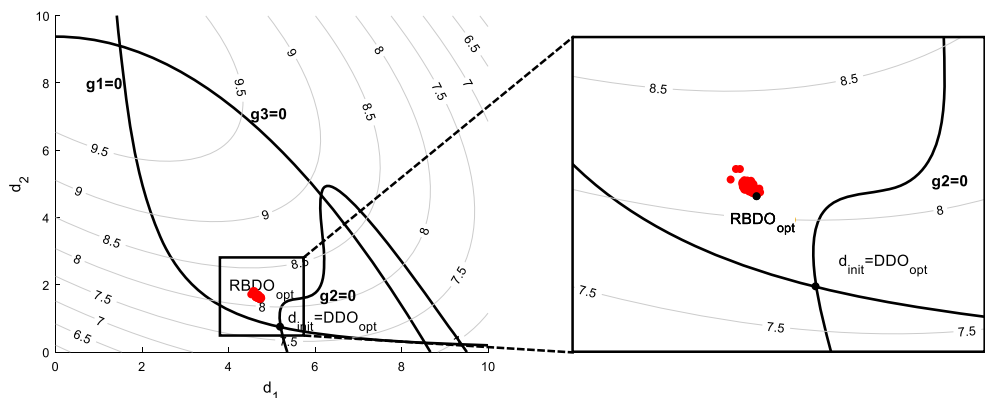
Fig. 9 Histogram of operation cost and design cost at optimum under sampling uncertainty

Fig. 10 Boxplot of optimum points

does not simply ignore the design cost which is small relative to the operation cost. The design cost function is a convex function, the minimum design cost can be found at $n = 0.2236$. However, the design cost itself is not minimized by the proposed RBDO formulation. Instead, the RBDO tries to find an optimum at which the total cost can be minimized. The range of the samples found at the optimum is from around 2100 to over 3300 which is relatively wide. This is because the design cost is a mildly linear function of n.

At each optimum, the probability calculation has been carried out to calculate the mean of the probability and confidence limit of the distribution, which is shown in Fig. 11. At the optimum, $P_F(G_3) = 0\%$ because it is inactive. Therefore the box plot is created only for $P_F(G_1)$ and $P_F(G_2)$. The mean probability of failure is now less than the target, which is 2.275%. Instead, the RBDO tries to match the target probability with the conservative estimate. It is also observed that the probabilistic constraint was actually violated just in one case out of 100 trials. The confidence level of the problem is 97.5%, and therefore only one or two cases have been expected to violate the constraints, which is consistent with the result. If the same number of samples were used for the RBDO, then around half of the probability calculation would not meet the target probability, which the RBDO clearly avoids when a conservative estimate of P_F is applied as in Fig. 11.

Fig. 11 Boxplot of mean probability of failure at optimum
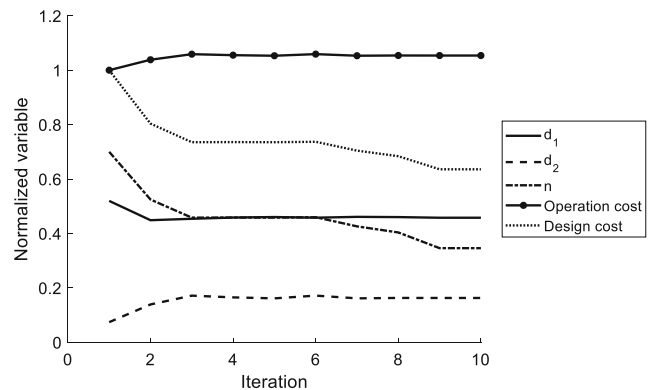
Fig. 12 Design history of reliability-based design optimization under sampling uncertainty

Table 4 Convergence of RBDO under sampling uncertainty

Number of samples	Mean			Variance		
	Operation cost	d_1	d_2	Operation cost	d_1	d_2
10,000	8.1336	4.6703	1.6451	3.28E-04	1.11E-03	1.60E-03
100,000	8.1064	4.7128	1.5838	9.59E-05	3.16E-04	4.86E-04
500,000	8.0984	4.7233	1.5661	3.50E-05	1.78E-04	1.81E-04
RBDO Optimum	8.0919	4.7324	1.5544	–	–	–

The optimization history shows an example of how the RBDO performs the tradeoffs. Figure 12 shows the design iteration history of the trial 1, where the design variables are normalized by the upper limit of their own and the cost is normalized by the initial value. Note that the operation cost increases because the RBDO starts from the deterministic design optimum. In an early iteration, the RBDO tries to change both the design points and the number of samples to satisfy the target. Once the design variables reach near the optimum point, only the number of samples is adjusted to satisfy the conservative reliability constraints. Thus, both the shifting design and shaping uncertainty that were suggested in Fig. 2 are manifested.

Lastly, the convergence of the RBDO is traced while increasing the number of samples. In Table 4, the mean and standard deviation of the objective function and the optimum location are calculated based on 100 repetition. As seen in the table, the standard deviation decreases as more samples are applied. Also, the mean of objective function approaches the aleatory-only RBDO optimum when the sample size is increased. This is because the sampling uncertainty decreases as more samples are applied.

6 Conclusion

This paper showed that epistemic uncertainty could be controlled during reliability-based design optimization (RBDO). Then, the optimization algorithm found a tradeoff between changing designs and shaping uncertainty. This paper proposed three key components to make such a tradeoff possible. (1) The objective function must depend on both the design variables and the epistemic uncertainty. In particular, in this paper the objective function composed of the operation cost and development cost. The former depends on design variables, while the latter depends on the number of samples used in the design stage. (2) In RBDO, the conservative estimate of the probability of failure is defined using the mean and the upper confidence limit, which are obtained from samples and from the normality assumption. (3) The sensitivity of the conservative probability of failure is derived with respect to design variables as well as the number of samples. It is shown that the proposed sensitivity is much more accurate than that

of the finite difference method and actually close to the analytical sensitivity. A simple RBDO example showed that once the design variables reach near the optimum point, the number of samples is adjusted to satisfy the conservative reliability constraints. This example showed that not only shifting design but also shaping uncertainty plays a critical role in the optimization process.

In fact, in addition to sampling uncertainty, there are other types of epistemic uncertainties. However, formulating such uncertainties in a form of probability distribution is not a trivial task while sampling uncertainty can be expressed using a distribution. Including other kinds of epistemic uncertainties will be explored in the future research.

Nomenclature **d**, design point; $G(\cdot)$, limit state function; $z_{1-\alpha}$, $1-\alpha$ level z-score; $\sigma(\cdot)$, standard deviation; $V(\cdot)$, variance; P_T, target probability of failure; y_{th}, threshold value of y; P_F, probability of failure; $f_x(\cdot)$, probability density function; $s(\cdot)$, score function; I_F, indicator function; Ω_F, failure domain

References

Agarwal H, Renaud JE, Preston EL, Padmanabhan D (2004) Uncertainty quantification using evidence theory in multidisciplinary design optimization. Reliab Eng Syst Saf 85(1-3):281–294

Bae S, Kim NH, Park C (2017) Confidence interval of Bayesian network and global sensitivity analysis. AIAA J 55(11):3916–3924. https://doi.org/10.2514/1.J055888

Bichon BJ, McFarland JM, Mahadevan S (2011) Efficient surrogate models for reliability analysis of systems with multiple failure modes. Reliab Eng Syst Saf 96(10):1386–1395

Cadini F, Gioletta A (2016) A Bayesian Monte Carlo-based algorithm for the estimation of small failure probabilities of systems affected by uncertainties. Reliab Eng Syst Saf 153:15–27

Choi KK, Kim NH (2004) Structural Sensitivity Analysis and Optimization I: Linear Systems. Springer, New York

Choi KK, Youn BD, Du L (2005) Integration of Reliability- and Possibility-Based Design Optimizations Using Performance Measure Approach. SAE World Congress, Detroit 2005-01-0342

Fraser DAS (1958) Statistics: an introduction. John Wiley & Sons Inc, Hoboken Chaps 2

Haldar A, Mahadevan S (2000) Reliability and Statistical Methods in Engineering Design. John Wiley & Sons Inc, New York

Helton JC, Oberkampf WL (2004) Alternative representations of epistemic uncertainty. Reliab Eng Syst Saf 85(1-3):1–10

Hofer E, Kloos M, Krzykacz-Hausmann B, Peschke J, Woltereck M (2002) An approximate epistemic uncertainty analysis approach in the presence of epistemic and aleatory uncertainties. Reliab Eng Syst Saf 77(3):229–238

Howard RA (1988) Uncertainty about probability: A decision analysis perspective. Risk Anal 8(1):91–98

Jiang Z, Chen S, Apley DW, Chen W (2016) Reduction of Epistemic Model Uncertainty in Simulation-Based Multidisciplinary Design. ASME J Mech Des 138(8):081403-1–081403-13. https://doi.org/10.1115/1.4033918

Lee I, Choi KK, Zhao L (2011) Sampling-based RBDO using the stochastic sensitivity analysis and dynamic kriging method. Struct Multidiscip Optim 44(3):299–317

Li Y, Chen J, Feng L (2013) Dealing with Uncertainty: A Survey of Theories and Practices. IEEE Trans Knowl Data Eng 25(11):2463–2482

Liang B, Mahadevan S (2011) Error and uncertainty quantification and sensitivity analysis in mechanics computational models. Int J Uncertain Quantif 1(2):147–161

Martinez FA, Sallak M, Schon W (2015) An efficient method for reliability analysis of systems under epistemic uncertainty using belief function theory. IEEE Trans Reliab 64(3):893–909

Mosleh A, Bier VM (1996) Uncertainty about probability: a reconciliation with the subjectivist viewpoint. IEEE Trans Syst Man Cybern Syst Hum 26(3):303–310

Nannapaneni S, Mahadevan S (2016) Reliability analysis under epistemic uncertainty. Reliab Eng Syst Saf 155:9–20

Park CY, Kim NH, Haftka RT (2014) How coupon and element tests reduce conservativeness in element failure prediction. Reliab Eng Syst Saf 123:123–136

Park CY, Kim NH, Haftka RT (2015) The effect of ignoring dependence between failure modes on evaluating system reliability. Struct Multidiscip Optim 52(2):251–268. https://doi.org/10.1007/s00158-015-1239-7

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

Tu J, Choi KK (1999) A New Study on Reliability-Based Design Optimization. ASME J Mech Des 121(4):557–564

Youn BD, Choi KK, Du L, Gorsich D (2006) Integration of Possibility-Based Optimization and Robust Design for Epistemic Uncertainty. ASME J Mech Des 129(8):876–882. https://doi.org/10.1115/1.2717232

Zhuang X, Pan R (2012) Epistemic uncertainty in reliability-based design optimization. In: 2012 Proc. Annual Reliability and Maintainability Symposium, Reno, pp. 1-6, https://doi.org/10.1109/RAMS.2012.6175496