AIAA JOURNAL
Vol. 45, No. 9, September 2007

Sensitivity Analysis with Mixture of Epistemic and Aleatory Uncertainties

Jia Guo* and Xiaoping Du†
University of Missouri—Rolla, Rolla, Missouri 65409-0050

DOI: 10.2514/1.28707

The study on epistemic uncertainty due to the lack of knowledge has received increasing attention in risk assessment, reliability analysis, decision making, and design optimization. Different theories have been applied to model and quantify epistemic uncertainty. Research on sensitivity analysis for epistemic uncertainty has also been initialized. Sensitivity analysis can identify the contributions of individual input variables with epistemic uncertainty to the model output. It then helps guide the collection of more information to reduce the effect of epistemic uncertainty. In this paper, an effective sensitivity analysis method for epistemic uncertainty is proposed when both epistemic and aleatory uncertainties exist in model inputs. This method employs the unified uncertainty analysis framework to calculate the plausibility and belief measures. The gap between belief and plausibility measures is used as an indicator of the effect of epistemic uncertainty on the model output. The Kolmogorov–Smirnov distance between the two measures is used to quantify the main effect and the total effect of each independent variable with epistemic uncertainty. By the Kolmogorov–Smirnov distance, the importance of each variable is ranked. The feasibility and effectiveness of the proposed method is demonstrated with two engineering examples.

Nomenclature

Bel = belief
C = subset of intervals
d_{KS} = KS distance
F = cumulative distribution function (CDF)
f = probability density function (PDF)
G = output of a performance function
g = performance function
ME = main effect
ME_{pf} = main effect on the probability of failure
m_Y = basic probability assignment (BPA)
P = probability
Pl = plausibility
p_f = probability of failure
R = reliability
TE = total effect
TE_{pf} = total effect on the probability of failure
$\mathbf{U}$ = vector of standard normal variables
U = standard normal variable
u = realization of U
$\mathbf{u}^*$ = most probable point in U space (MPP)
$\mathbf{X}$ = vector of random variables
X = random variable
x = realization of X
$\mathbf{x}^*$ = most probable point in X space (MPP)
$\mathbf{Y}$ = vector of variables with epistemic uncertainty
y = realization of Y
β = reliability index
Φ = cumulative distribution function of a standard normal distribution
Φ^{-1} = inverse function of Φ
ϕ = probability density function of a standard normal distribution

I. Introduction

UNCERTAINTY is ubiquitous in any engineering system, at any stage of product development, and throughout a product life cycle. Examples of uncertainty are manufacturing imprecision, usage variations, imperfect knowledge, and variability associated with loading, material properties, and geometric dimensions. Such uncertainties have a significant impact on product performance. A small variation in environment or design variables may lead to a significant quality loss. The ignorance of uncertainty may cause erroneous decision making, low robustness and reliability, costly warranty, low customer satisfaction, and even catastrophe [1–5]. With the intensive requirement of high product quality and reliability, understanding, identifying, and managing various uncertainties have become imperative.

Uncertainty can be viewed as the difference between the present state of knowledge and the complete knowledge (Fig. 1). It is classified into aleatory and epistemic types [6].

Aleatory uncertainty, also referred to as irreducible, objective, or stochastic uncertainty, describes the inherent variability associated with a physical system or environment [7–9]. Aleatory uncertainty is modeled by random variables or stochastic processes by probability theory if information is sufficient to estimate probability distributions. For example, for a cantilever beam in Fig. 2, aleatory uncertainty exists in the dimensions b, h, and l (due to manufacturing imprecision), external force Q (due to variations in operation), and material properties (due to the stochastic physical nature). All of the above quantities can be modeled as random variables if adequate statistical data are available. Aleatory uncertainty has been intensively researched and dealt with in a wide range of engineering fields.

Epistemic uncertainty, on the other hand, is due to the lack of knowledge about a physical system or environment [10,11]. In the above beam example, if we use different theories to calculate the stress and deflection, we may end up with different results. The reason is that each theory relies on various assumptions, which may not be completely valid. Epistemic uncertainty therefore exists in the model structure. Also, if the data of the external force Q are scarce, the distribution of Q may not be precisely known. This indicates that epistemic uncertainty may also exist in a parameter. Epistemic

Received 4 November 2006; revision received 2 March 2007; accepted for publication 11 March 2007.

*Graduate Assistant, Department of Mechanical and Aerospace Engineering, 1870 Miner Circle; jgfw4@umr.edu.

†Assistant Professor, Department of Mechanical and Aerospace Engineering, 1870 Miner Circle; dux@umr.edu (Corresponding Author).

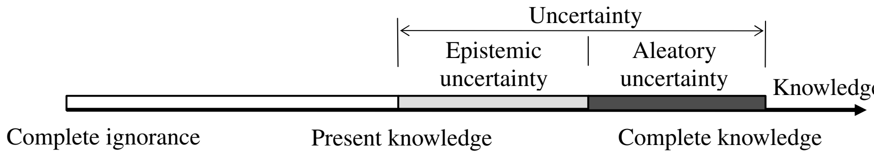

Fig. 1 Uncertainty types.

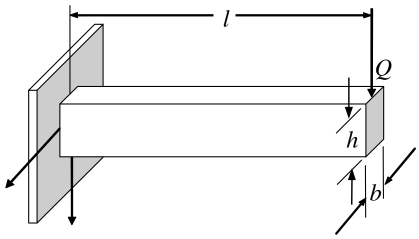

Fig. 2 A cantilever beam.

uncertainty is reducible because the collection of more information or an increase of knowledge would help decrease the level of uncertainty. In this work, we only focus on epistemic parameter uncertainty.

Different theories have been used to handle epistemic uncertainty. The theories include probability theory and nonprobability theories such as evidence theory [12], possibility theory [13,14], and fuzzy set theory [15]. Evidence theory is widely used to deal with epistemic uncertainty. Intervals with evidence theory interpretation are especially of interest in engineering applications [9]. Although there has been a longtime debate on whether probability theory is universal for handling all types of uncertainty, intervals do exist in many engineering applications, and their use is well justified in a vast amount of literature [8]. For example, for the above beam problem, the mean of the distribution of the external force Q may be given by a confidence interval with limited samples. Engineers often specify their design variables in the form of nominal value $\pm$ tolerance. More interval examples are given by Du [16] and Du et al. [17].

Evidence theory is the generalization of probability theory and possibility theory [18,19]. It can handle limited or even conflicting information. Most importantly, it is able to combine aleatory and epistemic uncertainties in a straightforward way [19,20]. Exploratory research on epistemic uncertainty by evidence theory has recently been conducted, including studies in risk assessment, decision making, and design optimization [8,14,17–25].

Most of the research focuses on uncertainty quantification and uncertainty analysis. A few investigations [18,26,27] have been conducted to explore sensitivity analysis with epistemic uncertainty. The purpose of such sensitivity analysis is to quantify the contribution of the input epistemic uncertainty to the model output. Bae et al. [18,26] develop a sensitivity analysis method for belief and plausibility measures. The method provides useful information to guide the future acquisition for more accurate reliability analysis and to reveal the most significant contributing factors in a sequential design phase. Helton et al. [27] propose a three-step sampling-based sensitivity analysis for epistemic uncertainty. In their work, an initial exploratory analysis is employed to evaluate the model behavior, and then stepwise analyses are followed to show the incremental effects of uncertain variables on belief and plausibility measures.

The above sensitivity analysis methods deal with only epistemic uncertainty. In practical engineering applications, both aleatory and epistemic uncertainties often occur simultaneously. Under this situation, a single probability measure (for instance, reliability) will not be available. Instead, its plausibility and belief measures must be used. Both of the measures will be discussed in the next section. The difference between the belief measure and plausibility measure indicates the effect of epistemic uncertainty. If the difference is too large, it will be difficult to make decisions. In this case, more information is needed to reduce the effect of epistemic uncertainty. Collecting more information on all the variables with epistemic uncertainty is costly. Collecting additional information on only the most important variables will be more efficient. Identifying variables with epistemic uncertainty that have the highest contribution to the uncertainty effect is the focus of sensitivity analysis in this paper. Because the proposed sensitivity analysis needs to quantify the uncertain characteristics of a model output given aleatory and epistemic uncertainties in model inputs, the unified uncertainty analysis framework [16] is used.

This paper is organized as follows. Brief introductions to sensitivity analysis, evidence theory, and unified uncertainty analysis are provided in Sec. II. The proposed sensitivity analysis method is discussed in Sec. III. In Sec. IV, two examples are used for demonstration. Conclusions and future work are given in Sec. V.

II. Sensitivity Analysis with Epistemic Uncertainty

A. Sensitivity Analysis

Sensitivity analysis identifies the input uncertain variables that have the highest contribution to the uncertainty in output variables. So far most of the research focuses on sensitivity analysis for aleatory uncertainty, which is mainly modeled by probability theory. Such sensitivity analysis with a probabilistic representation is usually named *probabilistic sensitivity analysis*. Various probabilistic sensitivity analysis methods have been reported in a wide range of literature, including differential analysis [28,29], variance-based methods [30], sampling-based methods [30], and a relative entropy-based method [31]. Among them, the variance-based method is popular, which derives from the decomposition of the total variance of a model output into variances due to different input variables and their combinations. The Fourier amplitude sensitivity test (FAST) [32,33], correlation ratios [34], importance measures [35], and Sobol's indices [36] belong to this type of method.

Generally, these methods work well with the probabilistic representation. However, how to apply these methods to obtain the sensitivity information from epistemic uncertainty has not been well studied.

As mentioned in the Introduction, Bae et al. [18,26] and Helton et al. [27] have conducted exploratory research on sensitivity analysis with epistemic uncertainty. In this work, we are interested in the independent epistemic variables, and our goal is to develop a new sensitivity analysis method for identifying the most important variables with epistemic uncertainty when both aleatory and epistemic uncertainties are present. We employ the unified uncertainty analysis [16] to quantify both types of uncertainty. We then perform sensitivity analysis to identify the main effect and total effect of each variable with epistemic uncertainty by the once-at-a-time (OAT) strategy [37,38] and the two-dimensional Kolmogorov–Smirnov (KS) distance [39]. Next, we provide a brief review of evidence theory and the unified uncertainty analysis.

B. Evidence Theory

Intervals are widely used to characterize epistemic uncertainty. They can be naturally handled by evidence theory [8]. A good example of intervals is the periodic monitoring [16]. Suppose the status of a system is monitored at discrete time instants $t_0, t_1, t_2, \ldots$. If a failure is detected at t_{i+1}, then the failure could occur at any time in the interval between t_i and t_{i+1}. In this case, we may not be able to determine the exact distribution of the failure time. But we can collect information to estimate the probability of the failure occurrence over each time interval. The probability assigned to an interval is defined as the basic probability assignment (BPA) in evidence theory. For example, for 20 systems, if 2 and 5 failures occurred over $[t_4, t_5]$ and $[t_9, t_{10}]$, respectively, the BPAs of intervals $[t_4, t_5]$ and $[t_9, t_{10}]$ would be $2/20 = 0.1$ and $5/20 = 0.4$, respectively.

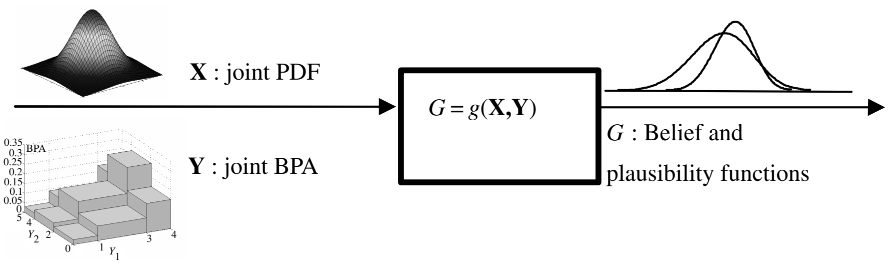

Fig. 3 The unified uncertainty analysis framework.

In this paper, we use Y to denote a variable with epistemic uncertainty. For brevity, we will call Y an epistemic variable in the remainder of the paper. We also use this same symbol Y to represent its frame of discernment, which is the sample space containing all the possible values of Y. We use $\mathcal{P}(Y)$ to denote the power set, the set that contains all the possible distinct subsets of Y. We also use A to denote an element of the power set.

In evidence theory, a BPA is a mapping function, $\mathcal{P}(Y) \to [0, 1]$, satisfying the following three axioms:

1)

$$m_Y(A) \geq 0 \quad \text{for any } A \in \mathcal{P}(Y) \tag{1}$$

2)

$$m_Y(\emptyset) = 0 \tag{2}$$

3)

$$\sum_{A \in \mathcal{P}(Y)} m_Y(A) = 1 \tag{3}$$

For two epistemic variables Y_1 and Y_2, if the change in Y_1 does not affect Y_2, and vice versa, Y_1 and Y_2 are said to be independent. Similar to the joint probability in probability theory, for two independent epistemic variables Y_1 and Y_2, their joint BPA is also used. The joint BPA is defined by

$$m_Y(C) = \begin{cases} m_{Y_1}(A) \cdot m_{Y_2}(B) & \text{when } C = A \times B \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $A \in \mathcal{P}(Y_1)$, $B \in \mathcal{P}(Y_2)$, $\mathbf{Y} = Y_1 \times Y_2$, and $C \in \mathcal{P}(Y)$. $\mathbf{Y} = Y_1 \times Y_2$ denotes the joint space of Y_1 and Y_2.

Because of the interval nature, a single probability measure is not available. Instead, two measures, belief and plausibility measures, are used in evidence theory. In this paper, we consider that the BPAs of epistemic variables are from nonconflicting items of evidence and that only one BPA exists for one interval of an epistemic variable. Under these conditions, belief and plausibility measures can be considered as the lower and upper bounds of a probability measure [40]. Let a performance G be expressed abstractly by a performance function $G = g(\mathbf{Y})$, where $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_{n_Y})$ is the vector of epistemic variables. Let an event E be defined by the performance less than a specific limit state c, namely, $E = \{\mathbf{Y} \mid g(\mathbf{Y}) < c\}$. Also let $m_{\mathbf{Y}}$ be the joint BPA over a frame $\mathbf{Y} = Y_1 \times Y_2 \times \cdots \times Y_{n_Y}$. The belief measure Bel and the plausibility measure Pl of the event $E \in \mathbf{Y}$ induced by $m_{\mathbf{Y}}$ are calculated by

$$Bel(E) = \sum_{A \in E} m_{\mathrm{Y}}(A) \tag{5}$$

and

$$Pl(E) = \sum_{A \cap E \neq \emptyset} m_{\mathrm{Y}}(A) \tag{6}$$

respectively.

$Bel(E)$ is interpreted as the degree of belief if the event E would occur. As shown in Eq. (5), it is calculated by adding the BPAs of the subsets entirely within the region $g(\mathbf{Y}) < c$. As indicated in Eq. (6), the degree of plausibility $Pl(E)$ is calculated by adding the BPAs of the subsets that are completely in the region $g(\mathbf{Y}) < c$ and the BPAs of the subsets that intersect with the region. The true probability $Pr\{g(\mathrm{Y}) < c\}$ is bounded by $Bel(E)$ and $Pl(E)$ under the abovementioned condition.

Next, we give a short review of the unified uncertainty analysis [16], which integrates probability and evidence theories to deal with the mixture of aleatory and epistemic uncertainties. The proposed sensitivity analysis relies on the unified uncertainty analysis.

C. Unified Uncertainty Analysis

A framework of unified uncertainty analysis is given in Fig. 3 [16]. The inputs to the framework are variables $\mathbf{X}$ with aleatory uncertainty defined by probability density functions (PDF) and epistemic variables $\mathbf{Y}$ represented by BPAs. Both types of uncertainty in the model inputs $\mathbf{X}$ and $\mathbf{Y}$ are propagated through the model $g(\mathbf{X}, \mathbf{Y})$ to the model output G. The outcomes of the uncertainty analysis are cumulative belief and plausibility functions (CBF and CPF).

Let the subsets of $\mathbf{Y}$ be denoted by $\mathbf{C}_{\mathbf{Y}i} (i = 1, 2, \ldots, n)$ with the corresponding joint BPA $m_Y(\mathbf{C}_{\mathbf{Y}i})$. After appropriate information aggregation [9,12], $\mathbf{C}_{Yi} (i = 1, 2, \ldots, n)$ can be disjoint. The entire input space therefore is partitioned into n mutually exclusive subsets $\mathbf{C}_{\mathbf{XY}i} = (\mathbf{X}, \mathbf{C}_{\mathbf{Y}i})$ $(i = 1, 2, \ldots, n)$. In probability theory, the cumulative distribution function (CDF) of G is defined by

$$F(c) = Pr(E) = Pr\{G = g(\mathbf{X}, \mathbf{Y}) < c\} \tag{7}$$

where F is the CDF of G at c.

Let the product space of $\mathbf{X} = X_1 \times X_2 \times \cdots \times X_{n_X}$ be discretized into k subsets (hypercubes) $\mathbf{C}_{\mathbf{X}j}$ $(j = 1, 2, \ldots, k)$ with $\Delta\mathbf{X} = \Delta X_1 \times \Delta X_2 \times \cdots \times \Delta X_{n_X}$, where ΔX_i $(i = 1, 2, \ldots, n_X)$ is the step size. Because the joint BPA of $\mathbf{C}_{\mathbf{X}j}$ is the probability of $\mathbf{X}$ in $\mathbf{C}_{\mathbf{X}j}$, the joint BPA of $\mathbf{X}$ is given by

$$m_{\mathbf{X}}(\mathbf{C}_{\mathbf{X}j}) = f_{\mathbf{X}}(\mathbf{x} \mid \mathbf{X} \in \mathbf{C}_{\mathbf{X}j})\Delta\mathbf{X} \tag{8}$$

where $f_{\mathbf{X}}(\cdot)$ is the joint PDF of $\mathbf{X}$.

The joint BPA of $\mathbf{X}$ and $\mathbf{Y}$ is then derived as

$$\begin{aligned} m_{\mathbf{XY}}(\mathbf{C}_{\mathbf{Y}i}, \mathbf{C}_{\mathbf{X}j}) &= m_{\mathbf{Y}}(\mathbf{C}_{\mathbf{Y}i}) \sum_{j=1}^{k} m_{\mathbf{X}}(\mathbf{C}_{\mathbf{X}j}) \\ &= m_{\mathbf{Y}}(\mathbf{C}_{\mathbf{Y}i}) \sum_{j=1}^{k} f_{\mathbf{X}}(\mathbf{x} \mid \mathbf{X} \in \mathbf{C}_{\mathbf{X}j})\Delta\mathbf{X} \end{aligned} \tag{9}$$

The belief measure of the failure event is then calculated by

$$\begin{aligned} Bel(c) &= \sum_{\substack{i=1 \\ (\mathbf{C}_{\mathbf{Y}i}, \mathbf{C}_{\mathbf{X}j}) \in E}}^{n} m_{\mathbf{XY}}(\mathbf{C}_{\mathbf{Y}i}, \mathbf{C}_{\mathbf{X}j}) \\ &= \sum_{\substack{i=1 \\ (\mathbf{C}_{\mathbf{Y}i}, \mathbf{C}_{\mathbf{X}j}) \in E}}^{n} \left[m_{\mathbf{Y}}(\mathbf{C}_{\mathbf{Y}i}) \sum_{j=1}^{k} m_{\mathbf{X}}(\mathbf{C}_{\mathbf{X}j}) \right] \\ &= \sum_{\substack{i=1 \\ (\mathbf{C}_{\mathbf{Y}i}, \mathbf{C}_{\mathbf{X}j}) \in E}}^{n} \left[m_{\mathbf{Y}}(\mathbf{C}_{\mathbf{Y}i}) \sum_{j=1}^{k} f_{\mathbf{X}}(\mathbf{x} \mid \mathbf{X} \in \mathbf{C}_{\mathbf{X}j})\Delta\mathbf{X} \right] \end{aligned} \tag{10}$$

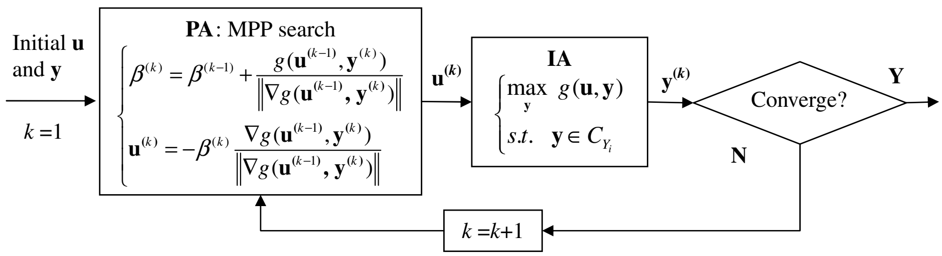

Fig. 4 Flowchart of MPP search in *Bel* calculation.

When k approaches infinity, the equation for the cumulative belief function (CBF), the degree of belief that the event $G < c$ would occur, becomes [16]

$$Bel(c) = F_G^{\min}(c) = \sum_{i=1}^{n} m_{\mathbf{Y}}(\mathbf{C}_{\mathbf{Y}i}) Pr\{G_{\max} < c \mid \mathbf{Y}_i \in \mathbf{C}_{\mathbf{Y}i}\} \quad (11)$$

By analogy, the plausibility measure function (CPF), the degree of plausibility that the event $G < c$ would occur, can be computed by

$$Pl(c) = F_G^{\max}(c) = \sum_{i=1}^{n} m_{\mathbf{Y}}(\mathbf{C}_{\mathbf{Y}i}) Pr\{G_{\min} < c \mid \mathbf{Y}_i \in \mathbf{C}_{\mathbf{Y}i}\} \quad (12)$$

respectively. $G_{\min}$ and $G_{\max}$ are, respectively, the global minimum and maximum values of G in the subset $\mathbf{C}_{\mathbf{Y}i}$ given the values of $\mathbf{X}$.

Equations (11) and (12) are derived from evidence theory by dividing the random variables into an infinite number of intervals. The same equation can also be derived from probability theory by using the total probability. See [16] for details. Equations (11) and (12) indicate that the evaluation of belief and plausibility measures with the mixture of probability distributions and BPAs is essentially the evaluation of the minimum and maximum probabilities of the performance function over the subsets of $\mathbf{Y}$. Therefore, traditional probabilistic analysis methods can be used for the unified uncertainty analysis. Hereby, we use the first-order reliability method (FORM) based uncertainty analysis method developed in [16].

D. FORM-Based Unified Uncertainty Analysis

FORM is used to calculate a CDF or the probability of failure when only random variables $\mathbf{X}$ exist. If the joint PDF of $\mathbf{X}$ is f_{X}, the probability of failure p_f is calculated by

$$p_f = F(c) = Pr\{G = g(\mathbf{X}) < c\} = \int_{g(\mathbf{X})<c} f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad (13)$$

FORM involves three steps to approximate the above integral: 1) transforming original random variables $\mathbf{X}$ to standard normal random variables $\mathbf{U}$, 2) searching the most probable point (MPP), and 3) calculating p_f.

Step 1: Transformation, which is given by

$$u_i = \Phi^{-1}\{F_{X_i}(x_i)\}, \qquad i = 1, 2, \ldots, n_X \quad (14)$$

where F_{X_i} is the CDF of X_i, and Φ^{-1} is the inverse CDF of a standard normal distribution.

Step 2: MPP search, where the MPP $\mathbf{u}^*$ is identified by

$$\min_{\mathbf{U}} \|\mathbf{U}\| \mid g(\mathbf{U}) = c \quad (15)$$

where $\|\cdot\|$ stands for the norm (length) of a vector. $\beta = \|\mathbf{u}^*\|$ is termed as a reliability index.

Step 3: Estimation of p_f, which is given by

$$p_f = \Phi(-\beta) \quad (16)$$

where Φ is the CDF of a standard normal distribution.

The key to FORM is the MPP search. The following recursive algorithm is used to search the MPP:

$$\begin{cases} \beta^{(k)} = \beta^{(k-1)} + \frac{g(\mathbf{u}^{(k-1)})}{\|\nabla g(\mathbf{u}^{(k-1)})\|} \\ \mathbf{u}^{(k)} = -\beta^{(k)} \frac{\nabla g(\mathbf{u}^{(k-1)})}{\|\nabla g(\mathbf{u}^{(k-1)})\|} \end{cases} \quad (17)$$

where $\nabla g(\mathbf{u}^{(k-1)})$ is the gradient of g at $\mathbf{u}^{(k-1)}$ and $\|\nabla g(\mathbf{u}^{(k-1)})\|$ is its magnitude, and k is the iteration counter.

The above process is called *probabilistic analysis* (PA) because only random variables are involved. As shown in Eqs. (11) and (12), we need to find the maximum and minimum values of G when interval variables $\mathbf{Y}$ exist. The process of finding the maximum and minimum G is called *interval analysis* (IA). Solving Eqs. (11) and (12) directly requires a double-loop procedure where PA and IA are nested [16]. Given a set of interval variables $\mathbf{Y}$, the MPP is searched by the algorithm in Eq. (17). Then interval analysis is performed to find the maximum and minimum performance function values with the random variables fixed at the MPP. This process repeats till convergence is reached. This double-loop procedure is computationally inefficient. To improve computational efficiency, we need to embed IA into the MPP search algorithm. In this work, we focus on black-box performance functions where closed-form functions are not applicable. Because the traditional interval arithmetic is not applicable to a black-box function, we employ nonlinear optimization to perform IA.

The flowchart for the minimum probability $Pr\{G_{\max} < c \mid \mathbf{Y}_i \in \mathbf{C}_{\mathbf{Y}i}\}$ in the CBF equation is given in Fig. 4. The solution is the MPP $\mathbf{u}^*$ where G is the maximum. The probability $Pr\{G_{\max} < c \mid \mathbf{Y}_i \in \mathbf{C}_{\mathbf{Y}i}\}$ in Eq. (11) is then computed by

$$Pr\{G_{\max} < c \mid \mathbf{Y}_i \in \mathbf{C}_{\mathbf{Y}i}\} = \Phi(-\beta) = \Phi(-\|\mathbf{u}^*\|) \quad (18)$$

For the plausibility calculation, the model of the MPP search is the same as in Fig. 4, and IA becomes a minimization problem.

III. Proposed Sensitivity Analysis Method

With only aleatory uncertainty, a single probability measure of a performance G can be obtained. With both aleatory and epistemic uncertainties, the probability bounds, belief measure and plausibility measure, can be obtained as shown in Fig. 5. The difference between belief and plausibility measures represents the effect of epistemic uncertainty. The wider the difference, the greater is the effect. If the difference is too wide, it will be difficult to make decisions.

For example, as shown in Fig. 5, the belief and plausibility are 0.016 and 0.64 at the limit state $G = 2$, respectively. If $G < 2$ is a failure event, then the minimum and maximum probabilities of failure p_f are 0.016 and 0.64, respectively. The large gap between the two bounds makes the decision process too difficult. If one used the belief ($p_f = 0.016$), the design might be highly risky because the true p_f may be much higher than the minimum value. If one used the plausibility ($p_f = 0.64$), however, the design might be too conservative. In this case, more information about the epistemic variables is needed to reduce their effect. How to effectively collect more information is critical. In this work, we develop a sensitivity analysis method to identify the most important epistemic variables that have the highest impact on design performance. With this

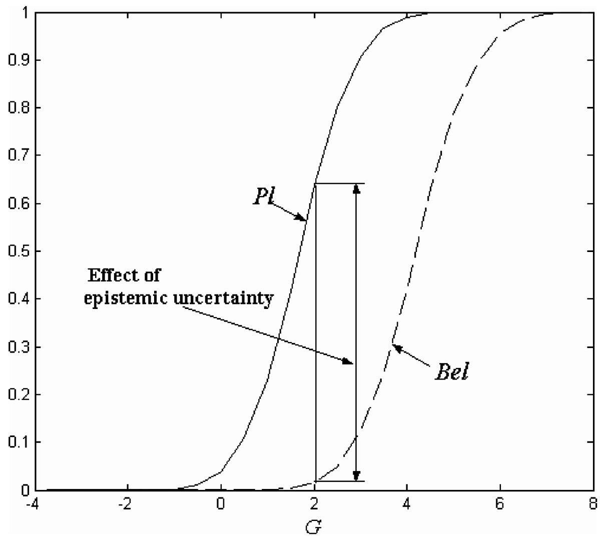

Fig. 5 Effect of epistemic uncertainty and aleatory uncertainty.

method, limited resources can be used to collect more information on the identified important epistemic variables.

We adopt the OAT strategy [38] to quantify the effect of each individual epistemic variable. The effect is measured by the difference between belief and plausibility measures. The difference is computed by the KS distance [39].

The OAT strategy belongs to the simplest class of screening methods. The impact of uncertainty in each variable is evaluated one by one [38]. The sensitivity analysis is conducted by keeping one epistemic variable while the other epistemic variables are fixed at their averages at one time. Then the impact of the varied variable on the performance can be isolated and evaluated. The average of an epistemic variable Y_j $(j = 1, 2, \ldots, n_Y)$ is calculated by

$$\bar{Y}_j = \sum_{i=1}^{n} m(C_{Y_i}) \frac{Y_{ij}^u + Y_{ij}^l}{2}, \qquad i = 1, 2, \ldots, n \tag{19}$$

where $m(C_{Y_i})$ is the BPA of the ith subset C_{Y_i}, Y_{ij}^u and Y_{ij}^l are the upper and lower bounds of Y_j on C_{Y_i}, respectively.

The KS distance is a measure used in the statistical test [39] and is defined as the maximum difference between the sample CDF and the hypothesized CDF. This distance measures how close the sample CDF to the hypothesized CDF. We adopt herein the same idea to measure the difference between CPF and CBF.

The proposed sensitivity analysis includes the following two steps:

Step 1—uncertainty analysis: the unified uncertainty analysis is performed to calculate CBF and CPF when both aleatory variables $\mathbf{X}$ and epistemic variables $\mathbf{Y}$ exist.

Step 2—OAT analysis: the main effect and the total effect of each epistemic variable are calculated. The main effect explores the impact on the performance from each single epistemic variable while the total effect measures the impact on the performance from the interactions of one epistemic variable with other epistemic variables.

To identify the main effect of the epistemic variable Y_i $(i = 1, 2, \ldots, n_Y)$, we fix the rest of the epistemic variables Y_j $(j = 1, 2, \ldots, n_Y, j \neq i)$ at their averages $\bar{Y}_j$ [see Eq. (19)]. Only Y_i is allowed to vary. To measure the total effect of the epistemic variable Y_i, we fix Y_i at its average $\bar{Y}_i$, and keep the rest of the epistemic variables Y_j $(j = 1, 2, \ldots, n_Y, j \neq i)$. After setting these different scenarios, we conduct the unified uncertainty analysis again to calculate CBF and CPF for each scenario. We then calculate the difference between CBF and CPF and rank the importance of epistemic variables by the difference.

In this work, we use sensitivity analysis for two applications, reliability analysis and uncertainty analysis for the entire range of a performance.

Application 1—reliability analysis. Let a failure mode be defined by the event where the performance G is less than a threshold c, namely, $G < c$. The probability of failure p_f can be calculated by Eq. (13) when only random variables $\mathbf{X}$ exist. When both aleatory and epistemic uncertainties are present, according to Eqs. (11) and (12), the minimum and maximum probabilities of failure are actually the CBF and CPF at c, namely,

$$p_f^{\min} = Bel(c) = F^{\min}(c) \tag{20}$$

and

$$p_f^{\max} = Pl(c) = F^{\max}(c) \tag{21}$$

The difference between $p_f^{\max}$ and $p_f^{\min}$ represents the effect of epistemic uncertainty on the probability of failure p_f. The difference is given by

$$d_{p_f} = p_f^{\max} - p_f^{\min} = Pl(c) - Bel(c) \tag{22}$$

The main effect of Y_i $(i = 1, 2, \ldots, n_Y)$ on the probability of failure is given by

$$ME_{p_f}^i = d_{p_f}^i \tag{23}$$

where $d_{p_f}^i$ is the difference between $p_f^{\max}$ and $p_f^{\min}$ when Y_i is kept as an epistemic variable and other variables Y_j $(j = 1, \ldots, n_Y, j \neq i)$ are fixed at their average $\bar{Y}_j$. ME_{pf}^i is computed by

$$\begin{aligned} ME_{\text{pf}}^i = d_{p_f}^i &= Pl(c \mid Y_j = \bar{Y}_j, j = 1, \ldots, n_Y, j \neq i) \\ &\quad - Bel(c \mid Y_j = \bar{Y}_j, j = 1, \ldots, n_Y, j \neq i) \end{aligned} \tag{24}$$

The smaller $d_{p_f}^i$ is, the weaker is the impact of Y_i on p_f, and therefore Y_i is less important.

The total effect of Y_i on d_{p_f} is given by

$$TE_{p_f}^i = d_{p_f}^{\sim i} \tag{25}$$

where $d_{p_f}^{\sim i}$ is the difference between $p_f^{\max}$ and $p_f^{\min}$ when Y_i is fixed at its average $\bar{Y}_i$ and the other variables Y_j $(j = 1, 2, \ldots, n_Y, j \neq i)$ are kept as epistemic variables. $TE_{p_f}^i$ is computed by

$$TE_{p_f}^i = d_{p_f}^{\sim i} = Pl(c \mid Y_i = \bar{Y}_i) - Bel(c \mid Y_i = \bar{Y}_i) \tag{26}$$

The smaller $d_{p_f}^{\sim i}$ means the larger influence of Y_i.

Application 2—uncertainty analysis over the entire range of the performance G. If we are interested in the effect of an epistemic variable on the entire range of the model output, we can calculate the KS distance between the CBF and CPF as follows:

$$d_{\text{KS}} = \max_c [Pl(c) - Bel(c)] \tag{27}$$

The equation implies that the KS distance is the maximum discrepancy between two curves of CBF and CPF as shown in Fig. 6.

The main effect of epistemic variable Y_i on CDF is calculated as

$$\begin{aligned} ME^i = d_{p_f}^i = \max_c [&Pl(c \mid Y_j = \bar{Y}_j, j = 1, 2, \ldots, n_Y, j \\ &\neq i) - Bel(c \\ &\mid Y_j = \bar{Y}_j, j = 1, 2, \ldots, n_Y, j \neq i)] \end{aligned} \tag{28}$$

where d_{KS}^i is the KS distance between CPF and CBF when Y_i is kept as an epistemic variable and other variables Y_j $(j = 1, 2, \ldots, n_Y, j \neq i)$ are fixed at their average $\bar{Y}_j$. The smaller d_{KS}^i is, the closer are CBF and CPF; namely, the impact of Y_i is weaker and Y_i is less influential. Therefore, the smaller ME^i is, the less significant is Y_i to the uncertainty of the performance.

The total effect of epistemic variable Y_i on CDF can be calculated as

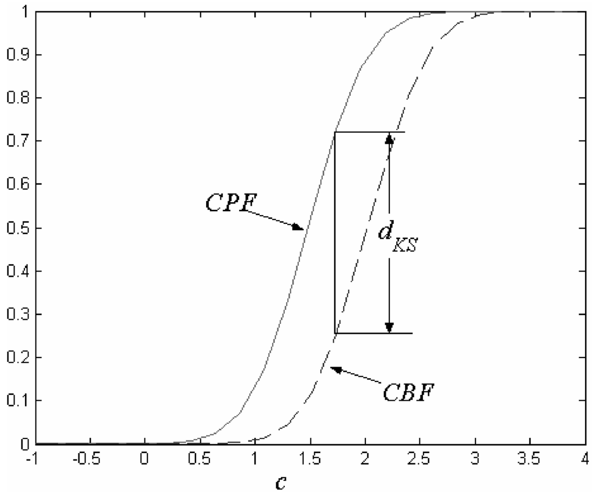

Fig. 6 KS distance.

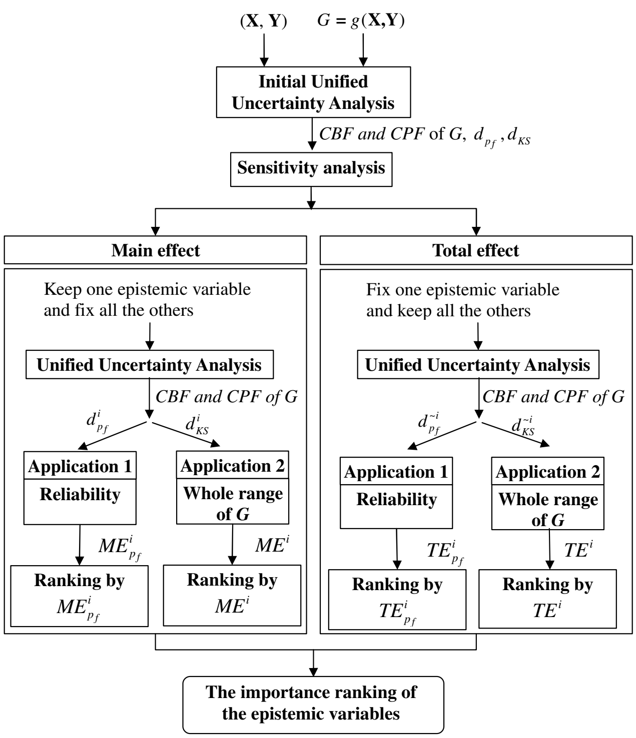

Fig. 7 Flowchart of the proposed sensitivity analysis.

$$TE^i = d^{\sim i}_{\mathrm{KS}} = \max_c [Pl(c \mid Y_i = \bar{Y}_i) - Bel(c \mid Y_i = \bar{Y}_i)] \quad (29)$$

where $d^{\sim i}_{\mathrm{KS}}$ is the KS distance when Y_i is fixed at its average $\bar{Y}_i$ and other variables Y_j $(j = 1, 2, \ldots, n_Y, j \neq i)$ are kept as epistemic variables. In this case, a smaller discrepancy between CPF and CBF implies higher influence of Y_i on G.

The flowchart of the proposed sensitivity analysis method is illustrated in Fig. 7.

From the above discussion, it is seen that one sensitivity analysis needs to call the unified analysis $2n_Y + 1$ times—one analysis is for the case with original uncertain variables, n_Y analyses are for the main effects of the n_Y epistemic variables, and the other n_Y analyses are for the total effects of the n_Y epistemic variables. The computation is intensive, and therefore efficiency is critical. To improve efficiency, we use the efficient MPP algorithm as shown in Eq. (17). In many engineering applications, a performance function is monotonic in terms of interval variables. In this case, it is not necessary to conduct nonlinear optimization for the interval analysis. However, it is difficult to know whether the performance function is monotonic because of the black-box model. We therefore perform optimization for the interval analysis in the first iteration. Thereafter, we check the Karush–Kuhn–Tucker (KKT) conditions [41] after the MPP is updated. If the KKT conditions are satisfied, there is no need to perform optimization again. We then proceed to the next iteration.

IV. Examples

A. Example 1: Crank–Slider Mechanism

A crank–slider mechanism is used in a construction machine as shown in Fig. 8 [16]. The length of the crank AB a, the length of the coupler BC b, the external force Q, the Young's modulus of the material of the coupler E, and the yield strength of the coupler S are random variables. Their distributions are given in Table 1.

Because of the harsh environment of the construction site, a precise distribution of the coefficient of friction μ between the ground NN and the slider C is not available, but its intervals and BPA are available based on the solicitation from experts. Because different installation positions of the slider are required in various construction sites, the intervals and BPA of the offset e are assigned based on limited historical data. Their BPAs are provided in Table 2, and the joint BPA is also visualized in Fig. 9.

The two performance functions are the safety margins for strength and buckling requirements of the coupler, which are defined by the difference between the material strength and the maximum stress, and the difference between the critical load and the axial load, respectively. The equations are obtained at one of the positions when the crank AB and the coupler BC overlap. The functions are given by

$$G_1 = g_1(\mathbf{X}, \mathbf{Y}) = S - \frac{4P(b-a)}{\pi\left(\sqrt{(b-a)^2 - e^2} - \mu e\right)\left(d_2^2 - d_1^2\right)}$$

and

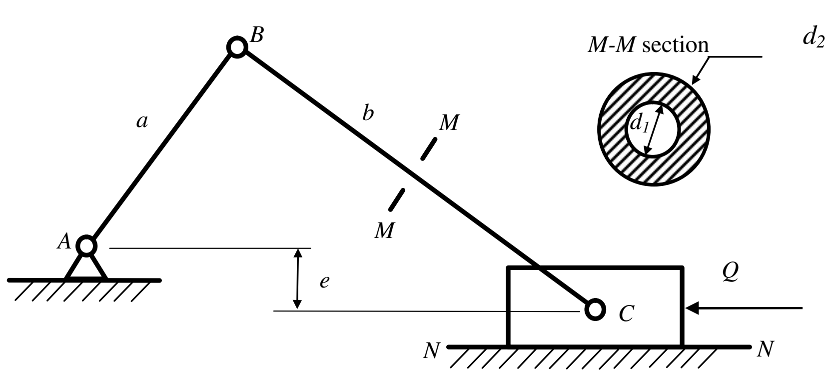

Fig. 8 A crank–slider mechanism.

Table 1 Random variables X

Variables	Symbols in Fig. 8	Mean	Standard deviation	Distribution
X_1	a	100 mm	0.01 mm	Normal
X_2	b	300 mm	0.01 mm	Normal
X_3	Q	250 kN	25 kN	Normal
X_4	E	200 GPa	30 GPa	Normal
X_5	S	390 MPa	39 MPa	Normal

Table 2 Uncertain variables with epistemic uncertainty

Variables	Symbols in Fig. 8	Intervals	BPA
		[100, 120]	0.2
Y_1	e, mm	[120, 140]	0.4
		[140, 150]	0.4
		[0.15, 0.18]	0.3
Y_2	μ	[0.18, 0.23]	0.3
		[0.23, 0.25]	0.4

$$G_2 = g_2(\mathbf{X}, \mathbf{Y}) = \frac{\pi^3 E\left(d_2^4 - d_1^4\right)}{64b^2} - \frac{P(b-a)}{\sqrt{(b-a)^2 - e^2 - \mu e}}$$

The failure events are defined by $E_1 = \{\mathbf{X}, \mathbf{Y} \mid G_1 < 0\}$ and $E_2 = \{\mathbf{X}, \mathbf{Y} \mid G_2 < 0\}$. Our goal is to find out the most significant epistemic variable (offset e or coefficient of friction μ) which has the most dominant effect on the performance functions G_1 and G_2.

We first perform the unified uncertainty analysis for the two failure modes. The result is given in Table 3. The difference d_{p_f} between the maximum and minimum probabilities of failure (or Pl and Bel) of G_1 is large, and the difference d_{p_f} of G_2 is almost zero. Therefore, the effect of epistemic uncertainty on failure mode 1 (G_1) cannot be neglected, and the effect of epistemic uncertainty on failure mode 2 (G_2) is negligible. Sensitivity analysis on failure mode 1 is then necessary. Hence we only conduct sensitivity analysis on G_1.

To confirm the accuracy of the united uncertainty analysis, we solve the problem by Monte Carlo simulation (MCS). The result is also provided in Table 3, where N is the number of function evaluations. N is used to measure computational efficiency. It is seen that the unified uncertainty analysis employed in this paper is very accurate and efficient.

We also perform the unified uncertainty analysis for the entire range of the two performance functions. The results of CBF and CPF for both G_1 and G_2 are shown in Table 4 and Figs. 10 and 11. It is also seen that the effect of epistemic uncertainty on G_1 is much larger than that on G_2 because the KS distance for G_1 is much larger than that for G_2. The numbers of function evaluations also indicate that the unified uncertainty analysis is more efficient than MCS.

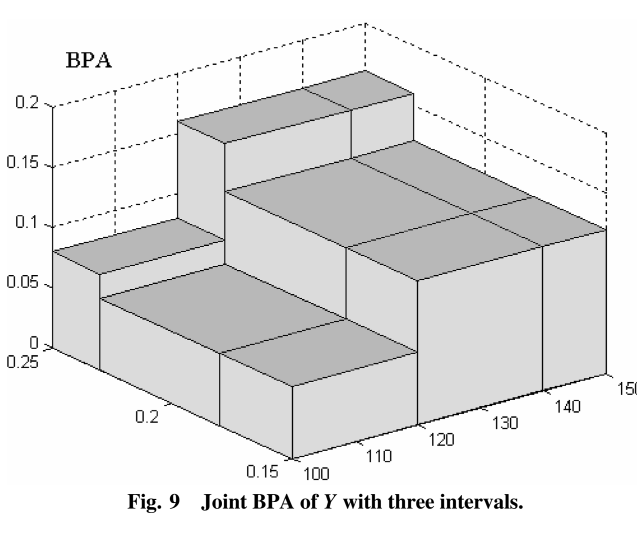

Fig. 9 Joint BPA of Y with three intervals.

Next we perform sensitivity analysis on G_1 to find out the most influential epistemic variable. In this example, there are only two epistemic variables Y_1 and Y_2; no total effect is therefore needed. Thus we only analyze the main effect of each variable.

Main effect of Y_1: Keep Y_1 as an epistemic variable and fix Y_2 at its average. The average of Y_2 is calculated by

$$\bar{Y}_2 = \frac{0.15 + 0.18}{2} \times 0.3 + \frac{0.18 + 0.23}{2} \times 0.3 + \frac{0.23 + 0.25}{2} \times 0.4 = 0.207$$

The CBF and CPF of G_1 are reevaluated by the unified uncertainty and are given in Fig. 12. The difference (main effect) between the maximum and minimum probabilities of failure $ME_{p_f}^1 = d_{p_f}^1$ and the KS distance for the entire distribution $ME^1 = d_{\mathrm{KS}}^1$ are given in Table 5.

Main effect of Y_2: Keep Y_2 as an epistemic variable and fix Y_1 at its average. The average of Y_1 is calculated by

$$\bar{Y}_1 = \frac{100 + 120}{2} \times 0.2 + \frac{120 + 140}{2} \times 0.4 + \frac{140 + 150}{2} \times 0.4 = 132$$

The CBF and CPF of G_1 are illustrated in Fig. 13, and $ME_{p_f}^2 = d_{p_f}^2$ and $ME^2 = d_{\mathrm{KS}}^2$ are also given in Table 5. The difference between the CBF and CPF is much narrower when Y_1 is fixed. The result indicates that the main effect of Y_1 is much greater than that of Y_2. Therefore Y_1 is the most influential contributor to the effect of epistemic uncertainty on the probability of failure p_f of G_1, and it is also true for the entire range of G_1.

If more information is needed to reduce the effect of epistemic uncertainty, we should collect more information on Y_1 instead of Y_2. After adequate information was collected on Y_1, Y_1 would become a random variable with only aleatory uncertainty. Suppose the available distribution of Y_1 is $N(125, 8.33)$ mm. Through the unified uncertainty analysis again, the gap between CBF and CPF of G_1 becomes much narrower as shown in Fig. 14 and Table 6.

If we did not conduct a sensitivity analysis, we might arbitrarily choose to collect more information on Y_2. Suppose the distribution of Y_2 is $N(0.2, 0.017)$ after more information is collected. As shown in Fig. 15 and Table 6, the reduced effect of epistemic uncertainty (the gap between the CPF and CBF) is much less significant compared to the situation where the epistemic uncertainty of Y_1 is eliminated (see Fig. 14).

For an easy comparison, all the information is provided in Table 6. The first row is the uncertainty analysis result with the original uncertain variables $\mathbf{X}$ and $\mathbf{Y}$. The second row is the result when Y_1 becomes aleatory while the third row is the result when Y_2 becomes aleatory. The table verifies that Y_1 is more important than Y_2 in terms of the effect on G_1.

B. Example 2: Crowned Cam Roller–Follower's Contact

A crowned cam roller/follower used in a transmission system is shown in Fig. 16 [42]. It has a gentle radius transverse to its rolling direction for eliminating the need for critical alignment of its axis with that of the cam. The roller radius is R_1 with a R_1' crown radius at 90 deg to the roller radius. The cam's radius of curvature at the point of maximum load is R_2 and is flat axially so its crown radius R_2' is infinite. The rotational axes of the cam and roller are parallel. The force is Q, normal to the contact plane. The materials of the roller and the follower are steel. Their Young's modulus E is 30×10^6 psi, and the Poisson's ratio ν is 0.28. Because of the elastic deformation, the contact patch is an ellipse, and the pressure distribution is a semiellipsoid, as illustrated in Fig. 17. R_1, R_1', and R_2 are random variables with the distributions listed in Table 7.

Table 3 d_{p_f} for G_1 and G_2

	United uncertainty analysis				Monte Carlo simulation			
	$p_f^{\min}$ (Bel)	$p_f^{\max}$ (Pl)	d_{p_f}	N	$p_f^{\min}$ (Bel)	$p_f^{\max}$ (Pl)	d_{p_f}	N
G_1	0.007353	0.072707	0.065354	468	0.012453	0.094238	0.081785	4×10^6
G_2	≈0.0	≈0.0	≈0.0	488	≈0.0	≈0.0	≈0.0	4×10^6

Table 4 Results of unified uncertainty analysis for G_1

G_1	United uncertainty analysis		Monte Carlo simulation	
	CPF–CBF	N	CPF–CBF	N
−100	0.0008	468	0.0013	36×10^6
−50	0.0103	468	0.0156	36×10^6
0	0.0654	468	0.0818	36×10^6
50	0.2042	468	0.2050	36×10^6
100	0.3248	468	0.2804	36×10^6
150	0.2594	468	0.2322	36×10^6
200	0.0946	468	0.1017	36×10^6
250	0.0137	468	0.0181	36×10^6
300	0.0007	468	0.0011	36×10^6
	Total function calls	4212	Total function calls	36×10^6
	KS distance	0.3248	KS distance	0.2804

Because of limited information, an accurate measure or a distribution of Q is not available. Its intervals and BPA are available based on the solicitation from experts, which is provided in Table 8.

The half-width of the major axis a and the half-width of the minor axis b are determined by

$$a = k_a \sqrt[3]{\frac{3Q(m_1 + m_2)}{4A}}, \qquad b = k_b \sqrt[3]{\frac{3Q(m_1 + m_2)}{4A}}$$

where

$$m_1 = m_2 = \frac{1 - v^2}{E^2}, \qquad A = \frac{1}{2}\left(\frac{1}{R_1} + \frac{1}{R_1'} + \frac{1}{R_2} + \frac{1}{R_2'}\right)$$

The factors k_a and k_b are obtained from a table in [42] based on the value of ϕ, which is calculated by

$$\phi = \cos^{-1}\left(\frac{A}{B}\right)$$

$$B = \frac{1}{2}\left[\left(\frac{1}{R_1} - \frac{1}{R_1'}\right)^2 + \left(\frac{1}{R_2} - \frac{1}{R_2'}\right)^2 + 2\left(\frac{1}{R_1} - \frac{1}{R_1'}\right)\left(\frac{1}{R_2} - \frac{1}{R_2'}\right)\right]^{\frac{1}{2}}$$

ϕ is a random variable due to the randomness in R_1, R_1', and R_2, and k_a and k_b are tabulated in terms of ϕ. The values of k_a and k_b are not precisely known and are estimated within two intervals as shown in Table 8.

The performance function is the safety margin for shear yield strength of the roller and follower, defined by the difference between the shear yield strength and maximum shear stress, which is one-half of the tensile yield strength based on maximum shear-stress theory. The function is given by

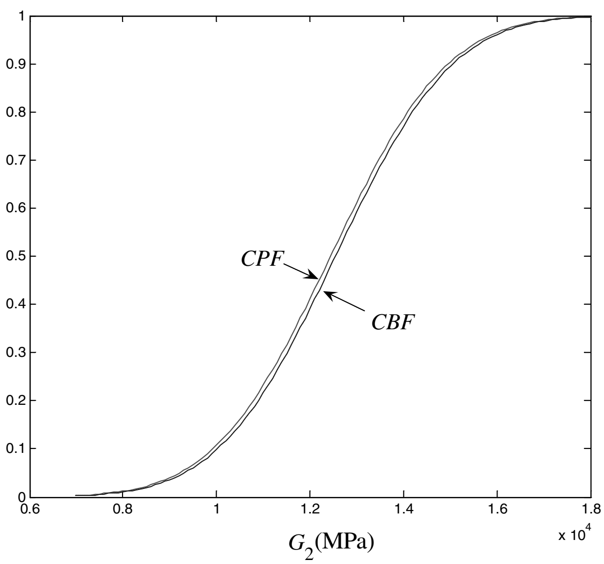

Fig. 11 Initial unified uncertainty analysis for G_2.

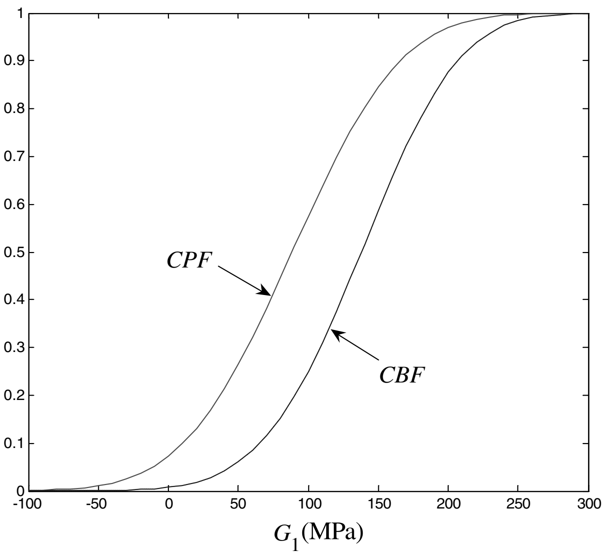

Fig. 10 Initial unified uncertainty analysis for G_1.

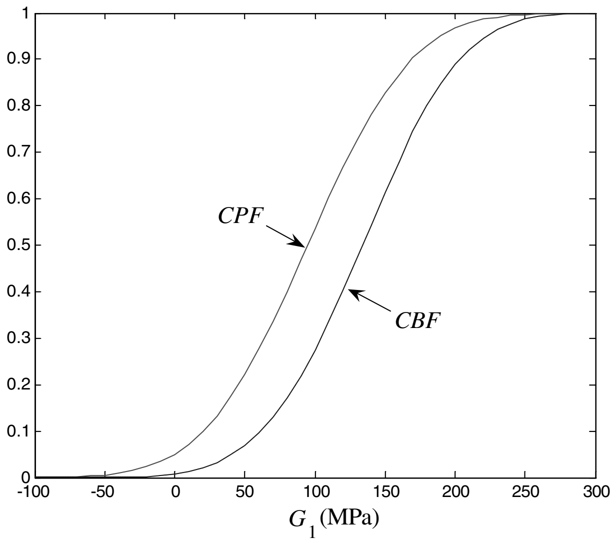

Fig. 12 CBF and CPF from the main effect analysis for Y_1.

Table 5 Main effect of each epistemic variable

Main effect	$p_f^{\min}$ (Bel)	$p_f^{\max}$ (Pl)	ME_{p_f}	ME
Y_1	0.008219	0.049754	0.041535	0.2979
Y_2	0.004638	0.008164	0.003527	0.0761

$$G = g(\mathbf{X}, \mathbf{Y}) = \tau_{\max} - S$$

where S is the shear yield strength, and $\tau_{\max}$ is defined by

$$\tau_{\max} = \max(\tau_a, \tau_u, \tau_{\text{ma}}, \tau_{\text{mi}})$$

where τ_a is the maximum shear stress at the contact surface, τ_u is the largest shear stress under the contact surface, and τ_{ma} and τ_{mi} are the shear stresses at the ends of the major and minor axes, respectively, on the contact surface. These stresses are calculated by

$$\tau_a = \left|\frac{\sigma_1 - \sigma_3}{2}\right|$$

where σ_1, σ_2, and σ_3 are principal stresses at the contact surface, calculated by

$$\sigma_1 = -\left[2v + (1-2v)\frac{b}{a+b}\right]p_{\max}$$

$$\sigma_2 = -\left[2v + (1-2v)\frac{a}{a+b}\right]p_{\max}$$

$$\sigma_3 = -p_{\max}$$

where

$$p_{\max} = \frac{3Q}{2\pi ab}$$

and

$$\tau_u = 0.34 p_{\max}$$

$$\tau_{\text{ma}} = (1-2v)\frac{k_3}{k_4^2}\left(\frac{1}{k_4}\tanh^{-1}k_4 - 1\right)p_{\max}$$

where

$$k_3 = \frac{b}{a}$$

$$\tau_{\text{mi}} = (1-2v)\frac{k_3}{k_4^2}\left(1 - \frac{k_3}{k_4}\tanh^{-1}\left(\frac{k_4}{k_3}\right)\right)p_{\max}$$

where

$$k_4 = \frac{1}{a}\sqrt{a^2 - b^2}$$

The failure event is defined by $E = \{\mathbf{X}, \mathbf{Y} \mid G < 0\}$. The CBF and CPF are calculated by the unified uncertainty analysis, and their difference is shown in Fig. 18. Also, a comparison with MCS is provided in Table 9. The result indicates that the impact of epistemic uncertainty is large and cannot be ignored. In the table, N is the number of function evaluations.

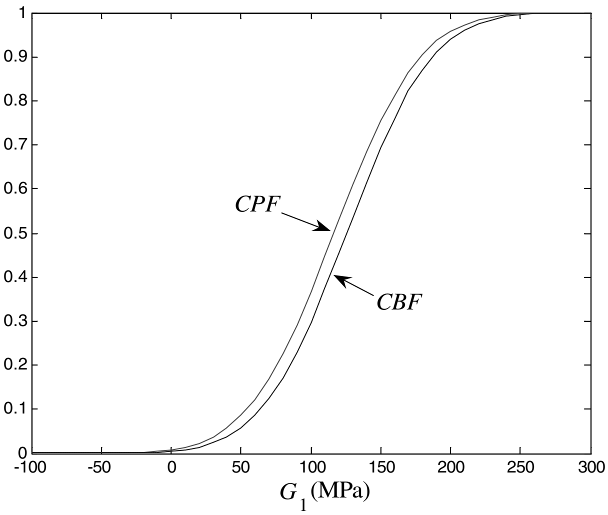

Fig. 13 ***CBF*** **and** ***CPF*** **from the main effect analysis for** $\boldsymbol{Y_2}$**.**

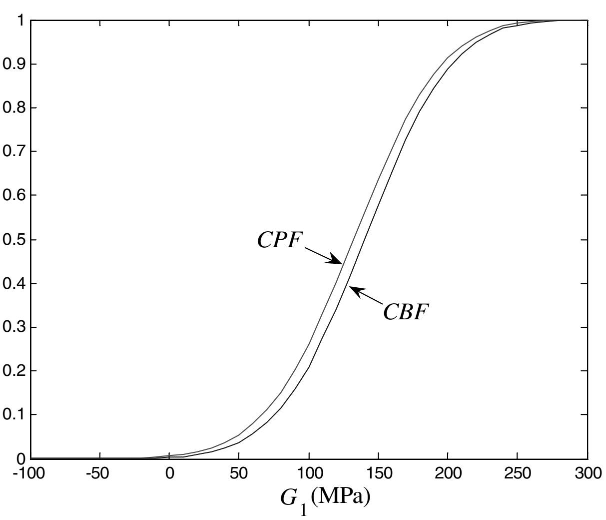

Fig. 14 ***CPF*** **and** ***CBF*** **of** $\boldsymbol{G_1}$ **after more information on** $\boldsymbol{Y_1}$ **is collected.**

Sensitivity analysis is then conducted to identify the most influential variables. Total and main effects of each variable can be obtained, respectively.

We conduct the main effect analysis as follows:

Main effect of Y_1: Keep Y_1 as an epistemic variable and fix Y_2 and Y_3 at their averages.

Main effect of Y_2: Keep Y_2 as an epistemic variable and fix Y_1 and Y_3 at their averages.

Main effect of Y_3: Keep Y_3 as an epistemic variable and fix Y_1 and Y_2 at their averages.

The respective results for the above three analyses are found in Figs. 19–21 and Table 10. As for the difference between the maximum and the minimum probabilities of failure, because $ME_{p_f}^1 = d_{p_f}^1$ is the largest, Y_1 therefore has the highest impact on p_f.

Table 6 $ME_{p_f}^i$ and ME^i

Scenarios	$p_f^{\min}$ (Bel)	$p_f^{\max}$ (Pl)	$ME_{p_f}^i$	ME^i
Y_1: 3 intervals, Y_2: 3 intervals	0.007353	0.072707	0.065354	0.3288
Y_1: aleatory, Y_2: 3 intervals	0.003379	0.005881	0.002502	0.0622
Y_1: 3 intervals, Y_2: aleatory	0.007507	0.045416	0.037909	0.2625

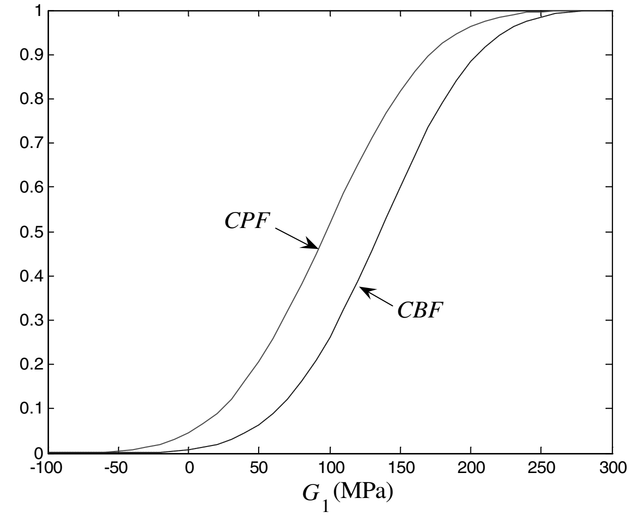

Fig. 15 ***CPF* and *CBF* of G_1 after more information on Y_2 is collected.**

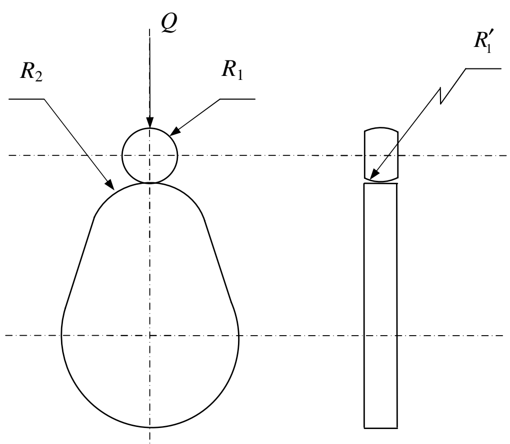

Fig. 16 A crowned cam roller/follower under load *Q*.

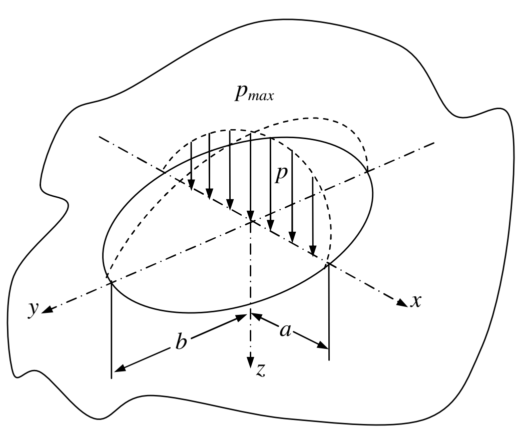

Fig. 17 Contact patch (*p* denotes the pressure distributed on the contact patch).

As for the KS distance between the CBF and CPF, because $ME^1 = d_{\mathrm{KS}}^1$ is the largest, Y_1 is also the most influential epistemic variable to the effect of epistemic uncertainty on G.

We next perform the total effect analysis as follows:

Total effect of Y_1: Keep Y_2 and Y_3 as epistemic variables and fix Y_1 at its average.

Total effect of Y_2: Keep Y_1 and Y_3 as epistemic variables and fix Y_2 at its average.

Total effect of Y_3: Keep Y_1 and Y_2 as epistemic variables and fix Y_3 at its average.

Table 7 Random variables *X*

Variable	Symbols in problem	Mean	Standard deviation	Distribution
X_1	R_1	1 in.	0.01 in.	Normal
X_2	R_1'	20 in.	0.2 in.	Normal
X_3	R_2	3.46 in.	0.0346 in.	Normal
X_4	S	37.5 ksi	0.375 ksi	Normal

Table 8 Uncertain variables with epistemic uncertainty

Variable	Symbols in problem	Intervals	BPA
Y_1	k_a	[3.50, 3.60]	1
Y_2	k_b	[0.434, 0.440]	1
Y_3	Q, lb	[246, 254]	1

The results are given in Figs. 22–24 and Table 11. It can be seen that $TE_{p_f}^1 = d_{p_f}^{\sim 1}$ and $TE^1 = d_{\mathrm{KS}}^{\sim 1}$ are smallest, and therefore Y_1 is most influential, which is consistent with the conclusion from the main effect analysis.

To confirm the above sensitivity analyses results, we assume that more information could be collected on Y_1, Y_2, and Y_3. The distributions of Y_1, Y_2, and Y_3 after gathering more information are $N(3.55, 0.036)$, $N(0.437, 0.0044)$, and $N(250, 2.5)$, respectively. The unified uncertainty analysis is performed when one epistemic variable becomes aleatory. The CBF and CPF of G are shown in Fig. 25, and ME_{p_f} and ME in each case are provided in Table 12. It is seen that from Fig. 25 the gap between CBF and CPF becomes narrowest after the epistemic uncertainty in Y_1 is eliminated. The result confirms that collecting more information on the most influential variable Y_1 has the highest contribution for reducing the effect of epistemic uncertainty. The second highest contribution is from gathering more information on Y_2. Collecting more information on Y_3 has the least contribution.

It is seen that after the epistemic uncertainty of the most important variable Y_1 has been eliminated, the gap between CBF and CPF is still large. The elimination of the epistemic uncertainty of one more variable may be needed. Because Y_2 is more important than Y_3, more information on Y_2 should be collected if further action needs to be taken.

V. Conclusions

An effective sensitivity analysis method is developed to identify the most important input variables with epistemic uncertainty when

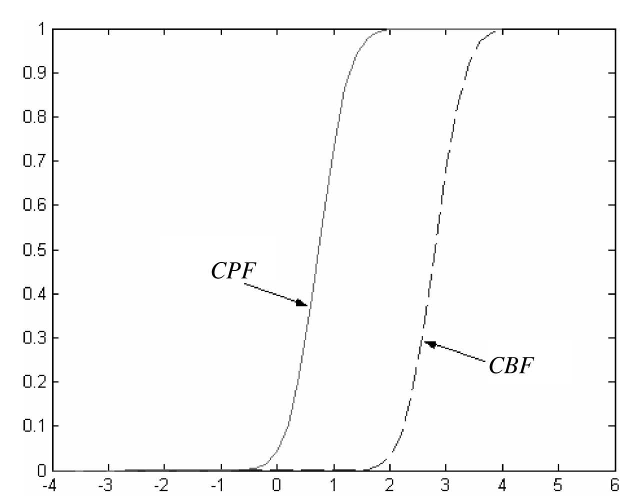

Fig. 18 Initial unified uncertainty analysis of *G*.

Table 9 Results of unified uncertainty analysis

G, ksi	United uncertainty analysis		Monte Carlo simulation	
	CPF–CBF	*N*	*CPF–CBF*	*N*
−1	0	102	0	88×10^6
−0.50	0.0018	88	0.0018	88×10^6
0	0.0412	88	0.0407	88×10^6
0.5	0.2890	83	0.2869	88×10^6
1	0.7331	83	0.7309	88×10^6
1.5	0.9627	84	0.9621	88×10^6
2.0	0.9675	84	0.9679	88×10^6
2.5	0.7575	83	0.7593	88×10^6
3.0	0.3205	83	0.3229	88×10^6
3.5	0.0517	88	0.0524	88×10^6
4.0	0.0027	88	0.0027	88×10^6
	Total function calls	954	Total function calls	88×10^6

aleatory uncertainty also exists. The importance of an epistemic variable is measured by its effect on the model output, including its main effect and total effect. These effects are indicated by the difference between belief and plausibility measures of an output variable. After the sensitivity analysis, all the epistemic variables are ranked by their importance. Then by collecting more information on the dominant epistemic variables, the effect of epistemic uncertainty can be reduced in the most efficient way as shown in the paper.

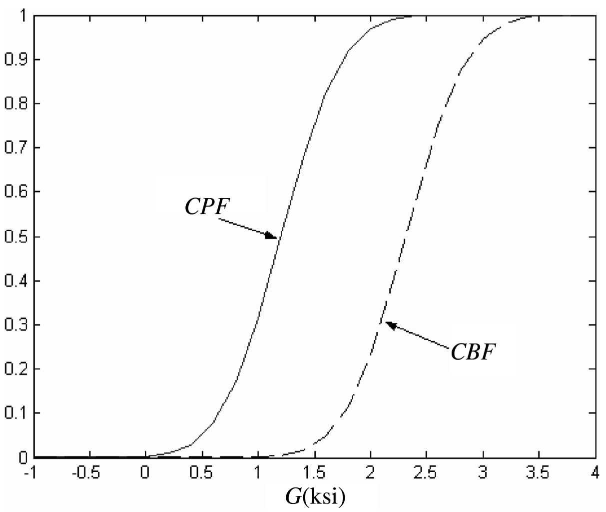

Fig. 19 ***CBF*** **and** ***CPF*** **from the main effect analysis for** Y_1**.**

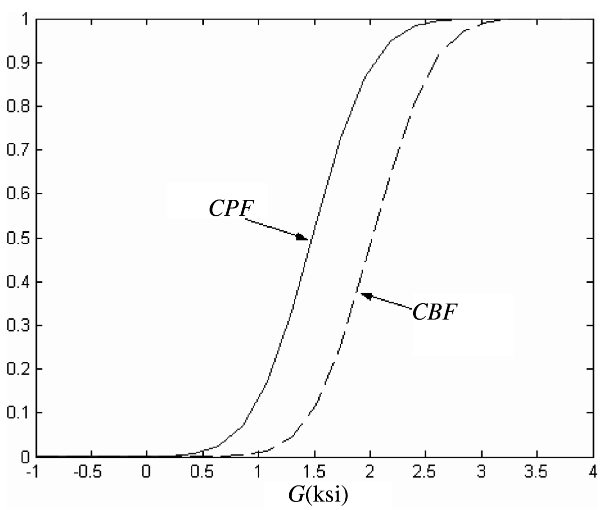

Fig. 20 ***CBF*** **and** ***CPF*** **from the main effect analysis for** Y_2**.**

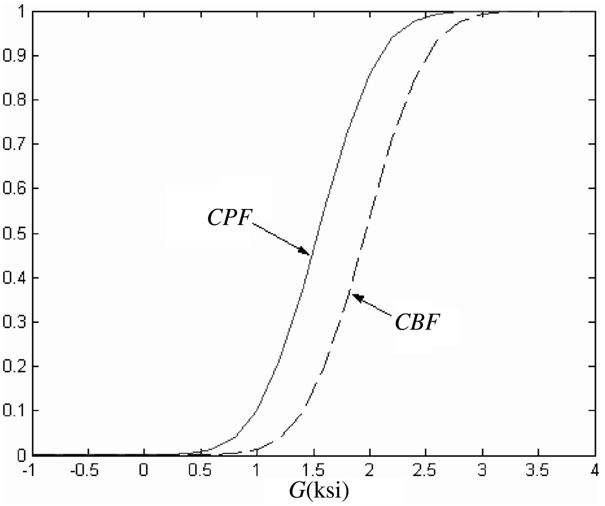

Fig. 21 ***CBF*** **and** ***CPF*** **from the main effect analysis for** Y_3**.**

Table 10 Main effect of each epistemic variable

Main effect of	$p_f^{\min}$ (*Bel*)	$p_f^{\max}$ (*Pl*)	ME_{p_f}	*ME*
Y_1	2.91*E*-08	0.002245	0.002245	0.80278
Y_2	1.01*E*-06	0.00024	0.000239	0.47207
Y_3	2.08*E*-06	0.000144	0.000142	0.37393

In the proposed sensitivity analysis procedure, a once-at-a-time strategy is used to set up different scenarios for the input epistemic variables to study their main effects and total effects. Then plausibility and belief measures of an output variable are calculated under each scenario by the unified uncertainty analysis framework. The Kolmogorov–Smirnov distance is used to quantify the discrepancy between the plausibility measure and belief measure, namely, the effect of epistemic uncertainty on the output. By comparing the main effects and total effects of the epistemic variables, their importance is ranked.

The proposed sensitivity analysis method is based on the first-order reliability method. The advantages of the proposed methods are as follows: 1) engineers are familiar with the first-order reliability method; 2) it is easy to quantify the contributions of individual epistemic variables to the reliability or to the probability of failure; 3) because optimization is used for interval analysis, the result in general is more accurate than that from interval arithmetic; 4) the

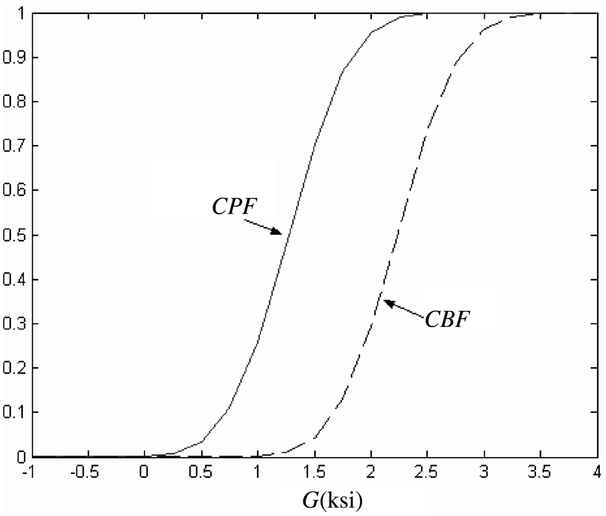

Fig. 22 ***CBF*** **and** ***CPF*** **from the total effect analysis for** Y_1**.**

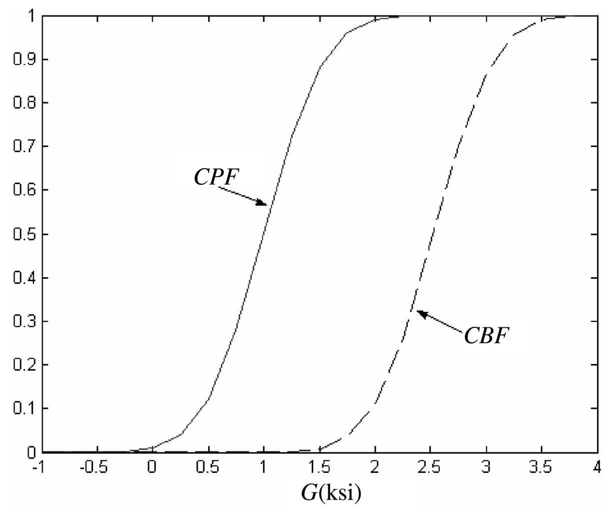

Fig. 23 **CBF and CPF from the total effect analysis for Y_2.**

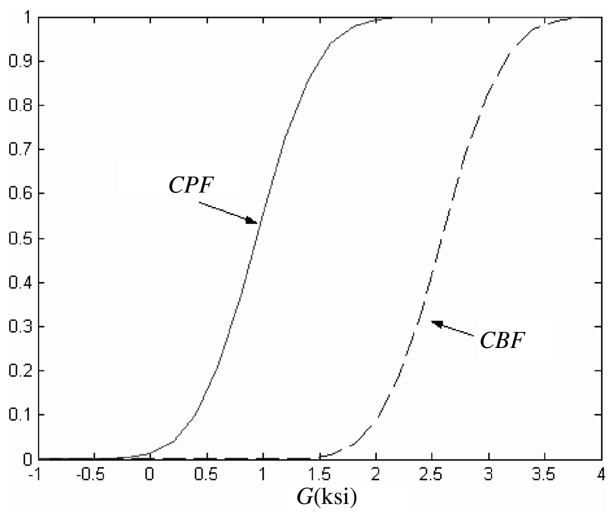

Fig. 24 **CBF and CPF from the total effect analysis for Y_3.**

process is efficient because the double-loop Monte Carlo simulation is not involved; 5) the proposed method is applicable to black-box models.

When using the proposed method, one should also consider the other features of the method. 1) The method assumes the global optimal solution for the interval analysis. The method may not provide an accurate solution if a global optima is not reached. 2) The efficiency of the method depends on the number of subsets of the epistemic variables because the first-order reliability method is performed for each subset. The efficiency also depends on the number of aleatory variables because the efficiency of the first-order reliability method is directly proportional to the number of aleatory (random) variables.

Compared to the traditional probabilistic sensitivity analysis, sensitivity analysis with the mixture of epistemic and aleatory uncertainties is much more computationally expensive. Our future work will be the improvement of computational efficiency. We will also study the sensitivity of aleatory uncertainty and its interaction with epistemic uncertainty.

Table 11 Total effect of each epistemic variable

Total effect of	p_f^{min} (Bel)	p_f^{max} (Pl)	TE_{p_f}	TE
Y_1	8.08*E*-08	0.00134	0.00134	0.73815
Y_2	1.66*E*-09	0.009275	0.009275	0.92589
Y_3	6.65*E*-10	0.013141	0.013141	0.94513

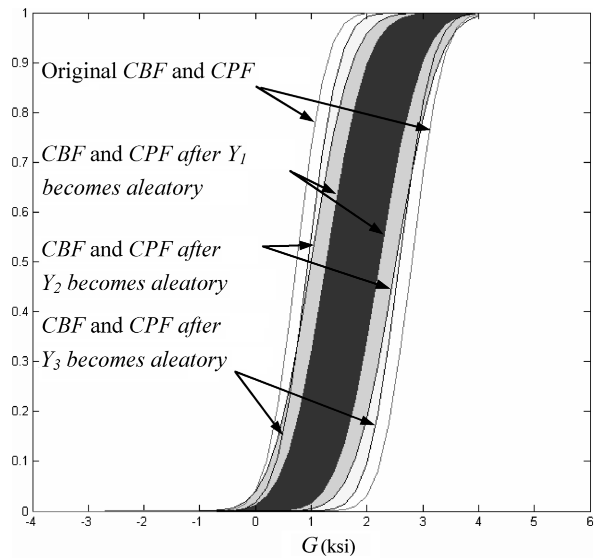

Fig. 25 Comparison of uncertainty effect.

Table 12 Unified uncertainty analysis for confirmation

Scenarios	p_f^{min} (Bel)	p_f^{max} (Pl)	$ME^i_{p_f}$	ME^i
Y_1, Y_2, and Y_3 are aleatory	2.74*E*-11	0.041172	0.041172	0.7836
Y_1 is aleatory	5.12*E*-05	0.012733	0.012682	0.5106
Y_2 is aleatory	5.89*E*-06	0.039711	0.039705	0.81014
Y_3 is aleatory	3.56*E*-09	0.0167	0.0167	0.93361

Acknowledgments

We are grateful for the partial support from the U.S. National Science Foundation Grant CMMI-040081 and the Intelligent Systems Center at the University of Missouri—Rolla. The authors would also like to thank the three anonymous reviewers for their valuable comments and suggestions.

References

[1] Oberkampf, W. L., Diegert, K. V., Alvin, K. F., and Rutherford, B. M., "Variability, Uncertainty, and Error in Computational Simulation," *AIAA/ASME Joint Thermophysics and Heat Transfer Conference*, ASME-HTD, AIAA, Reston, VA, 1998, Vol. 357, No. 2, pp. 259–272.

[2] Du, X., and Chen, W., "Methodology for Managing the Effect of Uncertainty in Simulation-Based Design," *AIAA Journal*, Vol. 38, No. 8, 2000, pp. 1471–1478.

[3] Du, X., and Chen, W., "Towards a Better Understanding of Modeling Feasibility Robustness in Engineering Design," *Journal of Mechanical Design*, Vol. 122, No. 4, 2000, pp. 357–583.

[4] Zang, T. A., Hemsch, M. J., Hilburger, M. W., Kenny, S. P., Luckring, J. M., Maghami, P., Padula, S. L., and Stroud, W. J., "Needs and Opportunities for Uncertainty-Based Multidisciplinary Design Methods for Aerospace Vehicles," NASA TM-2002-211462, 2002.

[5] Hong-Zhong, H., "Reliability Analysis Method in the Presence of Fuzziness Attached to Operating Time," *Microelectronics and Reliability*, Vol. 35, No. 12, 1995, pp. 1483–1487.

[6] Hoffman, F. O., and Hammonds, J. S., "Propagation of Uncertainty in Risk Assessment: The Need to Distinguish Between Uncertainty due to Lack of Knowledge and Uncertainty due to Variability," *Risk Analysis*, Vol. 14, No. 5, 1994, pp. 707–712.

[7] Oberkampf, W., Helton, J., and Sentz, K., "Mathematical Representation of Uncertainty," *42nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference and Exhibit*, 2001AIAA Non-Deterministic Approaches Forum, AIAA 2001-1645, 2001.

[8] Helton, J. C., Johnson, J. D., and Oberkampf, W. L., "An Exploration of Alternative Approaches to the Representation of Uncertainty in Model Predictions," *Reliability Engineering and System Safety*, Vol. 85, Nos. 1–3, 2004, pp. 39–71.
[9] Oberkampf, W. L., and Helton, J. C., "Investigation of Evidence Theory for Engineering Applications," *43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 4th Non-Deterministic Approaches Forum, AIAA Paper 2002-1569, 2002.
[10] Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., and Ferson, S., "Challenge Problems: Uncertainty in System Response Given Uncertain Parameters," *Reliability Engineering and System Safety*, Vol. 85, Nos. 1–3, 2004, pp. 11–19.
[11] Ayyub, B. M., and Chao, R. U., "Uncertainty Modeling in Civil Engineering with Structural and Reliability Applications," *Uncertainty Modeling and Analysis in Civil Engineering*, edited by Ayyub, B. M., CRC Press, Boca Raton, FL, 1997.
[12] Shafer, G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.
[13] Klir, G. J., and Wierman, M. J., *Uncertainty-Based Information—Elements of Generalized Information Theory*, Physica–Verlag, Heidelberg, 1999.
[14] Klir, G. J., "Generalized Information Theory: Aims, Results, and Open Problems," *Reliability Engineering and System Safety*, Vol. 85, Nos. 1–3, 2004, pp. 21–38.
[15] Zadeh, L., "Fuzzy Sets," *Information and Control*, Vol. 8, No. 3, 1965, pp. 338–353.
[16] Du, X., "Uncertainty Analysis with Probability and Evidence Theories," *Proceedings of the ASME 2006 International Design Technical Conferences & Computers and Information in Engineering Conference*, American Society of Mechanical Engineers, Fairfield, NJ, 2006.
[17] Du, X., Sudjianto, A., and Huang, B., "Reliability-Based Design with the Mixture of Random and Interval Variables," *Journal of Mechanical Design*, Vol. 127, No. 6, 2005, pp. 1068–1076.
[18] Bae, H. R., Grandhi, R. V., and Canfield, R. A., "Uncertainty Quantification of Structural Response Using Evidence Theory," *AIAA Journal*, Vol. 41, No. 10, 2003, pp. 2062–2068.
[19] Bae, H. R., Grandhi, R. V., and Canfield, R. A., "An Approximation Approach for Uncertainty Quantification Using Evidence Theory," *Reliability Engineering and System Safety*, Vol. 86, No. 3, 2004, pp. 215–225.
[20] Mourelatos, Z. P., and Zhou, J., "A Design Optimization Method Using Evidence Theory," *Journal of Mechanical Design*, Vol. 128, No. 4, July 2006, pp. 901–908.
[21] Agarwal, H., Renaud, J. E., Preston, E. L., and Padmanabhan, D., "Uncertainty Quantification Using Evidence Theory in Multi-disciplinary Design Optimization," *Reliability Engineering and System Safety*, Vol. 85, Nos. 1–3, 2004, pp. 281–294.
[22] Soundappan, P., Nikolaidis, E., Haftka, R. T., Grandhi, R., and Canfield, R., "Comparison of Evidence Theory and Bayesian Theory for Uncertainty Modeling," *Reliability Engineering and System Safety*, Vol. 85, Nos. 1–3, 2004, pp. 295–311.
[23] Bae, H.-R., Grandhi, R. V., and Canfield, R. A., "Epistemic Uncertainty Quantification Techniques Including Evidence Theory for Large-Scale Structures," *Computers and Structures*, Vol. 82, Nos. 13–14, 2004, pp. 1101–1112.
[24] Agarwal, H., Renaud, J. E., and Preston, E. L., "Trust Region Managed Reliability Based Design Optimization Using Evidence Theory," AIAA Paper 2003-1779, 2003.
[25] Bae, H.-R., Grandhi, R. V., and Canfield, R. A., "Reliability-Based Design Optimization Under Imprecise Uncertainty," AIAA Paper 2005-2069, 2005.
[26] Bae, H.-R., Grandhi, R. V., and Canfield, R. A., "Sensitivity Analysis of Structural Response Uncertainty Propagation Using Evidence Theory," *Structural and Multidisciplinary Optimization*, Vol. 31, No. 4, 2006, pp. 270–291.
[27] Helton, J. C., Johnson, J. D., Oberkampf, W. L., and Sallaberry, C. J., "Sensitivity Analysis in Conjunction with Evidence Theory Representations of Epistemic Uncertainty," *Reliability Engineering and System Safety*, Vol. 91, Nos. 10–11, 2006, pp. 1414–1434.
[28] Turanyi, T., "Sensitivity Analysis of Complex Kinetic Systems: Tools and Applications," *Journal of Mathematic Chemistry*, Vol. 5, No. 3, 1990, pp. 203–248.
[29] Cacuci, D. G., *Sensitivity and Uncertainty Analysis, Vol. 1: Theory*, CRC Press, Boca Raton, FL, 2003.
[30] Saltelli, A., Chan, K., and Scott, E. M., *Sensitivity Analysis*, Wiley, New York, 2000.
[31] Liu, H., Chen, W., and Sudjianto, A., "Relative Entropy Based Method for Probabilistic Sensitivity Analysis in Engineering Design," *Journal of Mechanical Design*, Vol. 128, No. 2, 2006, pp. 326–336.
[32] Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G., and Schaibly, J. H., "Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients. I. Theory," *Journal of Chemical Physics*, Vol. 59, No. 1, 1999, pp. 3873–3878.
[33] Saltlli, A., Tarantola, S., and Chan, K., "A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output," *Technometrics*, Vol. 41, No. 1, 1999, pp. 39–56.
[34] McKay, M., Morrison, J., and Upton, S., "Evaluating Prediction Uncertainty in Simulation Models," *Computer Physics Communications*, Vol. 117, Nos. 1–2, 1999, pp. 44–51.
[35] Homma, T., and Saltelli, A., "Importance Measures in Global Sensitivity Analysis of Nonlinear Models," *Reliability Engineering and System Safety*, Vol. 52, No. 1, 1996, pp. 1–17.
[36] Sobol, I., "Sensitivity Analysis for Nonlinear Mathematic Models," *Mathematical Modeling and Computational Experiment*, Vol. 1, No. 4, 1993, pp. 407–414.
[37] Sudjianto, A., Du, X., and Chen, W., "Probabilistic Sensitivity Analysis in Engineering Design Using Uniform Sampling and Saddlepoint Approximation," SAE Paper 2005-01-0344, 2005.
[38] Daniel, C., "One-at-a-Time-Plans," *Journal of American Statistical Association*, Vol. 68, No. 342, June 1973, pp. 353–360.
[39] Bain, L., and Engelhardt, M., *Introduction to Probability and Mathematical Statistics*, 2nd ed., Duxbury, Pacific Grove, CA, 1991.
[40] Shafer, G., "Perspectives on the Theory and Practice of Belief Functions," *International Journal of Approximate Reasoning*, Vol. 4, Nos. 5/6, 1990, pp. 323–362.
[41] Kuhn, H. W., and Tucker, A. W., "Nonlinear Programming," *Proceedings of the 2nd Berkeley Symposium*, University of California Press, Berkeley, 1951, pp. 481–492.
[42] Norton, R. L., *Machine Design—An Integrated Approach*, 2nd ed., Prentice–Hall, Upper Saddle River, NJ, 2000, pp. 472–480.

J. Samareh
Associate Editor