

# Unified uncertainty representation and quantification based on insufficient input data

Xiang Peng<sup>1</sup> · Jiquan Li<sup>1</sup> · Shaofei Jiang<sup>1</sup>

Received: 10 December 2016 / Revised: 10 May 2017 / Accepted: 14 May 2017 / Published online: 2 June 2017  
© Springer-Verlag Berlin Heidelberg 2017

**Abstract** Uncertainty quantification accuracy of system performance has an important influence on the results of reliability-based design optimization (RBDO). A new uncertain identification and quantification methodology is proposed considering the strong statistical variables, sparse variables, and interval variables simultaneously. Maximum likelihood function and Akaike information criterion (AIC) methods are used to identify the best-fitted distribution types and distribution parameters of sparse variables. The interval variables are represented with evidence theory. Finally, a unified uncertainty quantification framework considering the three types of uncertain design variables is put forward, and then the failure probability of system performance is quantified with belief and plausibility measures. The Kriging metamodel and random sampling method are used to reduce the computational complexity. Three examples are illustrated to verify the effectiveness of the proposed methodology.

**Keywords** Insufficient data · Uncertainty quantification · Distribution parameter analysis · Interval uncertainty · Akaike information criterion

**Electronic supplementary material** The online version of this article (doi:10.1007/s00158-017-1722-4) contains supplementary material, which is available to authorized users.

✉ Shaofei Jiang  
jsf75@zjut.edu.cn

Xiang Peng  
pengxiang@zjut.edu.cn

Jiquan Li  
lijq@zjut.edu.cn

<sup>1</sup> Key Laboratory of E&M, Zhejiang University of Technology, Hangzhou 310014, China

## 1 Introduction

Reliability analysis is concerned with the assessment of system performance in the presence of uncertainty. Reliability-based design optimization (RBDO) approach attempts to find the optimum structure such that the failure probability of system performance is satisfied (Zhu and Du 2016), and has been researched and applied widely by industry, government, and academia (Cho et al. 2016; Matsumura and Haftka 2013; Paulson and Starkey 2013; Suryawanshi and Ghosh 2015).

Uncertainties in the RBDO include input variable uncertainty and model uncertainty. The uncertain input variables can be divided into strong statistical variables, sparse variables, and interval variables according to the available amounts of input experimental data (Oberkampf et al. 2004). Strong statistical variables are aleatory uncertain variables, which are inherent and irreducible variation associated with the physical system (Mullins et al. 2016). Sparse variables and interval variables are epistemic uncertain variables, which derive from some levels of ignorance or incomplete information about the physical system or environment (Du 2006). Model uncertainty is also epistemic uncertainty, which is due to uncertainty in model parameters, numerical solution errors, and model form errors (Nannapaneni et al. 2016). Model uncertainty has an important influence on the reliability calculation, and has been researched widely (Arendt et al. 2012) (Jiang et al. 2016) (Hu et al. 2017). The paper focuses on the quantification of input variable uncertainty, so model uncertainty is not considered, and the design model is assumed to be deterministic.

Representation and management of input variable uncertainty are important in RBDO. Strong statistical variables are treated commonly as random variables with known probability distributions, such as Normal distribution, Gamma distribution, Beta distribution. Interval variables are represented using non-probabilistic

approaches, such as evidence theory, fuzzy theory, convex model theory, interval method (Bae et al. 2004; Beer et al. 2013; Simoen et al. 2015). The uncertainties of sparse variables are complicated because their distribution types and distribution parameters are uncertain. Though maximum likelihood estimation method (Sankararaman and Mahadevan 2011), Bayesian method (Xi et al. 2014), stochastic inverse method (Choi and Yoo 2016), nonparametric minimum power method (Chee 2017) have been presented, the accuracy of results is influenced by the distribution types. If the selected distribution types are not proper, the accuracy is very low. Then, Kang et al. (2016) proposed a sequential statistical modeling method to select appropriate candidate distribution types, but the method can accurately identify distribution types only if sufficient data are available. To solve this problem, a selection criterion of distribution types is proposed for the uncertainty representation of sparse variables in this paper.

The three types of input variable uncertainties are propagated to the uncertainty of system performance, so large amounts of methods have been proposed to integrate these uncertainties and quantify the influence on the system performance. The uncertainties of performance function considering the influences of both strong statistical variables and interval variables are quantified using belief and plausibility measures (Du 2008), probabilistic approach (Zaman et al. 2011), sampling-based worst case method (Yoo and Lee 2013), unified uncertain analysis method (Li et al. 2016), and probability box model (Liu et al. 2017). The performance uncertainty under sparse variables are calculated using Bayesian model averaging method and Bayesian hypothesis testing method (Sankararaman and Mahadevan 2013), possibility theory (Ren et al. 2015), and single-loop sampling approach (Nannapaneni et al. 2016). Although many methods have been proposed for uncertainty quantification, there is not a unified framework of uncertainty quantification considering these three types of input uncertain variables, simultaneously.

Bayesian theory is an attractive framework in uncertainty representation and management, and has been widely applied, such as identification of material parameters in high strength steel (Wang et al. 2015), construction of surrogate dictionary in elasticity problems (Contreras et al. 2016), representation of model uncertainty (Gal and Ghahramani 2016), construction of analysis model in credit scoring (Xia et al. 2017). Therefore, through using Bayesian theory, we proposed a unified uncertainty identification and quantification methodology considering these three types of uncertain variables due to insufficient input data. The construction of the work is

as follows: In Section 2, the three types of uncertainties are analyzed and represented using different presentation functions. The distribution types, distribution parameters and weight ratios of sparse variables are identified and represented in Section 3. In Section 4, the interval variables are represented using evidence theory. The unified uncertainty quantification framework is proposed, and the calculation flowrate is shown in Section 5 followed by three examples in Section 6.

## 2 Analysis of different types of uncertain input variables

There are three types of uncertain input variables, including strong statistical variables  $\mathbf{X}$ , sparse variables  $\mathbf{Y}$ , and interval variables  $\mathbf{Z}$  (Oberkampf et al. 2004). The three types of uncertain input variables are represented using different presentation forms.

### (1) Strong statistical variables $\mathbf{X}$

The random distribution types can be acquired easily or a large amount of experimental data is available, the uncertain variables  $\mathbf{X}$  are sufficient to derive an accurate statistical model. So,  $\mathbf{X}$  are specified by a probability distribution type  $\theta$  with precise distribution parameters  $\xi$ .

$$\mathbf{X} \sim \theta(\xi). \quad (1)$$

The distribution type  $\theta$  may be Normal distribution, Gamma distribution, Exponential distribution, etc. If  $\theta$  is Normal distribution type,  $\xi$  are mean  $\xi_1$  and standard deviation  $\xi_2$ .

### (2) Sparse variables $\mathbf{Y}$

Only a small number of experimental input point or interval data can be acquired, it is time-consuming and computational complex to acquire more input data, so the uncertain variables  $\mathbf{Y}$  are insufficient. A single distribution type may not fit the available input data, so the sum of multiple distribution types  $\theta_i$  is used to fit  $\mathbf{Y}$ .

$$\mathbf{Y} \sim \sum w_i \theta_i(\xi^i), \quad (2)$$

where the subscript  $i$  represents the  $i$ -th distribution type. The distribution type  $\theta_i$ , weight ratio  $w_i$  and distribution parameters  $\xi^i$  will be identified and represented in Section 3.

### (3) Interval variables $\mathbf{Z}$

**Table 1** The candidate distribution type

Distribution type	Probability density function	Parameters
$\theta_1$ : Normal	$f(y) = \frac{1}{\sqrt{2\pi\xi_2^2}} \exp\left(-\frac{(y-\xi_1)^2}{2\xi_2^2}\right)$	$\xi_1$ : location; $\xi_2$ : scale
$\theta_2$ : F	$f(y) = \frac{\Gamma(\frac{\xi_1+\xi_2}{2})}{\Gamma(\frac{\xi_1}{2})\Gamma(\frac{\xi_2}{2})} \left(\frac{\xi_1}{\xi_2}\right)^{\frac{\xi_1}{2}} \frac{y^{\xi_1-2}}{\left[1+\frac{\xi_1}{\xi_2}y\right]^{\frac{\xi_1+\xi_2}{2}}}$	$\xi_1$ : numerator degree $\xi_2$ : denominator degree
$\theta_3$ : Gamma	$f(y) = \frac{\xi_2^{\xi_1}}{\Gamma(\xi_1)} y^{\xi_1-1} e^{-\xi_2 y}$	$\xi_1$ : scale $\xi_2$ : rate
$\theta_4$ : Weibull	$f(y) = \frac{\xi_2}{\xi_1} \left(\frac{y}{\xi_1}\right)^{\xi_2-1} \exp\left(-\left(\frac{y}{\xi_1}\right)^{\xi_2}\right)$	$\xi_1$ : scale $\xi_2$ : shape
$\theta_5$ : Extreme value	$f(y) = \frac{1}{\xi_2} \exp\left(\frac{y-\xi_1}{\xi_2}\right) \exp\left(-\exp\left(\frac{y-\xi_1}{\xi_2}\right)\right)$	$\xi_1$ : location $\xi_2$ : scale
$\theta_6$ : Uniform	$f(y) = \begin{cases} \frac{1}{\xi_2-\xi_1} & \xi_1 \leq y \leq \xi_2 \\ 0 & \text{otherwise} \end{cases}$	$\xi_1$ : lower endpoint $\xi_2$ : upper endpoint
$\theta_7$ : Exponential	$f(y) = \xi_1 \exp(-\xi_1 y)$	$\xi_1$ : scale
$\theta_8$ : Log normal	$f(y) = \frac{1}{y\sqrt{2\pi\xi_2^2}} \exp\left(-\frac{(\ln y - \xi_1)^2}{2\xi_2^2}\right), y > 0$	$\xi_1$ : log mean; $\xi_2$ : log standard deviation

Interval variables  $\mathbf{Z}$  are also epistemic uncertain variables. Experimental input data are missing, so the interval variable  $z_i \in \mathbf{Z}$  is specified using evidence theory. The details will be explained in Section 4.1.

### 3 Parameter identification and representation of sparse variables

#### 3.1 Distribution parameter estimation under candidate distribution type

The probability density function of sparse variables  $\mathbf{Y}$  cannot be acquired directly, it is difficult to fit  $\mathbf{Y}$  using arbitrary fixed distribution type. So, the eight candidate distribution types are used, which are shown in Table 1.

Assuming the insufficient data of  $\mathbf{Y}$  contains  $m$  point data  $a_i (i = 1, \dots, m)$  and  $n$  interval data  $[b_j, \bar{b}_j] (j = 1, \dots, n)$ . Let  $f_y(y|\xi, \theta_k)$  denote the probability density function (PDF) under candidate distribution type  $\theta_k (k = 1, \dots, 8)$  and distribution parameters  $\xi$ . Firstly, for every candidate distribution type  $\theta_k$ , the likelihood estimation function

(Sankararaman and Mahadevan 2011)  $L(\xi, \theta_k)$  using the prescribed point data and interval data is constructed.

$$L(\xi, \theta_k) \sim \left[ \prod_{i=1}^m f_y(y = a_i | \xi, \theta_k) \right] \left[ \prod_{j=1}^n \int_{b_j}^{\bar{b}_j} f_y(y | \xi, \theta_k) dy \right]. \quad (3)$$

The maximum likelihood estimations of  $\xi$  under distribution type  $\theta_k$  are acquired by maximizing  $L(\xi, \theta_k)$ . Further, the uncertainty of distribution parameters  $\xi$  is calculated using Bayes' theorem. The probability density function  $f_\xi(\xi|\theta_k)$  of the distribution parameters  $\xi$  is expressed as

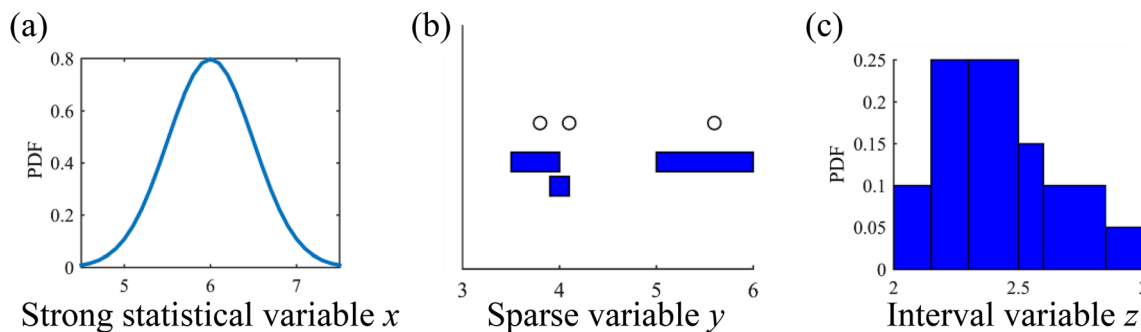
$$f_\xi(\xi|\theta_k) = \frac{L(\xi, \theta_k)}{\int L(\xi, \theta_k) d\xi}. \quad (4)$$

#### 3.2 Distribution type identification based on Akaike information criterion

The estimated distribution parameters under different distribution type are different based on the same insufficient input data. Therefore, Akaike information criterion (AIC) method is employed to estimate the best-fitted distribution types and corresponding weight ratios.

**Table 2** Intervals and BPAs of  $z$

Interval	[2, 2.15]	[2.15, 2.3]	[2.3, 2.5]	[2.5, 2.6]	[2.6, 2.85]	[2.85, 3]
BPA	0.1	0.25	0.25	0.15	0.2	0.05



**Fig. 1** Input experimental data of design variables

AIC is the relative quality measure of statistical models for a given set of data, which can compute the fitting degree of the candidate distributions. (Gelman et al. 2013) The AIC value of the candidate distribution type  $\theta_k$  is defined in (5)

$$AIC_k = 2num_k - 2\ln(L_{\max}(\xi, \theta_k)), \tag{5}$$

where,  $num_k$  is the amount of estimated distribution parameters in the candidate distribution type  $\theta_k$ ,  $L_{\max}(\xi, \theta_k)$  is the maximum value of the likelihood function  $L(\xi, \theta_k)$  under distribution type  $\theta_k$ , the subscript  $k(k=1, \dots, 8)$  presents the eight candidate distribution type defined in Section 3.1.

The AIC values of the eight candidate distribution types are calculated and denoted by  $AIC_1, AIC_2, \dots, AIC_8$ , respectively. Let  $AIC_{\min}$  be the minimum of these values. Then, the probability  $P_{\theta_{-k}}$  that the  $k$ -th model minimizes the estimated information loss can be interpreted in (6).

$$P_{\theta_{-k}} = \exp((AIC_{\min} - AIC_k)/2). \tag{6}$$

The distribution types with  $P_{\theta_{-k}} \geq 0.1$  (Taguri et al. 2014) are selected to represent the sparse variables  $\mathbf{Y}$ . The weight ratios  $w_k$  of these selected distribution types are proportional to the probability  $P_{\theta_{-k}}$ , and the summation of these weight ratios is 1.

$$w_k = \frac{P_{\theta_{-k}}}{\sum P_{\theta_{-k}}}. \tag{7}$$

**Table 3** Optimum distribution parameters under candidate distribution type

Distribution type	Parameter $\xi_1$	Parameter $\xi_2$	AIC Value
Normal distribution	4.4	0.8	22.62
F distribution	50	1.7	46.84
Gamma distribution	44.3	0.1	22.19
Weibull distribution	4.8	6.1	23.19
Extreme value distribution	4.8	0.8	24.01
Uniform distribution	3.7	5.6	18.35
Exponential distribution	4.5	—	36.54
Log normal distribution	1.6	0.3	25.81

As an example, suppose that the AIC values are 100, 102, 104, 106, 108, 110, 112 and 114, respectively. Then the probabilities of these distribution types  $P_{\theta}$  are 1, 0.368, 0.135, 0.05, 0.02, 0.01, 0.0025 and 0.0009, respectively. Therefore, the distribution type 1 (Normal distribution), distribution type 2 (F distribution) and distribution type 3 (Gamma distribution) are used, and the weight ratios are 0.665, 0.245, and 0.090, respectively.

### 3.3 Distribution parameters estimation using mixed distribution types

After determining the distribution types and corresponding weight ratios, the distribution parameters are determined using the Bayesian model averaging method (Nannapaneni et al. 2016).

The probability density function under multiple distribution types is calculated using (8).

$$f_y(y|\xi) = \sum w_k f_y(y|\xi, \theta_k), \tag{8}$$

where  $\theta_k$  is the selected distribution type,  $w_k$  is the corresponding weight ratio, and  $f_y(y|\xi, \theta_k)$  is the probability density function under distribution type  $\theta_k$ .

The combined likelihood function for insufficient input point data and interval data is constructed as follows.

$$L(\xi) = \left[ \prod_{i=1}^m f_y(y = a_i|\xi) \right] \left[ \prod_{j=1}^n \int_{b_j}^{\bar{b}_j} f_y(y|\xi) dy \right]. \tag{9}$$

The optimum distribution parameters  $\xi^*$  of the selected distribution types are calculated through maximizing  $L(\xi)$ . Further, the uncertainty of the distribution parameters  $\xi$  is calculated using Bayes' theorem. The probability density function  $f_{\xi}(\xi)$  is expressed in (10).

$$f_{\xi}(\xi) = \frac{L(\xi)}{\int L(\xi) d\xi}. \tag{10}$$

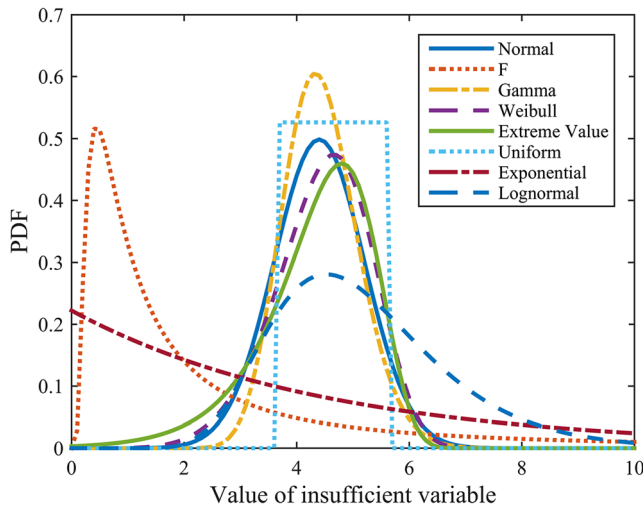


Fig. 2 Probability density function under candidate distribution types

### 3.4 Discrete representation of uncertain distribution parameters

To reduce the computational difficulty in the representation of sparse variables, the continuous uncertain distribution parameters  $\xi$  are dispersed in the neighborhood of optimum distribution parameters  $\xi^*$ . For example, there are  $mm$  parameters  $\xi_1, \xi_2, \dots, \xi_{mm}$  in  $\xi$ , and the optimum parameters are  $\xi_1^*, \xi_2^*, \dots, \xi_{mm}^*$ , respectively. In the dispersion of  $\xi_i, \xi_1, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_{mm}$  are fixed to be  $\xi_1^*, \dots, \xi_{i-1}^*, \xi_{i+1}^*, \dots, \xi_{mm}^*$ , respectively. The step to step flowrate to disperse  $\xi_i$  is shown as follows.

Step 1: The lower and upper bound of distribution parameter  $\xi_i$  are determined through analyzing the probability

density function  $f_{\xi}(\xi)$  in (10). In theory, the uncertain design space of  $\xi_i$  is  $(-\infty, +\infty)$ , it is impossible to obtain the bounds. So, the 99% confidence interval is used, the lower bound is set to be  $\xi_i^{\min} = \Gamma_{\xi_i}^{-1}(0.005)$ , and the upper bound is set to be  $\xi_i^{\max} = \Gamma_{\xi_i}^{-1}(0.995)$ , where  $\Gamma_{\xi_i}^{-1}(p)$  is the inverse cumulative distribution function of  $\xi_i$ .

- Step 2: The interval  $[\xi_i^{\min}, \xi_i^{\max}]$  is decomposed according to the inverse cumulative distribution function.  $\xi_i^j = \Gamma_{\xi_i}^{-1}(0.1j), j = 1, \dots, 9$  are calculated, and then the design space of  $\xi_i$  is decomposed to ten Bayesian evidence intervals,  $[\xi_i^{\min}, \xi_i^1], [\xi_i^j, \xi_i^{j+1}] (j = 1, \dots, 8)$  and  $[\xi_i^9, \xi_i^{\max}]$ , which have the same BPA  $m = 0.1$ .
- Step 3: The long interval is decomposed. If the interval length  $\xi_i^{j+1} - \xi_i^j > 0.2(\xi_i^{\max} - \xi_i^{\min})$ , the interval is decomposed to  $[\xi_i^j, (\xi_i^j + \xi_i^{j+1})/2]$  and  $[(\xi_i^j + \xi_i^{j+1})/2, \xi_i^{j+1}]$ , and the corresponding BPA are calculated using the cumulative distribution function of  $\xi_i$ .
- Step 4: If the length of the decomposed sub-interval exceeds  $0.2(\xi_i^{\max} - \xi_i^{\min})$ , Step 3 is implemented again.
- Step 5: After executing Step 4, the uncertain distribution parameters  $\xi$  are decomposed into  $nm$  sub-interval  $[\xi_i^{j-1}, \xi_i^j] (j = 1, \dots, nm)$ , where  $\xi_i^0 = \xi_i^{\min}$  and  $\xi_i^{nm} = \xi_i^{\max}$ . Then  $\xi_i$  is represented using discrete points  $(\xi_i^{j-1} + \xi_i^j)/2 (j = 1, \dots, nm)$  with the BPA of interval  $[\xi_i^{j-1}, \xi_i^j]$ .

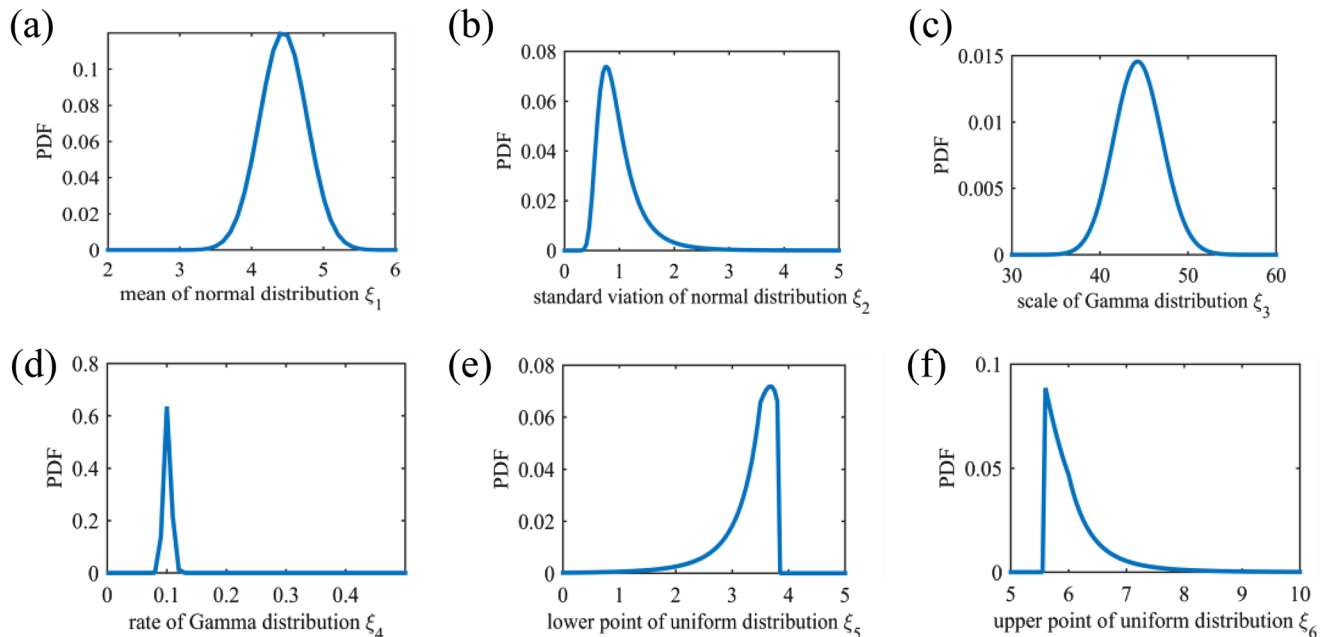


Fig. 3 Probability density function of distribution parameters

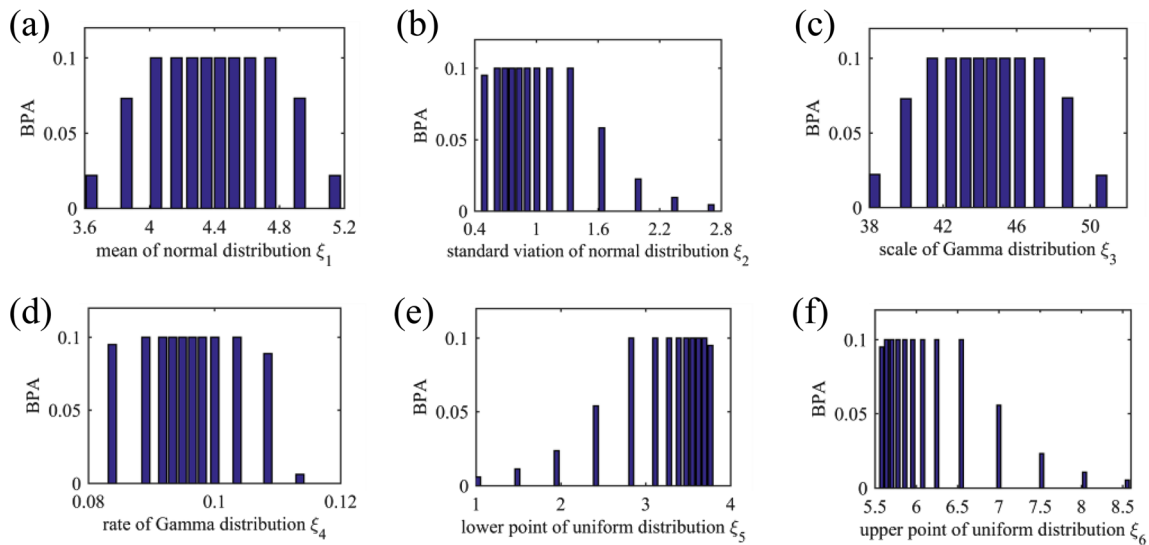


Fig. 4 Sampling points and BPAs of uncertain distribution parameters

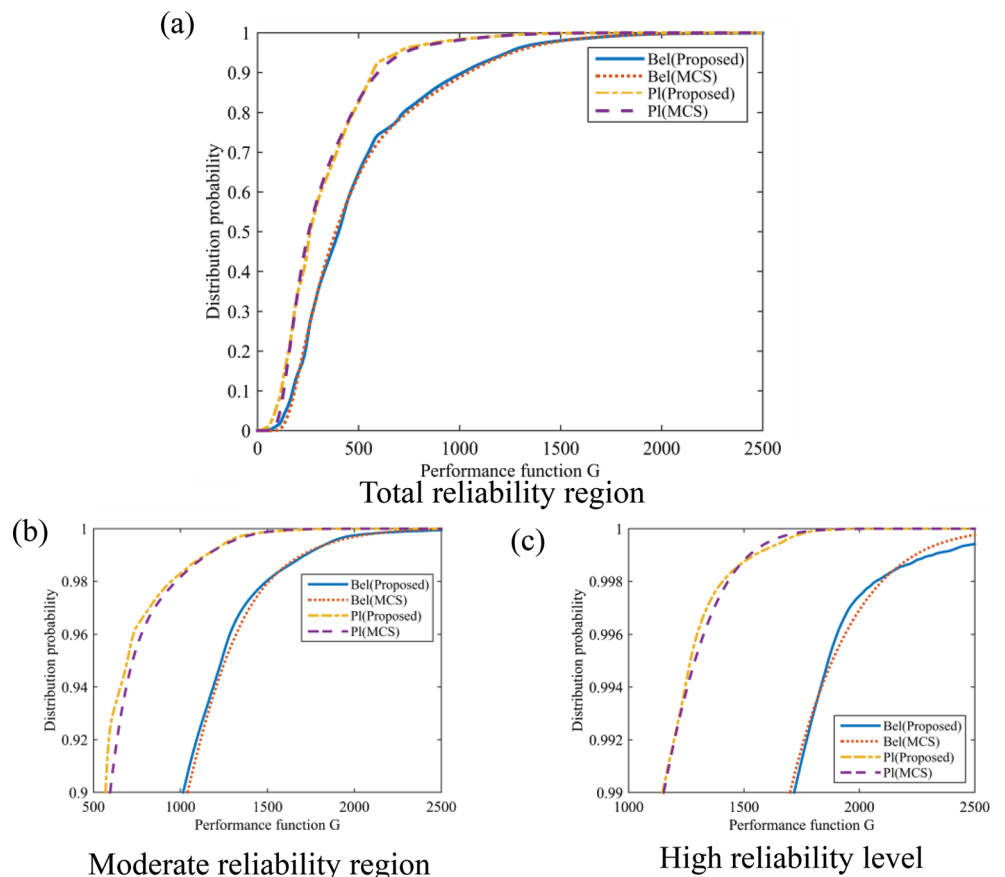
### 4 Representation of interval variables based on evidence theory

#### 4.1 Evidence theory

Evidence theory, also referred to as Dempster-Shafer (DS) theory, is presented by Shafer (Shafer 1976). It

combines evidence from different incomplete knowledge situations and arrives at a degree of belief that considers all available evidences. In evidence theory, the basic probability assignment (BPA) represents the uncertain distribution. There are three formations of BPAs, including general, Bayesian and consonant BPA, respectively (Tao et al. 2016).

Fig. 5 Belief and plausibility function of performance function G





**Table 4** Uncertainty quantification results of mathematical example

	Proposed method	Monto Carlo method
Belief value	0.8960	0.8881
Plausibility value	0.9830	0.9816
Evaluation number	$10^3$	$6 \times 10^7$

There are no experimental input data, and the BPAs are acquired from the expert knowledge, so the Bayesian BPA structure is used to represent interval variables  $\mathbf{Z}$ . Let  $\Psi$  denote the set of all possible values for the interval variable  $Z_i$ , and  $\gamma$  is the subset of  $\Psi$ . The BPA  $s_{Z_i}(\gamma)$  satisfies the following axioms (Du 2008).

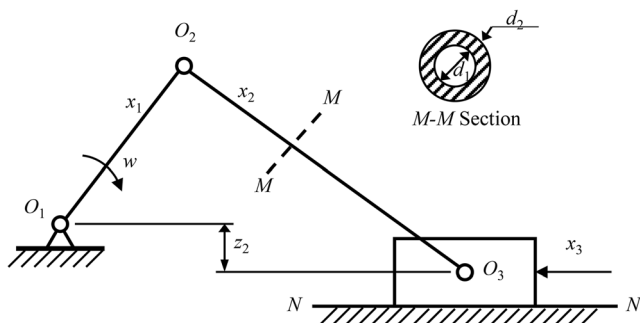
$$\begin{cases} s_{Z_i}(\gamma) \geq 0 & \text{if } \gamma \in \Psi \\ s_{Z_i}(\gamma) = 0 & \text{if } \gamma \notin \Psi \\ \sum_{\gamma \in \Psi} s_{Z_i}(\gamma) = 1 \end{cases} \quad (11)$$

While there are multiple interval variables, the multiple BPA structures can be aggregated using the rules of combination, which is similar to a joint probability in probability theory. For example, there are interval variables  $Z_1$  and  $Z_2$ , the joint BPA  $s_{\mathbf{Z}}(\gamma_{\mathbf{Z}})$  can be calculated using (12).

$$s_{\mathbf{Z}}(\gamma_{\mathbf{Z}}) = \begin{cases} s_{Z_1}(\gamma_1)s_{Z_2}(\gamma_2) & \text{when } \gamma_{\mathbf{Z}} = \gamma_1 \times \gamma_2 \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

**4.2 Uncertain measure of interval variables**

In the reliability analysis problem only considering interval variables  $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_q\}$ , the system performance function is expressed by  $G = g(\mathbf{Z})$ , and the failure model  $F$  is defined as  $F = \{\mathbf{Z} | g(\mathbf{Z}) \leq c\}$ , where  $c$  is a limit state.  $m_{\mathbf{Z}}(\gamma)$  is the joint BPA over the interval space  $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_q\}$ , the failure likelihood of the system performance function is quantified by belief measure  $Bel(F)$  and plausibility measure  $Pl(F)$ .



**Fig. 6** A crank-slider mechanism

**Table 5** Strong statistical variables X

Variables	Mean	Standard deviation	Distribution types
$x_1$	100 mm	0.01 mm	Normal
$x_2$	400 mm	0.01 mm	Normal
$x_3$	250 kN	25 kN	Normal

The belief measure  $Bel(F)$  is the lower bound of failure probability, which reflects the degree of belief that the failure event  $F$  would occur. It is calculated by adding the BPAs of subset  $\gamma$  entirely within the failure region  $F = \{\mathbf{Z} | g(\mathbf{Z}) \leq c\}$ , which is shown in (13)

$$Bel(F) = \sum_{\gamma \in F} s_{\mathbf{Z}}(\gamma) \quad (13)$$

The plausibility measure  $Pl(F)$  is the upper bound of failure probability, which is calculated by adding the BPAs of the subsets  $\gamma$  that are in the failure region and the BPAs of the subsets  $\gamma$  which intersect with the failure region. So  $Pl(F)$  is defined as follows.

$$Pl(F) = \sum_{\gamma \cap F \neq \emptyset} s_{\mathbf{Z}}(\gamma) \quad (14)$$

Then the true failure probability  $p_F$  is bounded in the interval between  $Bel(F)$  and  $Pl(F)$ , namely.

$$Bel(F) \leq p_F \leq Pl(F) \quad (15)$$

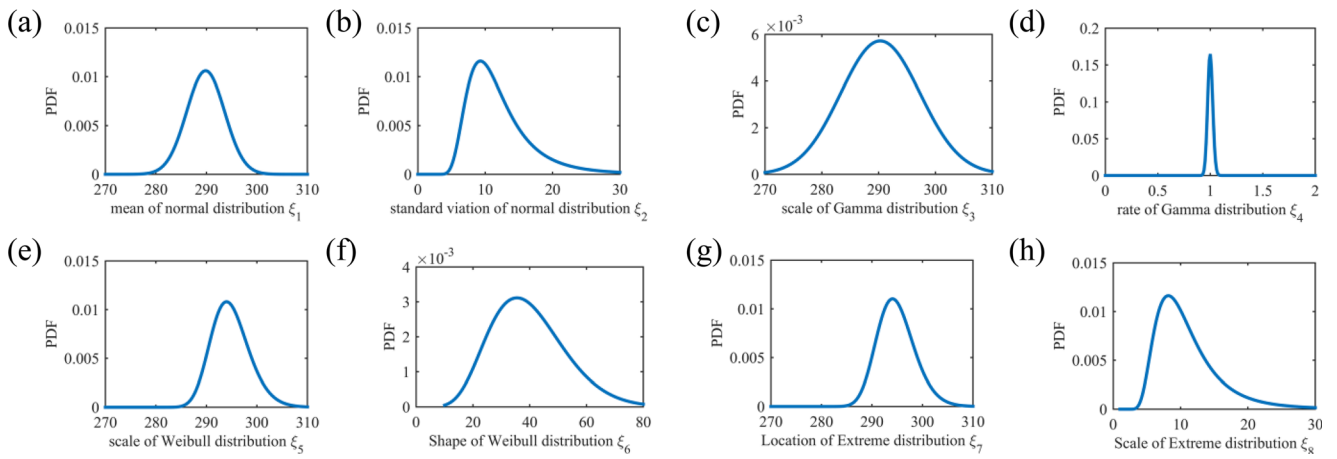
**5 Uncertainty quantification considering multiple types of uncertain input variables**

**5.1 Reliability calculation considering mixture uncertainties**

The system performance function considering three types of uncertainties can be expressed as  $G = g(\mathbf{X}, \mathbf{Y}(\xi), \mathbf{Z})$ .

**Table 6** Interval variables Z

Interval variables	Intervals	BPA
$z_1$	[100 20]	0.2
	[120 40]	0.4
	[140 50]	0.4
$z_2$	[0.15   0.18]	0.3
	[0.18   0.23]	0.3
	[0.23   0.25]	0.4



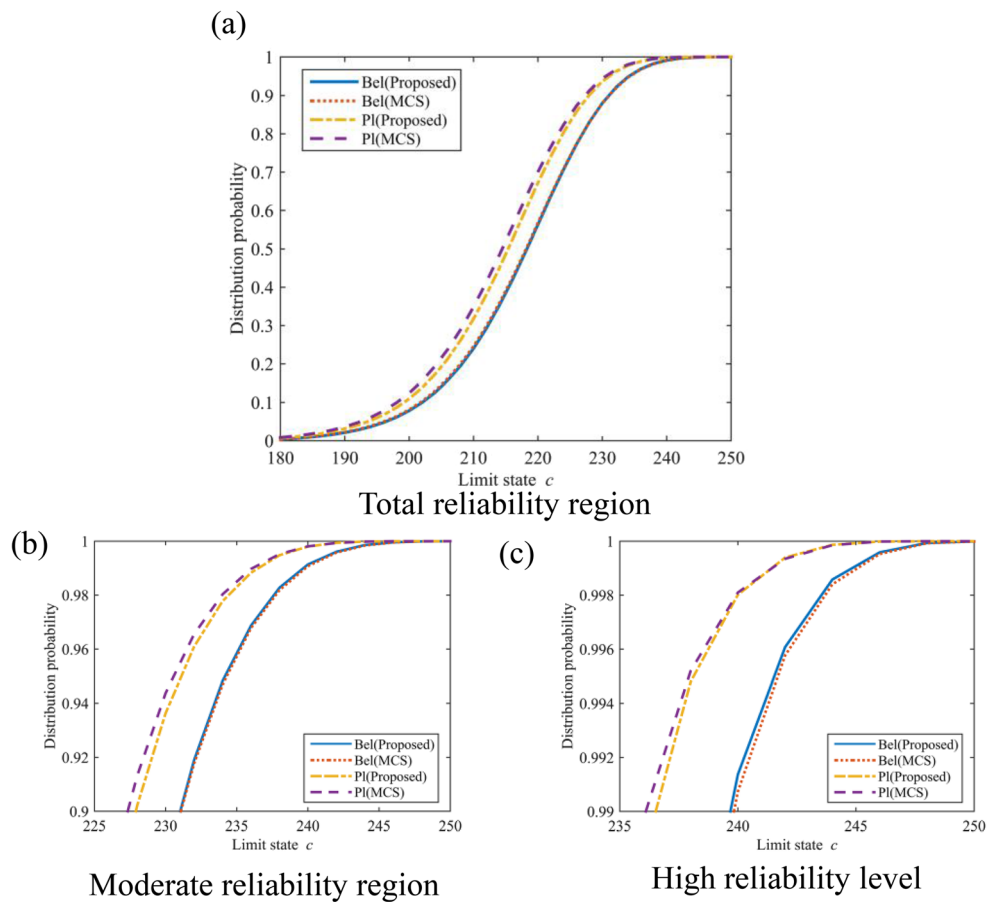
**Fig. 7** Probability density function of distribution paramters

The strong statistical variables  $\mathbf{X}$  are expressed using probability distribution function with determinate distribution parameters. The sparse variables  $\mathbf{Y}$  are determined with multiple distribution types, and the distribution parameters  $\xi$  are represented using discrete sampling points

with their BPAs. The interval variables  $\mathbf{Z}$  are quantified using intervals with their BPAs.

The uncertain distribution parameters  $\xi$  are represented using discrete sampling points  $(\xi^{i-1} + \xi^i)/2 (i = 1, \dots, nm)$  with the BPA of interval  $[\xi^{i-1}, \xi^i]$ , and the uncertain interval

**Fig. 8** Belief and plausibility measure of performance function G





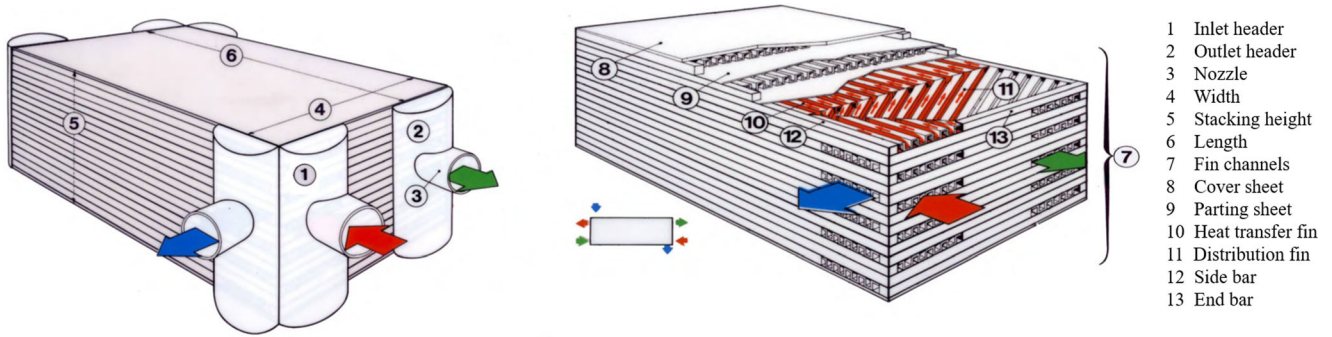


Fig. 9 The construction of heat exchanger

variables  $\mathbf{Z}$  are represented using intervals  $[\eta_j, \eta_{j+1}] (j = 1, \dots, nm)$  with their BPAs.

According to the theorem of total probability, the total failure probability  $p_F$  can be computed by

$$\begin{aligned}
 p_F &= \Pr\{g(\mathbf{X}, \mathbf{Y}(\xi), \mathbf{Z}) \leq c\} \\
 &= \sum_{i=1}^{nm} \sum_{j=1}^{nm} \Pr\{g(\mathbf{X}, \mathbf{Y}(\xi), \mathbf{Z}) \leq c | \xi \in [\xi^{i-1}, \xi^i], \mathbf{Z} \in [\eta_j, \eta_{j+1}]\} \Pr\{\xi \in [\xi^{i-1}, \xi^i]\} \Pr\{\mathbf{Z} \in [\eta_j, \eta_{j+1}]\} \\
 &\approx \sum_{i=1}^{nm} \sum_{j=1}^{nm} \Pr\left\{g\left(\mathbf{X}, \mathbf{Y}\left(\frac{\xi^{i-1} + \xi^i}{2}\right), \mathbf{Z}\right) \leq c | \mathbf{Z} \in [\eta_j, \eta_{j+1}]\right\} \Pr\{\xi \in [\xi^{i-1}, \xi^i]\} \Pr\{\mathbf{Z} \in [\eta_j, \eta_{j+1}]\} \\
 &= \sum_{i=1}^{nm} \sum_{j=1}^{nm} p_{F-ij} \Pr\{\xi \in [\xi^{i-1}, \xi^i]\} \Pr\{\mathbf{Z} \in [\eta_j, \eta_{j+1}]\}.
 \end{aligned} \tag{16}$$

where

$$p_{F-ij} = \Pr\left\{g\left(\mathbf{X}, \mathbf{Y}\left(\frac{\xi^{i-1} + \xi^i}{2}\right), \mathbf{Z}\right) \leq c | \mathbf{Z} \in [\eta_j, \eta_{j+1}]\right\}. \tag{17}$$

The minimum and maximum of  $p_{F-ij}$  are defined to be  $p_{F-ij}^L$  and  $p_{F-ij}^U$ , respectively. The Kriging model and random sampling method are used to calculate  $p_{F-ij}^L$  and  $p_{F-ij}^U$ . Optimal Latin Hypercube technique is used to obtain the sample points, and then the Kriging metamodel  $G$  of system performance function  $G$  is constructed using MATLAB Kriging Toolbox

DACE (Søren et al. 2017). In the sub-interval  $\xi \in [\xi^{i-1}, \xi^i]$  and  $\mathbf{Z} \in [\eta_j, \eta_{j+1}]$ , The  $\xi m$  random samplings of  $\xi \in [\xi^{i-1}, \xi^i]$  and  $zn$  random samplings of  $\mathbf{Z} \in [\eta_j, \eta_{j+1}]$  are obtained, and then the system performance under these sample points are calculated using the Kriging metamodel  $G$ . The minimum and maximum of  $G$  under every samples of  $\mathbf{Z}$  can be calculated and named as  $Min_{ij-q}$  and  $Max_{ij-q}$ , respectively, where  $q \in [1, \dots, zn]$ . Then  $p_{F-ij}^L$  and  $p_{F-ij}^U$  are the failure probability of  $Min_{ij-q}$  and  $Max_{ij-q}$ , and calculated in (18) and (19).

$$p_{F-ij}^L = \frac{NUM(Min_{ij-q} < c)}{zn}, \tag{18}$$

$$p_{F-ij}^U = \frac{NUM(Max_{ij-q} < c)}{zn}. \tag{19}$$

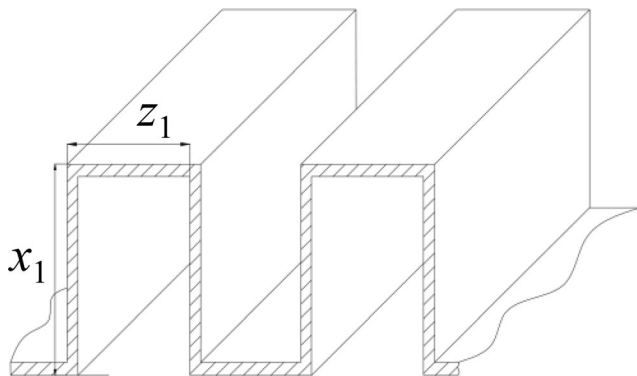


Fig. 10 The separating fin structure of hot stream

Table 7 Strong statistical variables for example 3

Variable	Mean	Standard deviation	Distribution
$x_1$ : Fin height of hot stream (mm)	9.5	0.06	Normal
$x_2$ : Fin height of cold stream (mm)	6.5	0.05	Normal

**Table 8** Interval variable for example 3

Variable	Intervals	BPA
$z_1$ : fin space ( $mm$ )	[1.5, 1.9]	0.2
	[1.9, 2.1]	0.6
	[2.1, 2.5]	0.2

After the calculation of  $p_{Fij}^L$  and  $p_{Fij}^U$ , the belief and plausibility of  $p_F$  are calculated in (20) and (21), respectively.

$$Bel(G) = P_F^L = \sum_{i=1}^{nm} \sum_{j=1}^{nm} p_{F_{-ij}}^L \Pr\{\xi \in [\xi^{i-1}, \xi^i]\} \Pr\{Z \in [\eta_j, \eta_{j+1}]\}, \quad (20)$$

$$Pl(G) = P_F^U = \sum_{i=1}^{nm} \sum_{j=1}^{nm} p_{F_{-ij}}^U \Pr\{\xi \in [\xi^{i-1}, \xi^i]\} \Pr\{Z \in [\eta_j, \eta_{j+1}]\}. \quad (21)$$

### 5.2 Calculation procedure of the proposed algorithm

The uncertainty identification and quantification of system failure probability are implemented using mixture of three types of uncertainties based on insufficient input data. The procedure of the proposed algorithm is summarized as follows.

- (1) The uncertain variables are decomposed to strong statistical variables  $\mathbf{X}$ , sparse variables  $\mathbf{Y}$ , and intervals variables  $\mathbf{Z}$  according to the available amount of input experimental data.
- (2) The strong statistical variables  $\mathbf{X}$  are represented using probability distribution function with determinate distribution parameters.
- (3) The interval variables  $\mathbf{Z}$  are represented by sub-intervals with basic probability assignment.
- (4) For sparse variables  $\mathbf{Y}$ , the distribution parameters under multiple distribution types are calculated through maximum of  $L(\xi, \theta_k)$  based on sparse input point and interval data. Then, the best-fitting distribution types  $\theta$  and corresponding weight ratios  $\mathbf{w}$  are calculated using AIC method, the distribution parameters  $\xi$  and their probability density function

**Table 9** Sparse variables for example 3

Variable	Sparse points	Sparse intervals
$y_1$ : Input temperature of hot stream (K)	(356, 360, 362)	[357361], [362363]
$y_2$ : Input velocity of hot stream (m/s)	(7.8, 8, 8.1, 8.3)	[7.5 7.9], [8.1, 8.4]
$y_3$ : Input temperature of cold stream (K)	(310, 315, 317)	[312316], [318320]
$y_4$ : Input velocity of cold stream (m/s)	(0.09, 0.1)	[0.095 0.11]

$f_{\xi}(\xi)$  are determined using the Bayesian model averaging method.

- (5) The continuous uncertain distribution parameters  $\xi$  are decomposed to reduce computation complexity.  $\xi$  are represented using discrete points  $(\xi^{i-1} + \xi^i)/2 (i = 1, \dots, nm)$  with the BPA of interval  $[\xi^{i-1}, \xi^i]$ .
- (6) The kriging model  $G$  of the system performance function  $G = g(\mathbf{X}, \mathbf{Y}(\xi), \mathbf{Z})$  is constructed using Optimal Latin Hypercube technique. In the  $ij$ -th sub-interval where  $\xi \in [\xi^{i-1}, \xi^i]$  and  $\mathbf{Z} \in [\eta_j, \eta_{j+1}]$ , the minimum value  $Min_{ij-q}$  and maximum value  $Max_{ij-q}$  of  $G$  are calculated using the random sampling method, and then  $p_{fij}^L$  and  $p_{fij}^U$  are calculated.
- (7) The belief  $Bel(G)$  and plausibility  $Pl(G)$  of  $p_f$  are calculated using (20) and (21).

## 6 Numerical examples

### 6.1 Example 1-mathematical problem

A mathematical example is considered to demonstrate the effectiveness of the proposed method. The system performance function is  $G = (x + y)^z$ , where  $x$  is strong statistical variable,  $y$  is sparse variable, and  $z$  is interval variable.

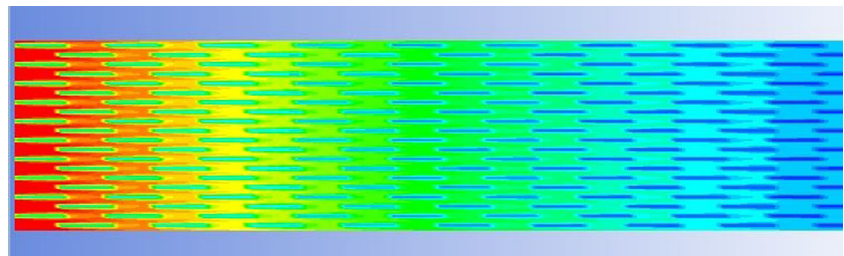
The strong statistical variable  $x$  is assumed to be Normal distribution with mean 6 and standard deviation 0.5, then  $x \sim N(6, 0.5^2)$ . The available data of  $y$  contain sparse points {3.8, 4.1, 5.6} and intervals [3.5, 4], [3.9, 4.1], [5, 6]. The intervals and their BPAs of  $z$  are shown in Table 2. The input data of these three variables are shown in Fig. 1.

The eight candidate distribution types  $\theta$  in Table 1 are used to identify distribution parameters  $\xi$  of  $y$ . The optimum distribution parameters under these distribution types are calculated through maximizing  $L(\xi, \theta_k)$  in (3), and the results are listed in Table 3. The probability density function of  $y$  with optimum distribution parameters under these candidate distribution types are shown in Fig. 2.

The best-fitted distribution types are selected using AIC method. The AIC values of these eight distribution types are also listed in Table 3, so the distribution type with minimum AIC value is uniform distribution, and the relative probability  $p_{\theta_k}$  of these eight distribution types are 0.1182,  $6.5155 \times 10^{-7}$ , 0.1464, 0.0890, 0.0591, 1,  $1.1240 \times 10^{-7}$ , and 0.024, respectively. According to the algorithm in Section 3.2, the best-fitted distribution types are Uniform distribution, Normal distribution, and Gamma distribution. Their weigh ratios  $\mathbf{w}$  are 0.7908, 0.0935, and 0.1157, respectively. So, the uncertainty of distribution parameters are calculated using (10), and the results are shown in Fig. 3.

To reduce the computational complexity, the uncertain distribution parameters are dispersed to discrete points

**Fig. 11** Temperature distribution in the heat exchanger



with their BPAs, the dispersed results of the six distribution parameters are shown in Fig. 4. The Kriging model of performance function  $G$  is constructed based on  $10^3$  samples generated using Optimal Latin Hypercube method, and the root model square error (RMSE) of the constructed Kriging model is less than 0.0258. The uncertainty of system performance function  $G$  is estimated using the proposed unified quantification framework considering three types of input uncertainties, and Monto Carlo method, respectively. The belief and plausibility measures under different reliability limit values  $c$  are calculated and shown in Fig. 5a. The comparison of belief and plausibility results in the moderate reliability ( $0.90 < \text{Reliability} < 0.99$ ) and high reliability ( $\text{Reliability} > 0.99$ ) levels are shown in Fig. 5b and Fig. 5c, respectively. The uncertainty of performance function  $G$  under limit state  $c = 1000$  are listed in Table 4. The proposed framework can obtain the accurate uncertainty quantification results of performance function, and the calculation number of system performance function decreases from  $6 \times 10^7$  to  $10^3$ , the computational complexity decreases seriously.

### 6.2 Example 2-crank slider mechanism

The example of a crank-slider mechanism in Ref. (Du 2008) is used to demonstrate the effectiveness and application of the proposed method. The mechanism is shown in Fig. 6. The length of the crank  $x_1$ , the length of the coupler  $x_2$  and the

external force  $x_3$  are strong statistical variables. The distributions of the strong statistical variables are given in Table 5. Difference from Ref. (Du 2008), the material Young's modulus  $y_1$  and the yield strength  $y_2$  of the coupler are assumed to be sparse variables. The available data of  $y_1$  are points  $\{195, 200, 204\}$ (GPa) and intervals  $[180, 197]$ ,  $[199, 209]$ ,  $[210, 212]$ (GPa). The available data of  $y_2$  are points  $\{280, 290, 298\}$ (MPa) and intervals  $[270, 282]$ ,  $[288, 292]$ ,  $[299, 313]$ (MPa). The friction coefficient  $z_1$  and the offset  $z_2$  are interval variables, their sub-intervals and BPAs are shown in Table 6. The internal diameter  $d_1$  and external diameter  $d_2$  of coupler are 25 mm and 60 mm, respectively.

The system performance function is defined by the difference between the critical load and the axial load, which is written in (22).

$$G = G(\mathbf{X}, \mathbf{Y}, \mathbf{Z}) = \frac{\pi^3 y_1 (d_2^4 - d_1^4)}{64 x_2^2} - \frac{x_3 (x_2 - x_1)}{\sqrt{(x_2 - x_1)^2 - z_2^2 - z_1 z_2}} \quad (22)$$

Through comparing the AIC values of the eight candidate distribution types, the uncertainty of the yield strength  $y_1$  is represented using weight sum of Normal distribution, Gamma distribution, Weibull distribution, and Extreme distribution with weight ratios of 0.352, 0.091, 0.284, and 0.273, respectively. The uncertainties of distribution parameters are shown in Fig. 7.

The failure event is defined by  $F = \{\mathbf{X}, \mathbf{Y}, \mathbf{Z} | G(\mathbf{X}, \mathbf{Y}, \mathbf{Z}) < c\}$ , where  $c$  is the limit state. The Kriging model is constructed based on  $2 \times 10^3$  samples, and the RMSE of the Kriging model is less than 0.0321. The belief and plausibility measures under different reliability limit values  $c$  are calculated using the proposed method and Monto Carlo method, and there are  $9 \times 10^4$  samples under every limit value  $c$  in MCS method. The results are shown in Fig. 8. In the range of limit state  $180 \leq c \leq 250$ , the belief and plausibility value of the proposed method are much the same as that of MCS method, which shows that the unified uncertainty quantification method presented in the paper is effective and practical.

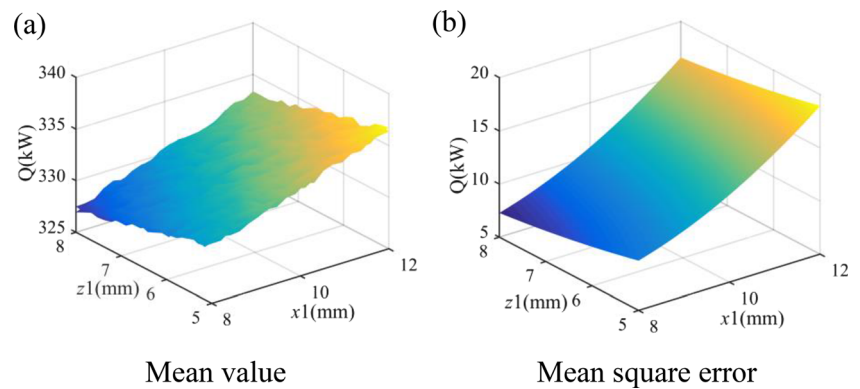
### 6.3 Example 3-heat exchanger

A practical heat exchanger structure shown in Fig. 9 is considered, which consists of inlet header, fin channels

**Table 10** Identification results of sparse variables  $y_1 \sim y_4$

Variables	Distribution type	Weight ratio	Distribution parameters	
$y_1$	Normal	0.3713	360	2
	Extreme	0.6287	361	2
$y_2$	Normal	0.2680	8	1
	Weibull	0.5674	8.1	10
	Extreme	0.1646	8.1	1
$y_3$	Normal	0.4612	315	3
	Weibull	0.0585	315.5	50
	Extreme	0.4803	316.5	3
$y_4$	Normal	0.4707	0.097	0.005
	Extreme	0.5293	0.098	0.004

**Fig. 12** The Kriging model of example 3



and outlet header. The hot streams and cold streams flow from inlet header to parallel fin channels, and then flow out in the outlet header. The fin channels, which has dimensions of  $1000 \times 480 \times 380$  mm (length  $\times$  width  $\times$  height), are the important heat transfer region of hot and cold streams, and the heat transfer rate between hot streams (air) and cold streams (water) through separating fins (Fig. 10) is an important index to evaluate its performance. The fin height of hot stream  $x_1$  and cold stream  $x_2$  are strong statistical variables because of manufacturing error, and their distributions are listed in Table 7. The fin space  $z_1$  of hot stream and cold stream is fabricated in the predetermined length intervals, so  $z_1$  is interval variable, its sub-intervals and BPA are listed in Table 8. Only sparse sampling point and interval data can be acquired using data acquisition machine for the input temperature and velocity of hot and cold streams  $y_1 \sim y_4$ , so  $y_1 \sim y_4$  are sparse variables, and their sampling values are listed in Table 9.

The heat transfer rate is the most important index for heat exchanger, so the limit state function can be defined in (23).

$$G = Q(x_1, x_2, y_1, y_2, y_3, y_4, z_1) > Q_m \quad (23)$$

Where  $Q$  is the heat transfer rate, which can be calculated after simulating the temperature distribution in the fin channel using finite element analysis (FEA) method, and  $Q_m = 500$  kW is the allowable minimum heat transfer rate. In the FEA, the simulation model is constructed and meshed using hexahedral elements in Gambit, the inlets of streams are set to be velocity inlet, and the

outlets are set to be outflow. The interface of flow regions and fin regions are set to be coupled to calculate the heat transfer rate. The meshed structures are analyzed in Fluent, and then the temperature and pressure distributions in the heat exchange are calculated. The simulation results of temperature distributions in the hot streams are plotted in Fig. 11.

The distribution types and distribution parameters of sparse variables  $y_1 \sim y_4$  are identified using the rules in Section 3, the results are shown in Table 10. The Kriging model is constructed based on the FEA results of  $2 \times 10^3$  samples, which is shown in Fig. 12. Then the presented unified uncertainty quantification method and the conventional MCS method are applied to solve the problem. The total number of the function calls by the proposed algorithm is  $5 \times 10^3$ , and the calculation can be finished in 2 h; while the total number of performance function evaluation is  $7.2 \times 10^8$  in MCS method, and the calculation will be finished in 3d. The belief and plausibility value of the proposed method and MCS method are listed in Table 11. The differences between the two results are less than 1.7%, which shows that the presented method in the paper is effective and practical.

## 7 Conclusions

A unified uncertainty quantification method considering strong statistical variables, sparse variables and interval variables simultaneously is proposed. To increase accuracy and decrease computational difficulty in the representation of sparse variables, mixed distribution types are identified and the calculated continuous uncertain distribution parameters are dispersed based on their probability density function. Then, a unified uncertainty quantification framework is constructed, the strong variables and sparse variables are represented through random samplings of distribution functions, and the interval variables are represented using evidence theory. The belief and plausibility of system performance function are calculated, and the kriging model of the performance function is used to increase computational efficiency. The

**Table 11** Belief and plausibility measure for example 3

	Proposed method	Monte Carlo method
Belief value	0.7510	0.7469
Plausibility value	0.9541	0.9385
Evaluation number	$5 \times 10^3$	$7.2 \times 10^8$



proposed method and MCS method are used to calculate the failure probability of three examples. The exact solutions and the results solved by the presented method are almost the same. Moreover, the required sample numbers of the proposed method is fewer than that in conventional method.

### Compliance with ethical standards

**Funding** This work was supported by the National Natural Science Foundation of China under grant [number: 51505421, 51375451, U1608256], and Zhejiang Provincial Natural Science Foundation of China under grant [number: LY15E050015].

### References

- Arendt PD, Apley DW, Chen W (2012) Quantification of model uncertainty: calibration, model discrepancy, and identifiability. *J Mech Des* 134:100908. doi:10.1115/1.4007390
- Bae HR, Grandhi RV, Canfield RA (2004) Epistemic uncertainty quantification techniques including evidence theory for large-scale structures. *Comput Struct* 82:1101–1112. doi:10.1016/j.compstruc.2004.03.014
- Beer M, Ferson S, Kreinovich V (2013) Imprecise probabilities in engineering analyses. *Mech Syst Signal Process* 37:4–29. doi:10.1016/j.ymsp.2013.01.024
- Chee CS (2017) A mixture model-based nonparametric approach to estimating a count distribution. *Comput Stat Data Anal* 109:34–44. doi:10.1016/j.csda.2016.11.012
- Cho H, Choi KK, Gaul NJ, Lee I, Lamb D, Gorsich D (2016) Conservative reliability-based design optimization method with insufficient input data. *Struct Multidiscip Optim* 54(6):1609–1630. doi:10.1007/s00158-016-1492-4
- Choi CK, Yoo HH (2016) Stochastic inverse method to identify parameter random fields in a structure. *Struct Multidiscip Optim* 54(6):1557–1571. doi:10.1007/s00158-016-1534-y
- Contreras AA, Olivier PLM, Wilkins A, Omar MK (2016) Multi-model polynomial chaos surrogate dictionary for Bayesian inference in elasticity problems. *Probabilist Eng Mech* 46:107–119. doi:10.1016/j.probengmech.2016.08.004
- Du X (2006) Uncertainty analysis with probability and evidence theories. Paper presented at the ASME 2006 international design engineering technical conference, USA
- Du X (2008) Unified uncertainty analysis by the first order reliability method. *J Mech Des* 130:091401. doi:10.1115/1.2943295
- Gal Y, Ghahramani Z (2016) Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In international conference on machine learning (pp. 1050–1059).
- Gelman A, Hwang J, Vehtari A (2013) Understanding predictive information criteria for Bayesian models. *Stat Comput* 24:997–1016. doi:10.1007/s11222-013-9416-2
- Hu Z, Ao D, Mahadevan S (2017) Calibration experimental design considering field response and model uncertainty. *Comput Methods Appl Mech Eng* 318:92–119. doi:10.1016/j.cma.2017.01.007
- Jiang Z, Chen S, Apley DW, Chen W (2016) Reduction of epistemic model uncertainty in simulation-based multidisciplinary design. *J Mech Des* 138:081403. doi:10.1115/1.4033918
- Kang YJ, Lim OK, Noh Y (2016) Sequential statistical modeling method for distribution type identification. *Struct Multidiscip Optim*. doi:10.1007/s00158-016-1567-2
- Li G, Lu Z, Li L, Ren B (2016) Aleatory and epistemic uncertainties analysis based on non-probabilistic reliability and its kriging solution. *Appl Math Model* 40:5703–5716. doi:10.1016/j.apm.2016.01.017
- Liu X, Yin L, Hu L, Zhang Z (2017) An efficient reliability analysis approach for structure based on probability and probability box models. *Struct Multidiscip Optim*. doi:10.1007/s00158-017-1659-7
- Matsumura T, Haftka RT (2013) Reliability based design optimization modeling future redesign with different epistemic uncertainty treatments. *J Mech Des* 135:091006. doi:10.1115/1.4024726
- Mullins J, Ling Y, Mahadevan S, Sun L, Strachan A (2016) Separation of aleatory and epistemic uncertainty in probabilistic model validation. *Reliab Eng Syst Saf* 147:49–59. doi:10.1016/j.res.2015.10.003
- Nannapaneni S, Hu Z, Mahadevan S (2016) Uncertainty quantification in reliability estimation with limit state surrogates. *Struct Multidiscip Optim*. doi:10.1007/s00158-016-1487-1
- Oberkampf WL, Helton JC, Joslyn CA, Wojtkiewicz SF, Ferson S (2004) Challenge problems: uncertainty in system response given uncertain parameters. *Reliab Eng Syst Saf* 85:11–19. doi:10.1016/j.res.2004.03.002
- Paulson EJ, Starkey RP (2013) Development of a multistage reliability-based design optimization method. *J Mech Des* 136:011007. doi:10.1115/1.4025492
- Ren Z, Cho H, Yeon J, Koh CS (2015) A new reliability analysis algorithm with insufficient uncertainty data for optimal robust design of electromagnetic devices. *IEEE Trans Magn* 51:1–4. doi:10.1109/tmag.2014.2360753
- Sankararaman S, Mahadevan S (2011) Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data. *Reliab Eng Syst Saf* 96:814–824. doi:10.1016/j.res.2011.02.003
- Sankararaman S, Mahadevan S (2013) Distribution type uncertainty due to sparse and imprecise data. *Mech Syst Signal Process* 37:182–198. doi:10.1016/j.ymsp.2012.07.008
- Shafer G (1976) A mathematical theory of evidence. Princeton university press, Princeton
- Simoen E, De Roeck G, Lombaert G (2015) Dealing with uncertainty in model updating for damage assessment: a review. *Mech Syst Signal Process* 56–57:123–149. doi:10.1016/j.ymsp.2014.11.001
- Søren NL, Hans BN and Jacob S (2017) DACE a matlab kriging toolbox (Version 2.0). <http://www2.imm.dtu.dk/projects/dace/>. Accessed 7 February 2017
- Suryawanshi A, Ghosh D (2015) Reliability based optimization in aero-elastic stability problems using polynomial chaos based metamodels. *Struct Multidiscip Optim*. doi:10.1007/s00158-015-1322-0
- Taguri M, Matsuyama Y, Ohashi Y (2014) Model selection criterion for causal parameters in structural mean models based on a quasi-likelihood. *Biometrics* 70:721–730. doi:10.1111/biom.12165
- Tao YR, Cao L, Huang ZHH (2016) A novel evidence-based fuzzy reliability analysis method for structures. *Struct Multidiscip Optim* 55:1237–1249. doi:10.1007/s00158-016-1570-7
- Wang H, Zeng Y, Yu X, Li G, Li E (2015) Surrogate-assisted Bayesian inference inverse material identification method and application to advanced high strength steel. *Inverse Probl Sci Eng* 24(7):1133–1161. doi:10.1080/17415977.2015.1113960
- Xi Z, Youn BD, Jung BC, Yoon JT (2014) Random field modeling with insufficient field data for probability analysis and design. *Struct Multidiscip Optim* 51:599–611. doi:10.1007/s00158-014-1165-0
- Xia Y, Liu C, Li Y, Liu N (2017) A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. *Expert Syst Appl* 78:225–241. doi:10.1016/j.eswa.2017.02.017
- Yoo D, Lee I (2013) Sampling-based approach for design optimization in the presence of interval variables. *Struct Multidiscip Optim* 49:253–266. doi:10.1007/s00158-013-0969-7
- Zaman K, McDonald M, Mahadevan S (2011) Probabilistic framework for uncertainty propagation with both probabilistic and interval variables. *J Mech Des* 133:021010. doi:10.1115/1.4002720
- Zhu Z, Du X (2016) Reliability analysis with monte carlo simulation and dependent kriging predictions. *J Mech Des* 138:121403. doi:10.1115/1.4034219