Bayesian Reliability Analysis With Evolving, Insufficient, and Subjective Data Sets

Pingfeng Wang
e-mail: pfwang@umd.edu

Byeng D. Youn
e-mail: bdyoun@umd.edu

Zhimin Xi
e-mail: zxi@umd.edu

University of Maryland,
College Park, MD 20742

Artemis Kloess
R&D Center,
General Motors Corp.,
Warren, MI 48090
e-mail: artemis.kloess@gm.com

This paper presents a new paradigm of system reliability prediction that enables the use of evolving, insufficient, and subjective data sets. The data sets can be acquired from expert knowledge, customer survey, inspection and testing, and field data throughout a product life-cycle. In order to handle such data sets, this research integrates probability encoding methods to a Bayesian updating mechanism. The integrated tool is called Bayesian Information Toolkit. Subsequently, Bayesian Reliability Toolkit is presented by incorporating reliability analysis to the Bayesian updating mechanism. A generic definition of Bayesian reliability is introduced as a function of a predefined confidence level. This paper also finds that there is no data-sequence effect on the updating results. It is demonstrated that the proposed Bayesian reliability analysis can predict the reliability of door closing performance in a vehicle body-door subsystem, where available data sets are insufficient, subjective, and evolving. [DOI: 10.1115/1.4000251]

1 Introduction

In the past three decades, engineering analysis and design methods have advanced to improve reliability of an engineering product system while considering uncertainties in the system. However, little attention has been made to data modeling with evolving, insufficient, and subjective data sets. In this paper, we refer the data, which are not static but evolve with time as "evolving data," refer the data, which are not sufficient to fully characterize random behavior as insufficient data and, similarly, refer the data, which pertain to or perceived only by individuals as subjective data. To be clear, aleatory uncertainty is defined as the uncertainty, which arises because of unpredictable variation in the performance of the system, whereas epistemic uncertainty is defined as the uncertainty, which is due to a lack of knowledge about the behavior of the system that is conceptually resolvable. More specifically, aleatory uncertainties are considered to be represented by statistical distributions, whereas epistemic uncertainties are considered to be represented by limited data sets in this paper. Most probabilistic analysis and design approaches still depend on the assumed probabilistic models of system inputs without engaging raw data. The research that predicts product reliability with evolving, insufficient, and subjective data sets is strongly in demand for engineering analysis and design. This has been acknowledged as one of the challenging problems in the field of engineering design and uncertainty analysis and have not been resolved yet [1]. Also, the subjective data analysis, for example, customer satisfaction data analysis, for engineering design has been identified as one of the future research directions of engineering design [2]. Insufficient data have become a major bottleneck in engineering analysis [3]. In addition, subjective data are an essential component to understand the interface between customers and products for product design [4,5].

Acquiring and modeling uncertainty data are essential for quantifying uncertain behaviors of engineered systems in reliability analysis and design optimization process. Subjective data, such as the opinions from experts and customers, are quite important at an early stage of system design. To deal with subjective data, the expert opinion approach has been further developed for uncertainty quantification of engineered systems in a wide range of engineering problems [6–10]. The expert opinion approach is basically a data collection scheme, which includes selection of experts, expert interview session, and probability encoding. In this approach, the probability encoding technique is a process of obtaining statistical information of interested subjective data in the form of probability density functions or simple probability values and occurrence rates. Spetzler et al. [11] provided an overview of different probability encoding methods. These methods will be further discussed and integrated with a Bayesian updating technique in Sec. 2. This integration forms a more sophisticated tool of handling the evolving subjective data sets, referred to as Bayesian Information Toolkit (BIT).

It is common to encounter insufficient data sets in practical engineering applications. When available data is insufficient, the classical probability theory may be improper to model uncertainties because it may lead to a result with a relatively low confidence. To deal with insufficient data sets, different methods have been developed for reliability analysis and design optimization. Methods are based on various nondeterministic theories: the possibility theory [12–16], the evidence theory [17–19], and Bayes' theory [20–22]. Although different methods have been developed to deal with subjective and insufficient data sets, evolving data sets have little been considered in these methods. Since Bayes theory provides a systematic framework of aggregating and updating uncertain information, this paper presents reliability analysis method based on the Bayes theory, referred to as Bayesian Reliability Toolkit (BRT).

To ensure the reliability of the product system, diverse design methodologies have been developed, such as reliability-based design optimization (RBDO) [23–26], possibility-based design optimization (PBDO) [27,28], evidence-based design optimization (EBDO) [29], and Bayesian RBDO [22]. Besides, some recent publications [30–33] delivered rigorous studies to deal with all kinds of uncertainty (e.g., aleatory/epistemic, discrete/continuous, and statistical/fuzzy) for system analysis and design. Such research activities have focused on how to assess reliability effectively by simply assuming nondeterministic models of random system inputs without engaging raw data [24,34,35]. Among these design methodologies, Bayesian approaches have been widely used in many engineering and science fields, where data are progressively accumulated. For example, Bayesian reliability analysis has been applied to series systems of binomial (safe or fail) subsystems and components [36], to reliability assessment of power

Contributed by the Design Automation Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received August 2, 2008; final manuscript received August 17, 2009; published online October 13, 2009. Assoc. Editor: Panos Y. Papalambros.

systems [37], to the effectiveness of reliability growth testing [38], to robust tolerance control and parameter design in the manufacturing process [39], and to input uncertainty modeling [40]. Two advanced Bayesian (maximum likelihood and parsimony) methods have been compared for molecular biology applications [41]. Bayesian updating has been implemented using the Markov chain Monte Carlo (MCMC) simulation for structural models and reliability assessment [42]. Dynamic object oriented Bayesian networks have been proposed for complex system reliability modeling [43].

Despite numerous efforts, it has been a great challenge to predict uncertain system performances while considering evolving, insufficient, and subjective data sets. The objective of this research is thus to establish a new paradigm of reliability prediction that enables the use of evolving, insufficient, and subjective data sets. The data sets come from expert knowledge, customer survey, inspection and testing, and field data over the entire product lifecycle. In order to handle such data sets, this research integrates probability encoding methods to a Bayesian updating mechanism. It is referred to as BIT. Subsequently, BRT is presented by incorporating reliability analysis to the Bayesian updating mechanism. A generic definition of Bayesian reliability is introduced as a function of a predefined confidence level. This paper also investigates the effect of the sequence of evolving data sets on Bayesian updating and finds that there is no data-sequence effect on the updating results. Integrating with the authors' earlier work in Bayesian reliability-based design optimization [22], this paper presents a Bayesian Information, Reliability and Design (BIRD) software.

In this paper, the proposed approach is applied for reliability prediction of door closing performance in a vehicle door system. The vehicle door system is of special concern due to its frequency of use and its engineering challenge with respect to design, assembly, and operation. A considerable amount of engineering effort is spent conducting hardware-based or analytical experiments to generate information for supporting engineering decisions during the vehicle development process. At the conceptual stage, the uncertainty characterization of this information is largely based on expert judgment and data from current or past designs. As the design matures, analysis results and test data are collected to quantify the uncertainty; however, data are usually of limited sample size. The door seal design engineer, for example, needs to know the requirements for a door seal system that isolates the passenger compartment from the external environment while simultaneously allowing the door to be closed with minimal effort. A door system design must satisfy a multitude of functional and engineering requirements. The functional requirements are deduced from the voice of the customer and include, for example, excellent exterior appearance/fit, interior quietness, protection from water leaks and dust intrusion, and an easy to open/close door. The functional requirements must be translated into measurable engineering requirements, and the engineering solutions should be simple and include manufacturing restrictions. Due to the inherent uncertainties associated with the voice of the customer, manufacturing processes, material properties, etc., engineers must seek an appropriate performance evaluation metric and corresponding method that can incorporate and evaluate the effect of those uncertainties.

2 Bayesian Information, Reliability, and Design Toolkit

This section presents the integration of probability encoding methods and reliability analysis to the Bayesian updating mechanism.

2.1 Bayesian Updating Techniques. As mentioned earlier, the evolving, insufficient, and subjective data sets can be obtained through either measurement or survey during the product lifecycle. To make use of such information for the purposes of product performance evaluation and design, BIT employs a Bayesian updating technique. This subsection gives an introduction of the Bayesian updating technique and describes the effect of the sequence of the evolving data sets on Bayesian updating.

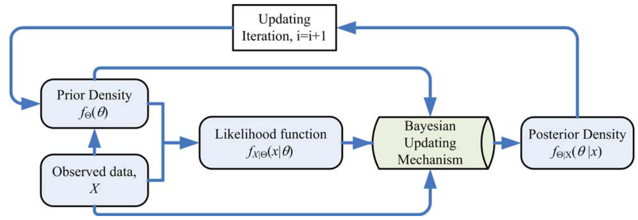

Fig. 1 Process of Bayesian updating

Let X be a random variable with a probability density function $f(x,\theta)$, $\theta \in \Omega$. From the Bayesian point of view, θ is interpreted as a realization of a random variable Θ with a probability density $f_\Theta(\theta)$. The density function expresses what one thinks about the occurring frequency of Θ before any future observation of X is taken, that is, a prior distribution. Based on Bayes' theorem, the posterior distribution of Θ given a new observation X can be expressed as

$$f_{\Theta|X}(\theta|x) = \frac{f_{X,\Theta}(x,\theta)}{f_X(x)} = \frac{f_{X|\Theta}(x|\theta)\cdot f_\Theta(\theta)}{f_X(x)} \quad (1)$$

The Bayesian approach is used for updating information about the parameter θ. First, a prior distribution of Θ must be assigned before any future observation of X is taken. Then, the prior distribution of Θ is updated to the posterior distribution as the new data for X is employed. The posterior distribution is set to a new prior distribution, and this process can be repeated with an evolution of data sets. This updating process can be briefly illustrated in Fig. 1.

Let us consider a Bayesian normal inference model as the example to illustrate the Bayesian updating process. In this example, a random variable (X) follows a normal distribution with unknown mean value θ and known standard deviation σ. The random variable is realized with N samples $(x_1, x_2, \ldots, x_N)$, where N is small. The Bayesian inference is used to update the prior knowledge of the mean value θ based on the observed data. The likelihood function for these N observations is expressed as

$$p(X|\theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2\sigma^2}(x_i-\theta)^2\right] \quad (2)$$

In Bayesian probability theory, a class of prior probability distributions $f_\Theta(\theta)$ is said to be conjugate to a class of likelihood functions $f_{X|\Theta}(x|\theta)$ if the resulting posterior distributions $f_{\Theta|X}(\theta|x)$ are in the same family as $f_\Theta(\theta)$. For example, if the likelihood function is Gaussian, choosing a Gaussian prior ensures that the posterior distribution is also Gaussian. In this example, suppose a conjugate prior distribution for θ, which is a normal distribution with mean, u, and variance, τ^2, then due to the conjugate property, the posterior distribution can be obtained through the Bayesian updating process, which also follows the normal distribution with the mean and variance as [44]

$$\tilde{u} = \frac{\bar{X}\cdot N\cdot\sigma^{-2} + u\cdot\tau^{-2}}{N\cdot\sigma^{-2}+\tau^{-2}}, \quad \tilde{\tau}^2 = \frac{1}{N\cdot\sigma^{-2}+\tau^{-2}} \quad (3)$$

To apply the Bayesian updating technique to the modeling of evolving data sets, one may be interested in how the sequence of evolving data sets affects the final updating results. Suppose there are two sets of data X_A and X_B with sample sizes N_A and N_B, respectively, the effect of the sequence of the data sets on Bayesian updating is studied by comparing the final results of different updating sequences A-B and B-A. First the updating sequence

A-B is considered. When the data set X_A is used to update the prior information of θ the result of the first updating can be obtained according to Eq. (3) as

$$\tilde{u}_1 = \frac{\bar{X}_A \cdot N_A \cdot \sigma^{-2} + u \cdot \tau^{-2}}{N_A \cdot \sigma^{-2} + \tau^{-2}}$$

$$\tilde{\tau}_1^2 = \frac{1}{N_A \cdot \sigma^{-2} + \tau^{-2}} \tag{4}$$

By treating the results in Eq. (4) as prior distribution for θ, the updating process continues as the data set X_B is added. The parameters can be updated as

$$\tilde{u}_2 = \frac{\bar{X}_B \cdot N_B \cdot \sigma^{-2} + \tilde{u}_1 \cdot \tilde{\tau}_1^{-2}}{N_B \cdot \sigma^{-2} + \tilde{\tau}_1^{-2}} = \frac{(\bar{X}_B \cdot N_B + \bar{X}_A \cdot N_A) \cdot \sigma^{-2} + u \cdot \tau^{-2}}{(N_B + N_A) \cdot \sigma^{-2} + \tau^{-2}}$$

$$\tilde{\tau}_2^{\,2} = \frac{1}{N_B \cdot \sigma^{-2} + \tilde{\tau}_1^{-2}} = \frac{1}{(N_B + N_A) \cdot \sigma^{-2} + \tau^{-2}} \tag{5}$$

Following the same procedure for updating sequence B-A, the same final results as Eq. (5) will be obtained. Involving k sets of evolving data in the updating process, regardless of the updating sequence, the final updating results are

$$\tilde{u}_k = \frac{\sigma^{-2} \cdot \sum_{i=1}^{k} \bar{X}_i \cdot N_i + u \cdot \tau^{-2}}{\sum_{i=1}^{k} N_i \cdot \sigma^{-2} + \tau^{-2}}$$

$$\tilde{\tau}_k^2 = \frac{1}{\sum_{i=1}^{k} N_i \cdot \sigma^{-2} + \tau^{-2}} \tag{6}$$

A Bayesian inference model is called a conjugate model if the conjugate prior distribution is used. For conjugate Bayesian inference models, the updating results are independent of the sequence of data sets. Conjugate models of Bayesian updating are quite useful for uncertainty modeling with evolving data sets, since the prior and posterior distributions are given in a closed form. However, it is found that the Bayesian updating results often depend on the selection of a prior distribution in the conjugate models. Besides, the available conjugate Bayesian models are limited. To eliminate the dependency and the limitation, a nonconjugate Bayesian updating model can be developed using Markov chain Monte Carlo methods. This is, however, more computationally intensive.

2.2 Bayesian Information Toolkit. As mentioned earlier, a great challenge exists in dealing with evolving, insufficient, and subjective data sets while performing reliability analysis. BIT is developed by integrating probability encoding methods with the Bayesian updating technique to systematically elicit subjective data from subjects (e.g., experts and customers), to model the subjective data with statistical distributions, and to update the uncertainty distributions with evolving subjective data sets. The probability encoding methods are briefly reviewed and followed by the integration of probability encoding methods with Bayesian updating techniques. One example is used to demonstrate the proposed BIT.

2.2.1 Probability Encoding Methods. To systemically extract and quantify subjective information that comes from individual judgment about uncertain quantities, the probability encoding [11,45,46] methods are employed in BIT. The methods employ an interview process, and most are based on questions for which the answers can be represented as points on a cumulative distribution function. The different encoding methods used vary according to whether they ask a subject to assign probabilities (P), values (V), or both. The three basic types of the encoding methods are listed below.

- *P*-methods require the subject to respond by specifying points on the probability scale while the values are fixed.
- *V*-methods require the subject to respond by specifying points on the value scale while the probabilities remain fixed.
- *PV*-methods ask questions that must be answered on both scales jointly; the subject essentially describes points on the cumulative distribution.

The probability encoding methods consist of a set of questions that the subject responds to either directly by providing numbers or indirectly by choosing between simple alternatives or bets. In the direct response mode, the subject is asked questions that require numbers as answers which will be given in the form of either values or probabilities depending on the method being used. In the indirect response mode, the subject is asked to choose between two or more bets, for example, the probability "wheel" method [11]. The bets are adjusted until the subject is indifferent to choosing between them. This indifference can then be translated into a probability or value assignment. Besides choosing between bets, another procedure is to ask the subject to choose between events defined on the value scale for the uncertain quantity, where each event represents a set of possible outcomes for the uncertain quantity. Subjective data in examples shown in Sec. 3 were obtained using the direct probability encoding procedure.

2.2.2 Integration of Probability Encoding Methods With Bayesian Updating Technique. In this study, the probability encoding methods are used to elicit subjective survey data in the forms of the probability (P) and subjective response (V). To apply Bayesian updating technique for subjective data modeling, the subjective survey data (probabilities and subjective responses) need to be transferred to corresponding statistical parameter data. The inverse transformation of a parametric cumulative distribution function (CDF) maps the subjective survey data into the parameter data of the CDF. For example, each subjective data point (Z_p, P) is used to fit a parametric CDF and computes corresponding parameter data (mean μ and standard deviation σ) through an inverse transformation of the CDF. This Bayesian technique updates subjective data by treating all parameter data as a new data set.

The procedure below can be followed for modeling uncertainties with subjective data sets by integrating probability encoding methods with the Bayesian updating technique.

- Step 1: specify uncertain quantities interested.
- Step 2: choose probability encoding methods and prepare surveys.
- Step 3: interview subjects and collect subjective data sets.
- Step 4: choose suitable Bayesian updating models.
- Step 5: transfer subjective data to interested model parameter data by inverse CDF analysis.
- Step 6: updating the uncertainty model parameters by integrating model parameter data obtained from Step 5 using the Bayesian updating technique.

2.2.3 One Example for BIT. This section presents the example of the probability encoding technique using the *PV*-method. Suppose that the highest temperature in one day is a random variable, which follows a normal distribution with unknown mean value θ and known standard deviation of 2°C. The interested uncertain quantity is the mean value of the highest temperature, θ. The *PV*-method is used in this example, and two subjects are questioned about tomorrow's highest temperature value (V) and corresponding probabilities (P). Results are summarized in Table 1 and shown in Fig. 2. After this data acquisition process, a Bayesian updating model should be specified to aggregate these subjective

Table 1 Results of the temperature survey and inverse CDF analysis

Subject I			Subject II		
Temp.	Prob.	θ	Temp	Prob.	θ
22	0.100	24.563	23	0.080	25.810
25	0.160	26.989	25	0.100	27.563
26	0.400	26.507	27	0.250	28.349
27	0.550	26.749	28	0.600	27.493
28	0.750	26.651	29	0.850	26.927
29	0.900	26.437	30	0.950	26.710
30	0.950	26.710	31	0.990	26.348
31	0.995	25.850	33	0.999	26.816

data sets and to update the distribution of θ. In this example, a conjugate Bayesian normal inference model, as discussed in Sec. 2.1, is used, and θ itself is modeled as a normal distribution. With the Bayesian updating model, the subjective data sets can then be used for inverse CDF analysis to get the corresponding model parameter data (data for θ). Taking the first data set [22, 0.1] from Subject I as an example, an inverse transformation of the CDF is carried out by treating this data set as one point of the normal distribution CDF curve, which is

$$p = \int_{-\infty}^{V} f(x,\theta,\sigma)dx \tag{7}$$

where p equals 0.1, V equals 22, and σ equals 2 for this case. As the only unknown in Eq. (7) is θ, the inverse calculation of this equation can provide a data for θ as 24.563. Similar analysis is applied on each set of data for both Subjects I and II, which results in two sets of data of θ, as shown in Table 1.

These two sets of data will then be used by Bayesian updating to update the distribution of θ. For example, if a prior distribution $N(26,1^2)$ is assumed for θ with the first set of data from Subject I, the posterior distribution is updated to $N(26.932,0.2655^2)$, as posterior I shown in Fig. 3. Aggregating the second set of data from Subject II, the posterior distribution is then updated to $N(26.965,0.192^2)$, as the posterior II shown in Fig. 3.

2.3 Bayesian Reliability Toolkit. In many engineering applications, outcomes of events from repeated trials can be a binary manner, such as occurrence or nonoccurrence, success or failure, good or bad, etc. In such cases, random behavior can be modeled with a discrete probability distribution model. In addition, if the events satisfy the additional requirements of a Bernoulli sequence,

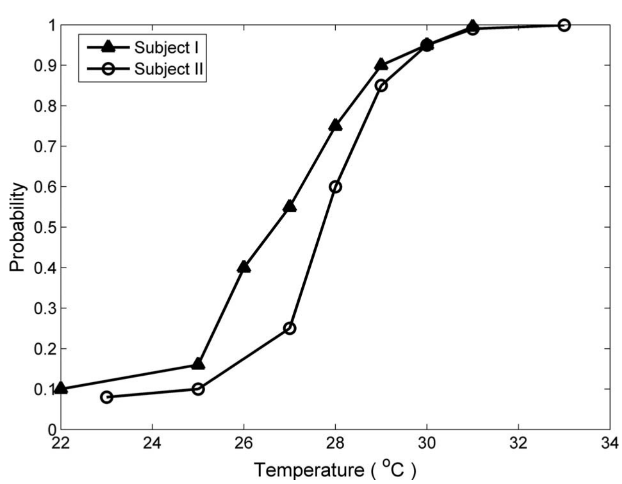

Fig. 2 Results of the temperature survey

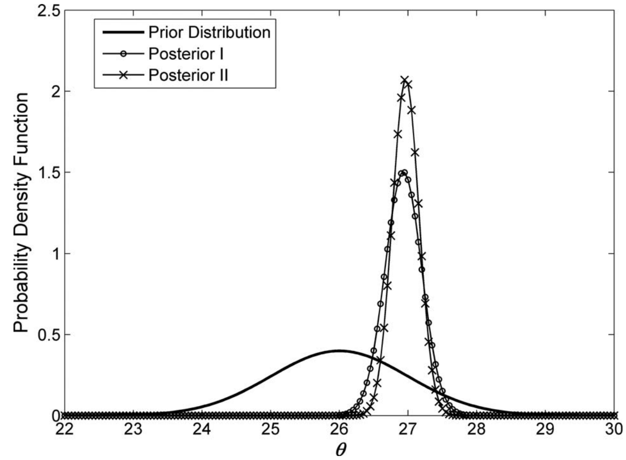

Fig. 3 Bayesian updating of the PDF of θ

that is to say, if the events are statistically independent and the probability of occurrence or nonoccurrence of events remains constant, they can be mathematically represented by the *binomial distribution* [47]. In other words, if the probability of an event occurrence in each trial is r and the probability of nonoccurrence is $(1-r)$, then the probability of x occurrences out of a total of N trials can be described by the probability mass function (PMF) of a binomial distribution as

$$\Pr(X=x,N|r) = \binom{N}{x} r^x(1-r)^{N-x}, \quad x=0,1,2,\ldots,N \tag{8}$$

where the probability of success identified in the previous test, r, is the parameter of the distribution.

In Eq. (8), the probability of x/N (x occurrences out of N trials) can be calculated when a prior distribution on r is provided. This inference process seeks to update r based on the outcomes of the trials. Given x occurrences out of a total of N trials, the probability distribution of r can be calculated using Bayes' rule as [48]

$$f(r|x) = \frac{f(x|r)f(r)}{\int_0^1 f(x|r)f(r)dr} \tag{9}$$

where $f(r)$ is the prior distribution of r, $f(r|x)$ is the posterior distribution of r, and $f(x|r)$ is the likelihood of x for a given r. The integral in the denominator is a normalizing factor to make the probability distribution proper. The prior distribution is known for r, prior to the current trials. In this paper, a uniform prior distribution is used to model r bounded in [0, 1]. However, it is possible to obtain a posterior distribution with any type of a prior distribution.

For Bayesian reliability predictions, both a prior reliability distribution (r) and the number (x) of safety occurrences out of the total number of test data sets N must be known. If the prior reliability distribution (r) is unavailable, it will be simply modeled with a uniform distribution, $r \sim U(a,\ b)$, where $a<b$ and $a,b \in [0,\ 1]$. At an early design stage, it can be modeled using reliability from the previous product designs or expert opinions. Alternatively, if the reliability distribution has been predicted with a data set in a precedent test, this reliability distribution will be used as the prior reliability distribution and updated to a posterior reliability distribution with new test data. Bayesian binomial inference model can be used to update the prior knowledge of reliability (r), which is a parameter of a binomial distribution. In this inference model, the binomial distribution likelihood function is used for test data, whereas the conjugate prior distribution of this likelihood function is used for reliability (r), which is a beta distribution. The PDF of the beta distribution is expressed as

$$f(r) = \frac{1}{B(\alpha,\beta)} r^{\alpha-1}(1-r)^{\beta-1}$$

$$\left(B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt \right) \quad (10)$$

where α and β are two parameters. For a simple case, $\alpha=\beta=1$ represents a uniform distribution over [0, 1]. If this uniform distribution is used as a prior distribution for r, the likelihood function $f(x|r)$ can be obtained using Eq. (8) and the posterior distribution $f(r|x)$ using Eq. (9). It follows a beta distribution with $\alpha = x+1$ and $\beta=N-x+1$. This posterior distribution represents the probability distribution of reliability, which is a function of x and N. With k sets of evolving testing data sets, the final updating result for r is also a beta distribution with

$$\alpha = 1 + \sum_{i=1}^{k} x_i \quad \text{and} \quad \beta = 1 + \sum_{i=1}^{k} (N_i - x_i) \quad (11)$$

As it is a conjugate Bayesian inference model, there is also no data-sequence effect on updating results.

When only epistemic uncertainties are engaged to assess reliability, its PDF can be modeled using the beta distribution in Eq. (10) by counting the number of safety occurrences, x. In general, both aleatory and epistemic uncertainties appear in most engineering design problems. In such situations, the PDF of reliability can be similarly obtained through a Bayesian reliability analysis. To build the PDF of reliability, reliability analysis must be performed at every data point for epistemic uncertainties while considering aleatory uncertainties. Different reliability measures, $R_k=R(x_{e,k})$, are obtained at different sample points for epistemic uncertainties. In Eq. (10), $\alpha=x+1$ and $\beta=N-x+1$, where $x=\Sigma R_k$. Then, the PDF of reliability r with a uniform prior distribution is updated to $R(\mathbf{X}_a, \mathbf{X}_e; \mathbf{d})$ as

$$R(\mathbf{X}_a, \mathbf{X}_e; \mathbf{d}) = f(r|\overline{\mathbf{x}}) = \frac{1}{\text{beta}(\alpha,\beta)} r^{a-1}(1-r)^{\beta-1}$$

$$\text{where} \quad \alpha = 1 + x, \quad \beta = N - x + 1, \quad \overline{\mathbf{x}} = \{\mathbf{x}_{e,1}, \ldots, \mathbf{x}_{e,N}\}$$

$$x = \sum R_k, \quad \text{and} \quad R_k = \Pr[g(X_a) \leq 0 | \mathbf{x}_{e,k}] \quad (12)$$

N is the number of finite data sets for epistemic uncertainties.

For design optimization, Bayesian reliability must satisfy two requirements: (a) sufficiency and (b) uniqueness. The sufficiency requirement means that the Bayesian reliability must be no larger than an exact reliability, when it is realized with a sufficient amount of data for input uncertainties. The uniqueness requirement means that the Bayesian reliability must be uniquely defined for the purpose of design optimization. To meet these two requirements, Bayesian reliability is generally defined with a confidence level of reliability prediction where the confidence level C_L of Bayesian reliability is defined as

$$C_L = \Pr(R > R_B) = \int_{R_B}^{1} f(r|\overline{\mathbf{x}}) dr = 1 - F_R(R_B) \quad (13)$$

The confidence level can be defined by system reliability and design engineers. With the predefined confidence level C_L, Bayesian reliability can be defined as

$$R_B = F_R^{-1}[1 - C_L] \quad (14)$$

From Eq. (14) it is clear that Bayesian reliability is based on the reliability distribution obtained from Eq. (12) with certain confidence level C_L. The reliability distribution might be different if different samples are used, and this is the fundamental characteristic of uncertainty modeling when the data employed for reliability analysis are only in a small size. Bayesian reliability is desirable since it is defined from the reliability distribution with a corresponding confidence level and accounts for reliability modeling error due to the lack of data. Moreover, Bayesian reliability can be updated as more data are acquired.

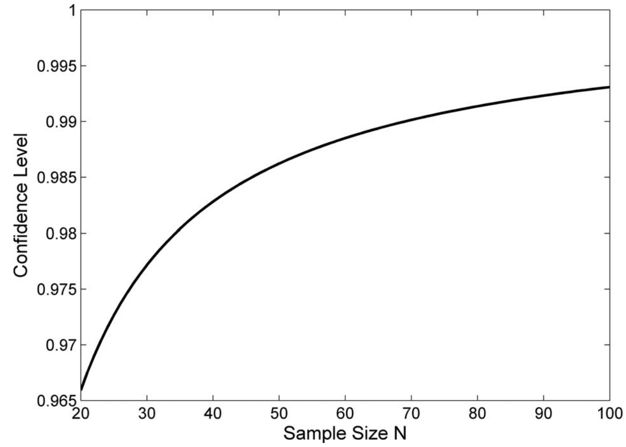

Fig. 4 Confidence level versus sample size for Bayesian reliability

In this study, extreme distribution theory for the smallest reliability value is employed to guarantee the sufficiency requirement. Then Bayesian reliability is defined as the median value of this extreme distribution obtained from the Beta distribution in Eq. (12) based on the extreme distribution theory. The remaining part of this subsection derives the Bayesian reliability formulation and its corresponding confidence level.

Based on the extreme distribution theory, the extreme distribution for the smallest reliability value is constructed from the reliability distribution, beta distribution. For random reliability R, which follows the beta distribution, $F_R(r)$, let ${}^1R$ be the smallest value among N data points for R, the CDF of the smallest reliability value, ${}^1R$, can be expressed as [49]

$$1 - F_{^1R}(r) = P(^1R > r) = P(^1R > r, ^2R > r, \ldots, ^NR > r) \quad (15)$$

Since the ith smallest reliability values, ${}^iR(i=1,\ldots,N)$, are identically distributed and statistically independent, the CDF of the smallest reliability value becomes

$$F_{^1R}(r) = 1 - [1 - F_R(r)]^N \quad (16)$$

Bayesian reliability, R_B, is defined as the median value of the extreme distribution shown in Eq. (16). Based on this definition, the formulation of Bayesian reliability can be obtained as the solution of the nonlinear equation, Eq. (16), by setting its left hand side to 0.5, which is

$$R_B = F_R^{-1}[1 - \sqrt[N]{1 - F_{^1R}(r^m)}] = F_R^{-1}[1 - \sqrt[N]{0.5}] \quad (17)$$

Based on this definition, the confidence level of the Bayesian reliability in Eq. (17) can be calculated as

$$C_L^B = 1 - F_R(R_B) = 1 - F_R(F_R^{-1}[1 - \sqrt[N]{0.5}]) = \sqrt[N]{0.5} \quad (18)$$

As shown in the above equation, the confidence level of Bayesian reliability is a function of the sample size (N) of epistemic uncertainties. Figure 4 shows the relationship between the sample size (N) and confidence level of the Bayesian reliability.

Based on the discussion above, Bayesian reliability analysis can be conducted using a numerical procedure as follows.

- Step 1: collect a limited data set for epistemic uncertainties where the data size is N.
- Step 2: calculate reliabilities (R_k) with consideration of aleatory uncertainties at all epistemic data points.

- Step 3: build a distribution of reliability using the beta distribution in Eq. (12) with aleatory and/or epistemic uncertainties.
- Step 4: select an appropriate confidence level, C_L, of Bayesian reliability.
- Step 5: determine the Bayesian reliability using Eq. (14).

In this paper, following the general procedure above, the confidence level of Bayesian reliability in Step 4 is selected based on Eq. (18), and the corresponding Bayesian reliability can then be calculated directly by Eq. (17).

2.4 Bayesian Design Toolkit. The BIRD software is developed by incorporating both BIT and BRT with Bayesian RBDO developed by the authors [22]. Although this paper is not focused on Bayesian RBDO, BDT will be reviewed as one of BIRD modules. Knowing that both aleatory and epistemic uncertainties exist in a system of interest, Bayesian RBDO can be formulated as

$$\text{minimize } C(\mathbf{X}_a,\mathbf{X}_e;\mathbf{d})$$

$$\text{subject to } P_B(G_i(\mathbf{X}_a,\mathbf{X}_e;\mathbf{d}) \le 0) \ge \Phi(\beta_{t_i})$$

$$i=1,\ldots,np; \quad \mathbf{X}_a \in R^{na}, \quad \mathbf{X}_e \in R^{ne}$$

$$\mathbf{d}^{\mathbf{L}} \le \mathbf{d} \le \mathbf{d}^{\mathbf{U}}, \quad \mathbf{d} \in R^{nd} \tag{19}$$

where $P_B(G_i(\mathbf{X}_a,\mathbf{X}_e;\mathbf{d})\le 0)=R_i^B(\mathbf{X}_a,\mathbf{X}_e;\mathbf{d})$ is the Bayesian reliability, and $G_i(\mathbf{X}_a,\mathbf{X}_e;\mathbf{d})\le 0$ is defined as a safety event. $\boldsymbol{C}(\mathbf{X}_a,\mathbf{X}_e;\mathbf{d})$ is the objective function; $\mathbf{d}=\boldsymbol{\mu}(\mathbf{X})$ is the design vector; $\mathbf{X}$ is the random vector; β_t is the prescribed reliability target; and np, nd, na, and ne are the numbers of probabilistic constraints, design variables, aleatory random variables, and epistemic random variables, respectively. If the parameters describing a random variable are controllable among all (both aleatory and epistemic) random variables, they are considered design variables. For instance, a random variable with a normal distribution may have two design variables, mean and standard deviation. It will be shown that the result from Bayesian RBDO asymptotically approaches that from the conventional (or frequentist) RBDO. In other words, frequentist RBDO is a special case of Bayesian RBDO because Bayesian RBDO is able to handle aleatory and/or epistemic uncertainties.

3 Case Studies

Two examples are employed to demonstrate the feasibility of Bayesian reliability analysis with evolving, insufficient, and subjective data sets: (1) a mathematical example and (2) door closing effort in the vehicle door system.

3.1 A Mathematical Example. Let $G(X_1,X_2)=3-X_1^2X_2/20 \le G_0$ be an inequality constraint with two random variables, where X_1 is an epistemic random variable and X_2 is an aleatory random variable, $X_2 \sim N(2.8,0.2^2)$. Besides, G_0 is a random parameter, which follows a normal distribution $N(2.0,0.05^2)$ and represents the uncertainty of the target performance. In this mathematical example, the distribution for G_0 is known; however, in most practical cases, this distribution should be determined by observation.

Twenty data values are randomly sampled for X_1 from an assumed normal distribution $N(2.9,0.2^2)$, as shown in Table 2. The table also shows the corresponding reliabilities $R_k=\Pr[G(X_2) \le G_0 | X_1(k)]$ for $k=1,\ldots,20$ that are computed from reliability analyses. For example, $X_1(1)=2.9277$, then $R_1=P(3-2.9277^2 * X_2/20 \le G_0)=0.97807$. Figure 5 shows the PDFs of the performance function $G(X_1=2.9277,\ X_2)$ and G_0. By carrying out the probability analysis for all 20 epistemic data, 20 probability values are then obtained, as shown in Table 2. From Table 2, the expected number of safe design points out of the 20 designs can be obtained from the sum of all 20 reliabilities, $x=\Sigma R_k = 17.4408$. As discussed in Sec. 2, the reliability can then be modeled with the beta distribution as beta(18.4408, 3.5592) at the design point, (2.9, 2.8). This is graphically shown in Fig. 6.

To validate the results, Monte Carlo simulation (MCS) (10,000 random samples) is conducted by assuming X_1 to follow $N(2.9,0.2^2)$. It gives the true reliability (0.8345) of the design point. Figure 6 shows that the reliability distribution gives a quite feasible estimate of the true reliability with both aleatory and epistemic uncertainties. In this example, a uniform distribution, $r \sim U(0,1)$, is used as the prior distribution of reliability. There-

Table 2 X_1 samples and probabilities

X_1	Probability	X_1	Probability
2.9277	0.97807	3.4741	1.00000
2.7605	0.76836	2.9575	0.98709
2.775	0.80247	2.9029	0.96671
3.1006	0.99929	2.9430	0.98323
2.8175	0.88239	2.8196	0.88559
2.5933	0.24267	2.9706	0.98986
3.1047	0.99936	2.7157	0.64237
2.9604	0.98775	2.6738	0.50406
3.1706	0.99986	2.8869	0.95693
2.9354	0.98082	2.8185	0.88392

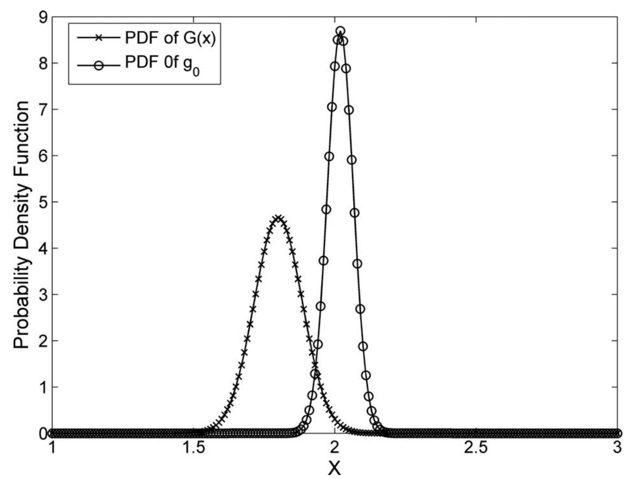

Fig. 5 PDFs for $G(X_1,X_2)$ and G_0

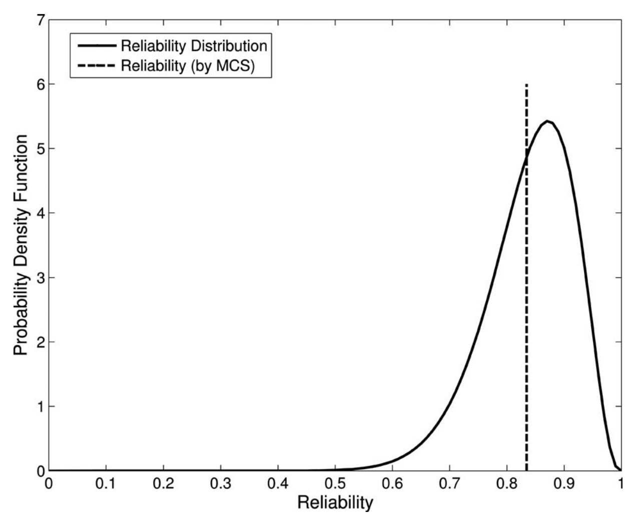

Fig. 6 Actual reliability and estimated reliability distribution

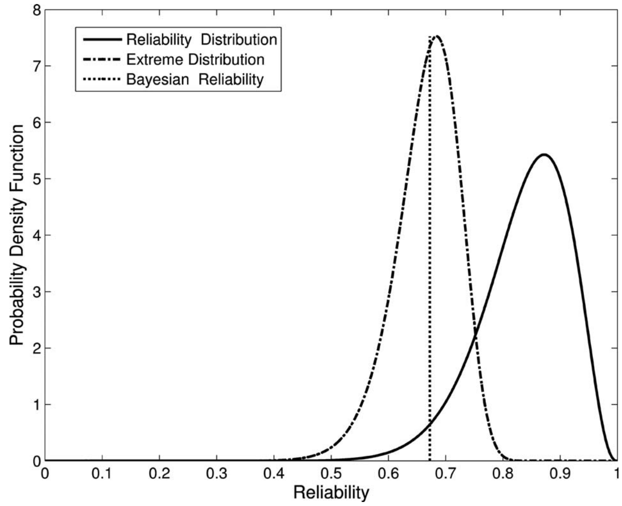

Fig. 7 Bayesian reliability for the mathematical example

fore, the reliability distribution appears to be widely distributed, but it becomes narrower if the prior distribution is more precisely given.

Using Eq. (20), the extreme distribution for the smallest reliability value is obtained as

$$F_{^1R}(r) = \Pr(^1R \le r) = 1 - \left[1 - \int_0^r \frac{1}{B(18.4408, 3.5592)} \theta^{17.4408}(1 - \theta)^{2.5592} d\theta\right]^{20}$$

From Eq. (17), Bayesian reliability is calculated as $P_B = 0.6725$. The beta distribution for reliability, its extreme distribution for the smallest reliability value, and the Bayesian reliability are shown in Fig. 7.

3.2 Bayesian Reliability Analysis for a Vehicle Door System. The demonstration problem used in this study is the body-door system of a passenger vehicle, as illustrated in Fig. 8. The vehicle door system is of special concern due to its frequency of use and its engineering challenge with respect to design, assembly, and operation. Variation exists in the compression load deflection (CLD) response of the seal, the gap between the body and door, as well as in attaching the door to the car body. Besides the presence of variation, the complexity of the system is high due to the nonlinear seal behavior and the dynamics of door closing. The detail of vehicle door system regarding the problem description, failure mechanism specification, physical model creation, and response surface construction can be found in Ref. [50]. The performance measure selected in this study to assess one aspect of door

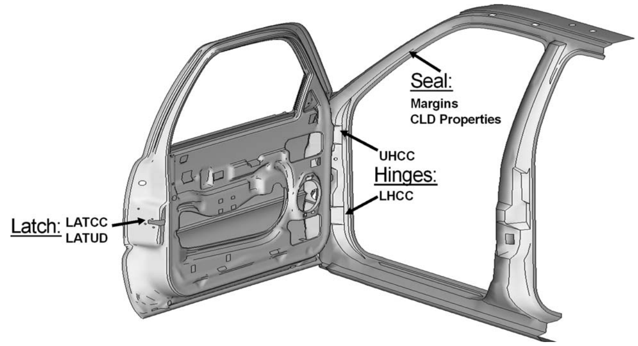

Fig. 8 Vehicle door system

system design is the door closing effort. The measurable quantity for this performance measure is the door closing velocity. A response surface for door closing velocity was created based on results from physics-based models, and the performance evaluation criteria were deduced from both expert opinions and voice of the customer information.

For the door system example in this study, 26 random input variables are used to specify the uncertainty of the system. Within these 26 random input variables, listed in Table 3, X_5, X_6, X_7, X_{25}, and X_{26} are aleatory variables, which, for this example, are assigned uniform distributions on different threshold values, as shown in the table. Except for these 5 random input variables, all others are epistemic variables with a total of 79 sets of measurement data. For illustrative purpose, these epistemic data are partially listed in Table 4.

In Secs. 3.2.1 and 3.2.2, we describe the modeling of the performance evaluation criteria, i.e., the marginal velocity, using the Bayesian updating technique introduced in Sec. 2 followed by the Bayesian reliability analysis carried out for the door closure problem.

3.2.1 Modeling of the Marginal Velocity. In this subsection, the marginal velocity, which serves as the criterion of the door performance evaluation, is modeled by using the Bayesian updating technique based on expert opinion and the customer data. From a hypothetical expert, the door closing velocity values for customer satisfaction should be, for example, within the range of 0 m/s to $v_{\max}$ m/s. Customer survey regarding the door closing velocity can be carried out by using the direct customer survey method [11], and illustrative results, which show the customer rejection rate (CRR) versus the door closing velocity (normalized by $v_{\max}$) are graphically shown in Fig. 9. For the modeling of the marginal velocity, CRR can be treated as the probability of the marginal velocity being smaller than a given a or $CRR = P(v_m \le a)$, where v_m is a random marginal velocity and a is within $[0, v_{\max}]$ based on expert opinion.

The procedure of marginal velocity modeling can be briefly

Table 3 Random variables and descriptions for the vehicle door system

Variable name	Description	Variable type
X_1	UHCC—upper hinge location in cross-car direction	Epistemic
X_2	LHCC—lower hinge location in cross-car direction	Epistemic
X_3	LATCC—latch location in cross-car direction	Epistemic
X_4	LATUD—latch location in up-down direction	Epistemic
X_5	Primary seal CLD property factor	$U(0.7, 1.3)$
X_6	Auxiliary seal CLD property factor	$U(0.7, 1.3)$
X_7	Cutline seal CLD property factor	$U(0.7, 1.3)$
X_8–X_{24}	Primary seal margin regions $1 \sim 17$	Epistemic
X_{25}	Auxiliary seal margin	$U(-1, 1)$
X_{26}	Cutline seal margin	$U(-1, 1)$

Table 4 Data of epistemic random variables for the vehicle door system

Variables	Data							
	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	…	Set 79
X_1	1.62	2.29	1.58	1.58	1.19	1.70	…	2.16
X_2	2.82	2.49	1.80	2.10	2.03	1.37	…	1.36
X_3	2.56	2.10	1.82	1.67	1.75	1.01	…	1.35
X_4	−0.38	−0.35	−0.01	−0.01	0.61	−0.44	…	−0.61
X_8	1.66	1.24	1.02	0.72	0.71	−0.06	…	0.56
X_9	1.078	0.77	0.59	0.28	0.12	−0.31	…	0.40
X_{10}	0.50	0.31	0.17	−0.15	−0.48	−0.57	…	0.24
X_{11}	1.24	0.74	0.43	0.11	−0.23	−0.02	…	0.96
X_{12}	−0.27	−0.31	−0.28	−0.66	−1.29	−0.09	…	0.28
X_{13}	0.03	0.16	−0.205	−0.29	−1.02	−0.31	…	0.11
X_{14}	0.33	0.63	−0.13	0.08	−0.75	−0.53	…	−0.05
X_{15}	0.50	0.79	0.06	0.22	−0.76	−0.35	…	0.14
X_{16}	0.89	1.01	0.87	0.27	−0.63	0.02	…	0.24
X_{17}	0.27	0.51	−0.01	−0.21	−1.57	−0.19	…	0.23
X_{18}	−0.35	0.01	−0.89	−0.69	−2.5	−0.39	…	0.23
X_{19}	−0.35	0.01	−0.89	−0.69	−2.5	−0.39	…	0.23
X_{20}	−0.44	−0.53	−1.27	−1.55	−2.93	−0.77	…	−0.38
X_{21}	−0.44	−0.53	−1.27	−1.55	−2.93	−0.77	…	−0.38
X_{22}	0.16	−0.03	−0.71	−0.86	−1.68	−0.17	…	0.12
X_{23}	0.76	0.47	−0.16	−0.18	−0.44	0.42	…	0.62
X_{24}	1.49	0.91	0.56	0.27	0.91	0.08	…	0.54

summarized into three steps. First, based on the customer data, one Bayesian inference model should be specified. For example, if the Bayesian normal inference model is used, the marginal velocity will be modeled as the mean value of the normal distribution, which is the conjugate distribution for this model. Second, based on the selected model, the CDF analysis can be carried out for the CDF/velocity data. After completing this analysis, the CDF data are then transferred to parameter data for the distribution. Third, with one prior distribution assumed, Bayesian updating can then be carried out with sets of parameter data.

In this study, the Bayesian normal inference model will be used, and the marginal velocity will be modeled as the mean value of a normal distribution. As introduced in Sec. 2 of this paper, we suppose that the marginal velocity also follows a normal distribution, which is the conjugate distribution of the normal inference model. Expert opinion is used in modeling the prior information on the marginal velocity. To properly model the normal distribution with the information from the expert, the six sigma region of the normal distribution is set to the interval, such that $[\mu-3\sigma=0,\ \mu+3\sigma=v_{\max}]$. Although the domain of the normal distribution is $[-\infty, +\infty]$, the contribution out of the bound $[0, v_{\max}]$ is negligible. Normalizing by $v_{\max}$, the distribution $N(0.5, 0.1667)$ is used for the prior distribution of this model.

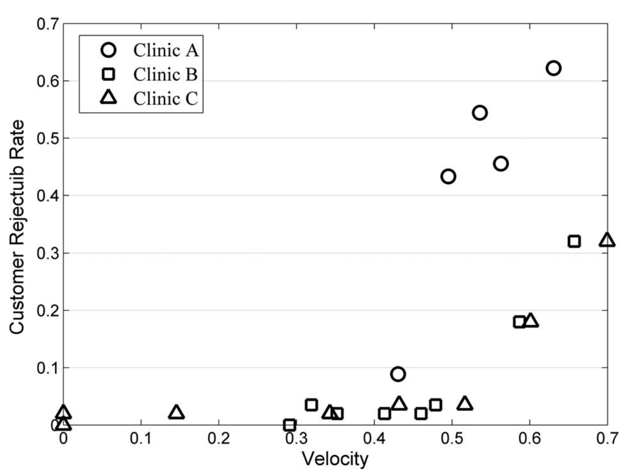

Fig. 9 Customer rejection rate

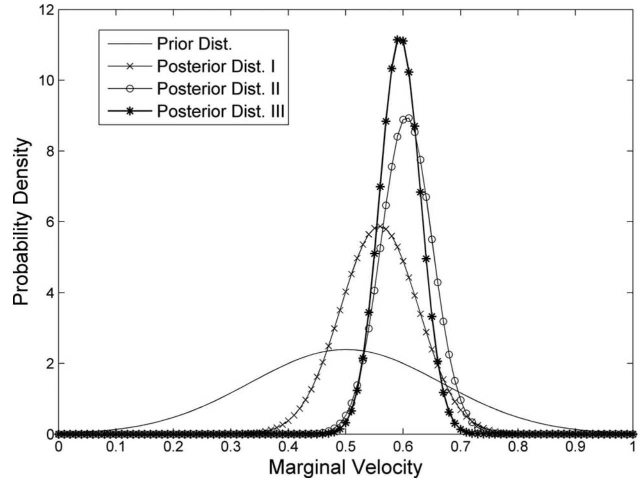

Fig. 10 Bayesian updating for the marginal velocity using a normal distribution

As an example to show how the CDF analysis is carried out, we use a set of data, e.g., normalized velocity is 0.43 and customer satisfaction rate is 91.8%, from clinic A's customer data shown in Fig. 9. The 8.2% customer rejection rate will be considered as the CDF value corresponding to the velocity value 0.43. As we suppose $\sigma=0.1667$, then based on the CDF data $Z_{0.082}=0.43$, we can determine that the mean value of the normal distribution is 0.662. For each set of customer data, the corresponding parameter data are determined by the CDF analysis. Three sets of parameter data are then obtained from the three sets of customer data after the CDF analysis. The Bayesian normal inference can be expressed as

$$\mu_1 = \left(\frac{\mu_0}{\sigma_0^2} + \frac{\sum_{i=1}^{N} X_i}{\sigma^2}\right) \Big/ \left(\frac{1}{\sigma_0^2} + \frac{N}{\sigma^2}\right)$$

$$\sigma_1^2 = \left(\frac{1}{\sigma_0^2} + \frac{N}{\sigma^2}\right)^{-1}$$

where μ_1 and σ_1 are parameters for the posterior distribution, whereas μ_0 and σ_0 are parameters for the prior distributions, X_i is the ith parameter data, and σ is the population variance. Based on the Bayesian normal inference, the PDFs for the marginal velocity can then be gradually refined by aggregating three different clinic data sets with the normal prior distribution, shown in Fig. 10. With the clinic-A data set, the first Bayesian model for the marginal velocity is the posterior distribution I, $N(0.559, 0.068^2)$, shown in Fig. 10. Then this posterior distribution is treated as the prior distribution and combined with the clinic-B data set, to obtain the second Bayesian model, posterior distribution II, $N(0.606, 0.0445^2)$. Similarly with the clinic-C data set, the final Bayesian model is obtained as the posterior distribution III, $N(0.5946, 0.0355^2)$, as shown in Fig. 10. Figure 11 shows the PDF and CDF of the final Bayesian model, $N(0.5946, 0.0355^2)$, for the marginal velocity.

3.2.2 Bayesian Reliability Analyses for a Vehicle Door System. Based on the marginal velocity PDF created, Bayesian reliability analysis is then carried out for the door closing effort problem with both aleatory and epistemic uncertainties. For a given set of input values, the performance response can be obtained from the response surface created based on the physical model [50]. Since Bayesian reliability analysis requires the probabilistic performance evaluation for each set of epistemic data, two

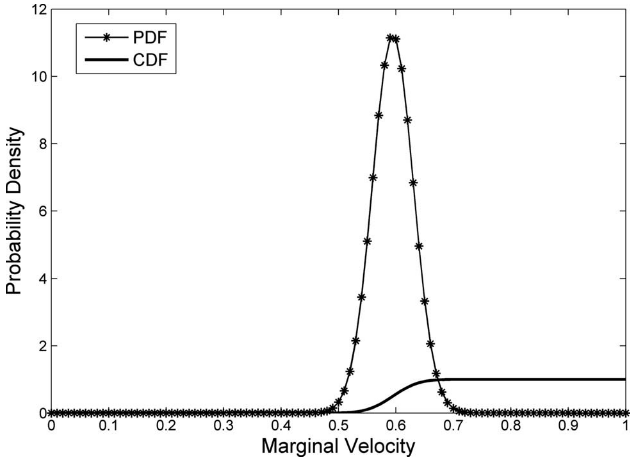

Fig. 11 Bayesian model for the marginal velocity using a normal distribution

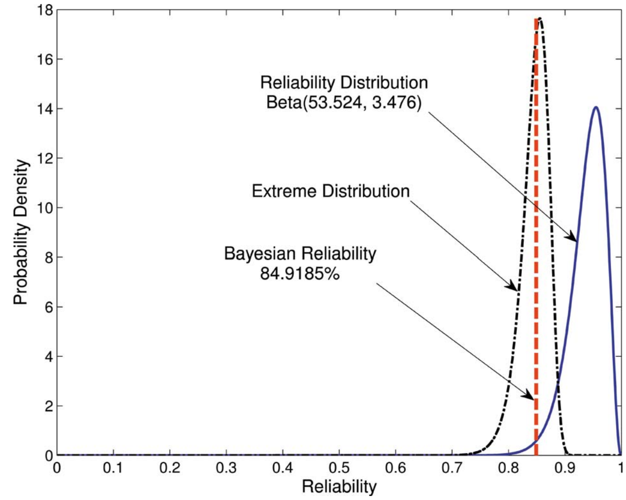

Fig. 12 Bayesian reliability with 55 sets data (by MCS)

different approaches, Monte Carlo simulation and eigenvector dimension reduction (EDR) method [35], are employed in this study to calculate the reliability for each set of epistemic data. EDR method is an efficient and accurate sensitivity free method for reliability analysis. Results for the door closing effort problem in this study from MCS and EDR are compared.

First, for each set of epistemic data, direct MCS is used to carry out the reliability analysis. For each aleatory variable (including the variable of marginal velocity), 10,000 random samples are generated and used for MCS. Table 5 shows the 55 reliabilities corresponding to the first 55 sets of epistemic data. Based on Table 5, we carried out the Bayesian reliability analysis and obtained the reliability distribution as beta (53.524, 3.476). Then by the Bayesian reliability definition described in Sec. 2.3, the extreme distribution of the smallest value for the beta distribution is constructed, and the Bayesian reliability is realized as 0.849185. Figure 12 shows the beta distribution, extreme distribution, and the Bayesian reliability value using MCS. With 24 new data sets involved for the epistemic random variables the Bayesian reliability is updated. The updated reliability distribution is beta (77.1869, 3.8131), and the Bayesian reliability is updated from the original 0.849185 to 0.880935. Table 6 shows the reliabilities corresponding to each set of the new involved data. Figure 13 shows the updated beta distribution, extreme distribution, and the Bayesian reliability using MCS.

As we can see from the Monte Carlo simulation method, the reliability analysis for each set of epistemic data can require a large amount of response performance evaluations depending on the simulation sample size (in this case 10,000). In order to make the calculation of the Bayesian reliability more efficient, the EDR method is used for the probability calculation for each set of epistemic data. By using EDR method, the total number of the response performance evaluation is reduced from 10,000 to $2n+1=13$. Based on the marginal velocity PDF created in Sec. 3.2.1, the reliability R_i of a certain design (Xa, Xe^i) can be formulated as $R_i=\Pr[V(Xa,Xe^i)-Vt\leq 0]$, where $V(Xa,Xe^i)$ is the performance velocity variable corresponding to a certain design (Xa, Xe^i), Xa is the aleatory variable set and Xe^i is the ith set of epistemic data, and Vt is the marginal velocity. Totally 55 different reliabilities corresponding to 55 different sets of epistemic uncertainties are realized, as shown in Table 7. Based on these results, the reliabil-

Table 5 55 reliabilities corresponding to 55 epistemic data sets (by MCS)

Rel. (1∼11)	Rel. (12∼22)	Rel. (23∼33)	Rel. (34∼44)	Rel. (45∼55)
0.9973	1.0000	0.9987	0.9995	0.9988
1.0000	0.9993	0.9970	0.9998	0.2703
0.9993	1.0000	1.0000	0.9999	0.9987
0.9945	1.0000	0.9951	0.9999	1.0000
0.8265	1.0000	0.9970	0.9974	0.9955
0.9996	0.9999	0.9899	0.9977	0.9937
0.9985	0.9991	0.9998	0.9918	0.9918
1.0000	0.9999	1.0000	0.9007	1.0000
1.0000	0.9993	0.9993	0.9976	0.9994
1.0000	1.0000	1.0000	0.9778	0.2109
1.0000	0.9999	0.9963	0.9730	0.4436

Table 6 24 reliabilities corresponding to 24 new data sets (by MCS)

Rel. (1∼6)	Rel. (7∼12)	Rel. (13∼18)	Rel. (19∼24)
0.9929	0.9996	0.8864	1.0000
0.9999	0.9999	1.0000	0.8973
0.9995	0.9996	0.9963	0.9842
0.9993	0.9989	0.9240	0.9866
0.9993	1.0000	1.0000	0.9998
0.9994	1.0000	1.0000	1.0000

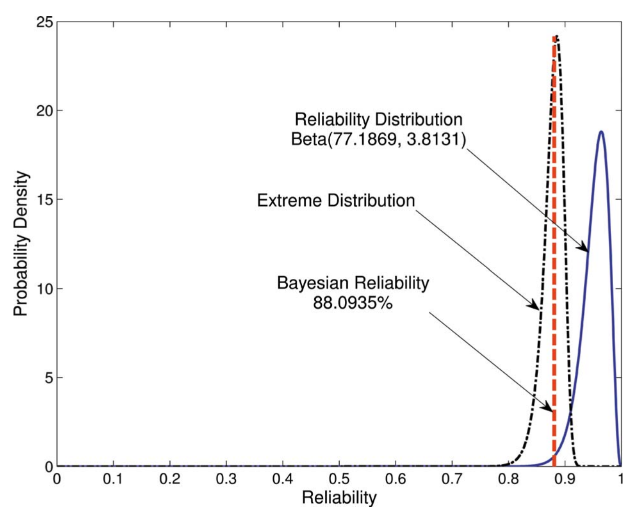

Fig. 13 Updated Bayesian reliability with 24 new data sets (by MCS)

Table 7 55 reliabilities corresponding to 55 epistemic data sets (by EDR)

Rel. (1∼11)	Rel. (12∼22)	Rel. (23∼33)	Rel. (34∼44)	Rel. (45∼55)
0.9978	1.0000	0.9991	0.9998	0.9993
1.0000	0.9997	0.9976	0.9998	0.2642
0.9996	1.0000	1.0000	1.0000	0.9992
0.9953	1.0000	0.9963	1.0000	1.0000
0.8243	1.0000	0.9977	0.9982	0.9963
0.9998	0.9999	0.9893	0.9984	0.9944
0.9991	0.9995	0.9998	0.9917	0.9915
1.0000	1.0000	1.0000	0.8938	1.0000
1.0000	0.9996	0.9996	0.9984	0.9997
1.0000	1.0000	1.0000	0.9755	0.2070
1.0000	0.9999	0.9971	0.9702	0.4394

ity distribution is obtained as beta (53.5076, 3.4924) from the Bayesian inference. Then by the Bayesian reliability definition, the extreme distribution of smallest value for the beta distribution is constructed, and the Bayesian reliability is realized as 0.848752. Figure 14 shows the beta distribution, extreme distribution, and the Bayesian reliability using the EDR method. With 24 new data sets involved, the Bayesian reliability is updated. The updated reliability distribution is beta (77.1567, 3.8433), and the Bayesian reliability is updated from the original 0.848752 to 0.880363. Table 8 shows the reliabilities corresponding to each set of the new involved data. Figure 15 shows the updated beta distribution, extreme distribution, and the Bayesian reliability using the EDR method.

A comparison of the results from using the two different probability analysis approaches shows that the EDR method maintains good accuracy and at the same time provides a higher computational efficiency compared with MCS. From the analysis results

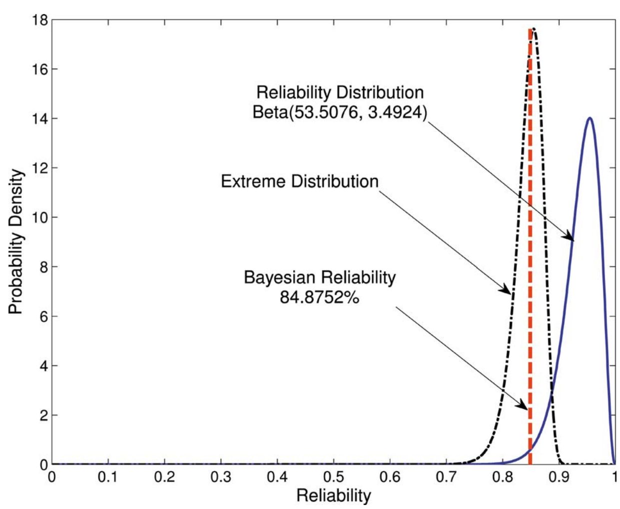

Fig. 14 Bayesian reliability with 55 data sets (by EDR)

Table 8 24 reliabilities corresponding to 24 new data sets (by EDR)

Rel. (1∼6)	Rel. (7∼12)	Rel. (13∼18)	Rel. (19∼24)
0.9928	0.9998	0.8814	1.0000
0.9999	0.9999	1.0000	0.8915
0.9997	0.9997	0.9969	0.9835
0.9996	0.9993	0.919	0.9867
0.9996	1.0000	1.0000	0.9999
0.9997	1.0000	1.0000	1.0000

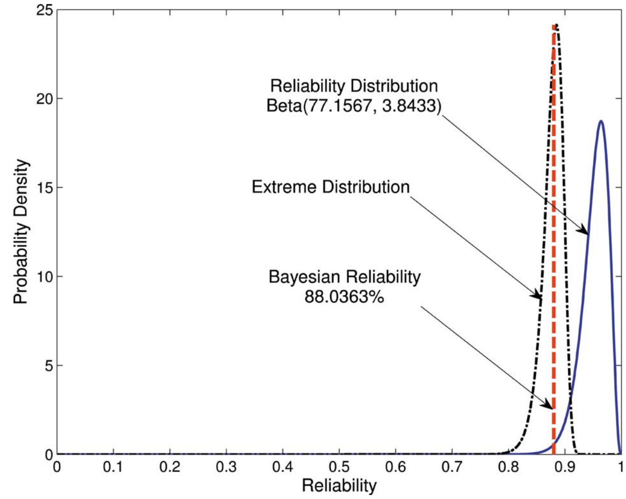

Fig. 15 Updated Bayesian reliability with 24 new data sets (by EDR)

obtained with both MCS and the EDR method, two points are clear: first, Bayesian reliability increases with the increase in the reliability value corresponding to each set of epistemic data; second, the updated Bayesian reliability increases with the addition of more epistemic data into the Bayesian reliability analysis. This is because the Bayesian reliability represents not only the design uncertainty of the system but also the uncertainty due to the limiting information represented by the epistemic uncertainties. As more data are involved, a better understanding of the characteristic of epistemic uncertainties can be expected and consequently a higher Bayesian reliability can be realized. Also, the Bayesian reliability analysis approach proposed in this paper offers a convenient and effective method for the performance evaluation of the problems involving several different types of uncertainty and where uncertainty data are continuously collected.

4 Conclusion

This research presented a new paradigm of reliability prediction that enables the use of evolving, insufficient, and subjective data sets. Bayes' theory was employed to deal with the evolving and insufficient data sets for data and reliability analyses. Moreover, the integration of the probability encoding methods (*P*-methods, *V*-methods, and *PV*-methods) to the paradigm enabled the use of subjective data sets for data and reliability analyses. The tools for data and reliability analyses are referred to as BIT and BRT, respectively. The generic definition of Bayesian reliability was presented, which is a function of a predefined confidence level. The confidence level can be defined as a function of the available sample size of epistemic uncertainties by reliability analysts or decision makers. The effect of the sequence of evolving data sets on Bayesian updating was carefully studied and found that there is no data-sequence effect on the updating results. Integrating with the authors' earlier work in Bayesian RBDO, this paper presents BIRD software. It was shown that the proposed Bayesian reliability analysis can predict the reliability of the door closing performance in the vehicle body-door subsystem where available data sets are insufficient, subjective, and evolving.

Acknowledgment

This research is partially supported by the U.S. National Science Foundation (NSF) under Grant No. GOALI-0729424, by the U.S. Army TARDEC under STAS Contract No. TCN-05122, and by General Motors under Grant No. TCS02723. The authors are grateful to Dr. John Cafeo from GM R&D Center for reviewing the paper and providing valuable comments.

References

[1] Roy, R., Hinduja, S., and Teti, R., 2008, "Recent Advances in Engineering Design Optimization: Challenges and Future Trends," CIRP Ann., **57**(2), pp. 697–715.

[2] Liu, S., and Boyle, I. M., 2009, "Engineering Design: Perspectives, Challenges, and Recent Advances," J. Eng. Design, **20**(1), pp. 7–19.

[3] Gunawan, S., and Papalambros P. Y., 2006, "A Bayesian Approach to Reliability-Based Optimization With Incomplete Information," ASME J. Mech. Des., **128**(4), pp. 909–918.

[4] Luo, L., Kannan, P. K., and Ratchford, B. T., 2008, "Incorporating Subjective Characteristics in Product Design and Evaluation," J. Mark. Res., **45**(2), pp. 182–194.

[5] Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., and Ferson, S., 2004, "Challenge Problems: Uncertainty in System Response Given Uncertain Parameters," Reliab. Eng. Syst. Saf., **85**(1–3), pp. 11–19.

[6] Mohammadi, J., Longinow, A., and Williams, T. A., 1991, "Evaluation of System Reliability Using Expert Opinions," *Structural Safety*, Elsevier, New York, Vol. 9, pp. 227–241.

[7] Clemen, R. T., and Winkler, R. L., 1999, "Combining Probability Distributions From Experts in Risk Analysis," Risk Anal., **46**(8), pp. 1100–1115.

[8] Ayyub, B. M., 2000, "Methods for Expert Opinion Elicitation of Probabilities and Consequences for Corps Facilities," U.S. Army Corps of Engineers, IWR Report No. 00-R-10.

[9] Ouchi, F., 2004, "A Literature Review on the Use of Expert Opinion in Probabilistic Risk Analysis," World Bank Policy Research, Working Paper No. 3201.

[10] Kumar, U. D., and Hinds, P., 2005, "Aggregation of Expert Opinion: Applications to Reliability Prediction and Spare Parts Provisioning," Proceedings of the CSER, Hoboken, NJ, Mar. 23–25.

[11] Spetzler, C. S., Von Holstein, S., and Carl-Axel, S., 1975, "Probability Encoding in Decision Analysis," Manage. Sci., **22**(3), pp. 340–358.

[12] Utkin, L. V., and Gurov, S. V., 1996, "A General Formal Approach for Fuzzy Reliability Analysis in the Possibility Context," Fuzzy Sets Syst., **83**, pp. 203–213.

[13] Bai, X. G., and Asgarpoor, S., 2004, "Fuzzy-Based Approaches to Substation Reliability Evaluation," Electr. Power Syst. Res., **69**(2–3), pp. 197–204.

[14] Du, L., Choi, K. K., Youn, B. D., and Gorsich, D., 2006, "Possibility-Based Design Optimization Method for Design Problems With Both Statistical and Fuzzy Input Data," ASME J. Mech. Des., **128**(4), pp. 928–935.

[15] Zhou, J., and Mourelatos, Z. P., 2008, "A Sequential Algorithm for Possibility-Based Design Optimization," ASME J. Mech. Des., **130**(1), p. 011001.

[16] Youn, B. D., Choi, K. K., and Gorsich, D., 2008, "Integration of Possibility-Based Optimization to Robust Design for Epistemic Uncertainty," ASME J. Mech. Des., **129**(8), pp. 876–882.

[17] Sentz, K., and Ferson, S., 2002, "Combination of Evidence in Dempster-Shafer Theory," Sandia National Laboratories, Apr., Technical Report No. SAND2002-0835.

[18] Bae, H.-R., Grandhi, R. V., and Canfield, R. A., 2003, "Uncertainty Quantification of Structural Response Using Evidence Theory," AIAA J., **41**(10), pp. 2062–2068.

[19] Helton, J. C., Johnson, J. D., Oberkampf, W. L., and Sallaberry, C. J., 2006, "Sensitivity Analysis in Conjunction With Evidence Theory Representations of Epistemic Uncertainty," Reliab. Eng. Syst. Saf., **91**(10–11), pp. 1414–1434.

[20] Coolen, F. P. A., and Newby, M. J., 1994, "Bayesian reliability Analysis With Imprecise Prior Probabilities," Reliab. Eng. Syst. Saf., **43**(1), pp. 75–85.

[21] Huang, H. Z., Zuo, M. J., and Sun, Z. Q., 2006, "Bayesian Reliability Analysis for Fuzzy Lifetime Data," Fuzzy Sets Syst., **157**(12), pp. 1674–1686.

[22] Youn, B. D., and Wang, P. F., 2008, "Bayesian Reliability-Based Design Optimization Using Eigenvector Dimension Reduction Method," Struct. Multidiscip. Optim., **36**(2), pp. 107–123.

[23] McDonald, M., and Mahadevan, S., 2008, "Reliability-Based Optimization With Discrete and Continuous Decision and Random Variables," ASME J. Mech. Des., **130**(6), p. 061401.

[24] Youn, B. D., Choi, K. K., and Park, Y. H., 2003, "Hybrid Analysis Method for Reliability-Based Design Optimization," ASME J. Mech. Des., **125**(2), pp. 221–232.

[25] Youn, B. D., and Choi, K. K., 2004, "Selecting Probabilistic Approaches for Reliability-Based Design Optimization," AIAA J., **42**(1), pp. 124–131.

[26] Kim, C., and Choi, K. K., 2008, "Reliability-Based Design Optimization Using Response Surface Method With Prediction Interval Estimation," ASME J. Mech. Des., **130**(12), p. 121401.

[27] Du, L., Choi, K. K., and Youn B. D., 2006, "Inverse Possibility Analysis Method for Possibility-Based Design Optimization," AIAA J., **44**(11), pp. 2682–2690.

[28] Du, L., and Choi, K. K., 2008, "An Inverse Analysis Method for Design Optimization With Both Statistical and Fuzzy Uncertainties," Struct. Multidiscip. Optim., **37**(2), pp. 107–119.

[29] Mourelatos, Z. P., and Zhou, J., 2006, "A Design Optimization Under Uncertainty Using Evidence Theory," ASME J. Mech. Des., **128**(4), pp. 901–908.

[30] Allen, J. K., and Choi, H. J., 2009, "A Meta-Modeling Approach for Uncertainty Analysis of Non-Deterministic Systems," ASME J. Mech. Des., **131**(5), p. 041008.

[31] Huang, H., and Zhang, X., 2009, "Design Optimization With Discrete and Continuous Variables of Aleatory and Epistemic Uncertainties," ASME J. Mech. Des., **131**(3), p. 031006.

[32] Schulz, V., and Schillings, C., 2009, "Problem Formulations and Treatment of Uncertainties in Aerodynamic Design," AIAA J., **47**(3), pp. 646–654.

[33] Youn, B. D., and Wang, P., 2009, "Complementary Intersection Method for System Reliability Analysis," ASME J. Mech. Des., **131**(4), p. 041004.

[34] Du, X., and Chen, W., 2004, "Sequential Optimization and Reliability Assessment for Probabilistic Design," ASME J. Mech. Des., **126**(2), pp. 225–233.

[35] Youn, B. D., Zhimin, X., and Wang, P. F., 2008, "Eigenvector Dimension-Reduction (EDR) Method for Sensitivity-Free Probability Analysis," Struct. Multidiscip. Optim., **37**(1), pp. 13–28.

[36] Fickas, E. T., Martz, H. F., and Waller, R. A., 1988, "Bayesian Reliability Analysis of Series Systems of Binomial Subsystems and Components," Technometrics, **30**(2), pp. 143–154.

[37] Yu, D. C., Nguyen, T. C., and Haddawy, P., 1999, "Bayesian Network Model for Reliability Assessment of Power Systems," IEEE Trans. Power Syst., **14**(2), pp. 426–432.

[38] Quigley, J., and Walls, L., 1999, "Measuring the Effectiveness of Reliability Growth Testing," Qual. Reliab. Eng. Int, **15**, pp. 87–93.

[39] Rajagopal, R., 2004, "Bayesian Methods for Robustness in Process Optimization," Ph.D. thesis, Industrial Engineering and Operations Research, Penn. State University.

[40] Chung, T. H., Mohamed, Y., and AbouRizk, S., 2004, "Simulation Input Updating Using Bayesian Techniques," Proceedings of the 2004 Winter Simulation Conference, pp. 1238–1243.

[41] Merl, D., Escalante, A., and Prado, R., 2005, "Comparison of Bayesian, Maximum Likelihood and Parsimony Methods for Detecting Positive Selection," Report No. AMS2005-03.

[42] Beck, J. L., and Au, S.-K., 2002, "Bayesian Updating of Structural Models and Reliability Using Markov Chain Monte Carlo Simulation," J. Eng. Mech., **128**, pp. 380–391.

[43] Weber, P., and Jouffe, L., 2006, "Complex System Reliability Modelling With Dynamic Object Oriented Bayesian Networks (DOOBN)," Reliab. Eng. Syst. Saf., **91**, pp. 149–162.

[44] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B., 2004, *Bayesian Data Analysis*, 2nd ed., Chapman & Hall/CRC, Chap. 2.

[45] Wallsten, T. S., and Budescu, D. V., 1983, "Encoding Subjective Probabilities: A Psychological and Psychometric Review," Manage. Sci., **29**(2), pp. 151–173.

[46] Winkler, R. L., 1967, "The Assessment of Prior Distributions in Bayesian Analysis," J. Am. Stat. Assoc., **62**, pp. 776–800.

[47] Haldar, A., and Mahadevan, S., 2000, *Probability, Reliability, and Statistical Methods in Engineering Design*, Wiley, New York, Chap. 4.

[48] Li, G. Y., Wu, Q.-G., and Zhao, Y. H., 2002, "On Bayesian Analysis of Binomial Reliability Growth," J. Japan Statist. Soc., **32**(1), pp. 1–14.

[49] Rao, S. S., 1992, *Reliability-Based Design*, McGraw-Hill, New York, Chap. 4.

[50] Kloess, A., Mourelatos, Z. P., and Meernik, P. R., 2004, "Probabilistic Analysis of an Automotive Body-Door System," Int. J. Veh. Des., **34**(2), pp. 101–125.