Struct Multidisc Optim
DOI 10.1007/s00158-016-1487-1

RESEARCH PAPER

Uncertainty quantification in reliability estimation with limit state surrogates

Saideep Nannapaneni[1] · **Zhen Hu**[1] · **Sankaran Mahadevan**[1]

Received: 7 December 2015 / Revised: 27 April 2016 / Accepted: 17 May 2016

Abstract Model-based reliability analysis is affected by different types of epistemic uncertainty, due to inadequate data and modeling errors. When the physics-based simulation model is computationally expensive, a surrogate has often been used in reliability analysis, introducing additional uncertainty due to the surrogate. This paper proposes a framework to include statistical uncertainty and model uncertainty in surrogate-based reliability analysis. Two types of surrogates have been considered: (1) general-purpose surrogate models that compute the system model output over the desired ranges of the random variables; and (2) limit-state surrogates. A unified approach to connect the model calibration analysis using the Kennedy and O'Hagan (KOH) framework to the construction of limit state surrogate and to estimating the uncertainty in reliability analysis is developed. The Gaussian Process (GP) general-purpose surrogate of the physics-based simulation model obtained from the KOH calibration analysis is further refined at the limit state (local refinement) to construct the limit state surrogate, which is used for reliability analysis. An efficient single-loop sampling approach using the probability integral transform is used for sampling the input variables with statistical uncertainty. The variability in the GP prediction (surrogate uncertainty) is included in reliability analysis through correlated sampling of the model predictions at different inputs. The Monte Carlo sampling (MCS) error, which represents the error due to limited Monte Carlo samples, is quantified by constructing a probability density function. All the different sources of epistemic uncertainty are quantified and aggregated to estimate the uncertainty in the reliability analysis. Two examples are used to demonstrate the proposed techniques.

Keywords Reliability · Epistemic · Surrogate · Gaussian process · Uncertainty Quantification

Nomenclature

$\mathbf{X}$	A vector of input random variables in a system
$g(\mathbf{X})$	Physics model
p_f^c	Failure probability of a component (single limit state)
$I(.)$	Indicator function for identifying samples in a failure region
p_f^s	Failure probability of a system (with multiple limit states)
n_{MCS}	Number of Monte Carlo samples for reliability analysis
$\hat{g}(\mathbf{X})$	Surrogate model built for the simulation model
$\mathbf{x}$	A realization of the vector of input random variables
$g_{ext}(\mathbf{X})$	Limit state function (either a component or a system limit state)
$\hat{g}_{ext}(\mathbf{X})$	A surrogate for the limit state (component or system)
EFF	Expected Feasibility Function
X	A single input random variable
x	A realization of an input random variable
$\boldsymbol{\Theta}$	A vector of distribution parameters
$\boldsymbol{\theta}$	A realization of the vector of distribution parameters
BMA	Bayesian model averaging
BHT	Bayesian hypothesis testing

✉ Sankaran Mahadevan
sankaran.mahadevan@vanderbilt.edu

[1] Department of Civil and Environmental Engineering, Vanderbilt University, Station B, Box 1831, Nashville, TN 37235, USA

Published online: 02 June 2016

w_k	Weight corresponding to the k^{th} distribution type in BMA
$\mathbf{D}$	Data available on an input random variable
$\mathbf{\Psi}$	Parameters of a simulation model
$\delta(\mathbf{X})$	True model discrepancy for the simulation model
$g_{obs}(\mathbf{X})$	Experimental observations
Y	Output of the physics model ($Y = g(\mathbf{X})$)
$F_Y(y)$	True CDF of model output
$\overline{F}_Y(y)$	Estimated CDF of model output
$g_{\mathrm{mod}\,el}(\mathbf{X})$	Simulation model to the physics model
$\hat{\delta}(\mathbf{X})$	An estimate for the model discrepancy
U_{MD}	Auxiliary variable to represent the uncertainty in model discrepancy estimate
u_{MD}	A realization of the auxiliary variable for model discrepancy
U_X	Auxiliary variable for aleatory uncertainty in an input random variable
u_X	A realization of the auxiliary variable for aleatory uncertainty in an input
$\boldsymbol{\varphi}$	A variable used to represent the set of input random variables, uncertain model parameters and uncertainty in model discrepancy prediction
U_{AK}	Coefficient of variation of surrogate prediction
p_f^{MCS}	Failure probability after inclusion of Monte Carlo simulation error
U_{MCS}	Auxiliary variable for the epistemic uncertainty in reliability estimate

1 Introduction

Optimization under uncertainty has been studied in two directions – (1) Reliability-based Design Optimization (RBDO), and (2) Robust Design Optimization (RDO). One of the crucial elements in an RBDO problem is reliability analysis. Techniques for reliability analysis can be categorized into analytical methods and simulation-based methods (Du 2010). Analytical methods such as First Order Reliability Method (FORM) and Second Order Reliability Method (SORM) approaches employ first order and second order approximations of the limit state (Haldar and Mahadevan 2000a). First-order and second-order bounds for system reliability estimates have been proposed based on first-order and second-order approximations of the limit states (Mahadevan et al. 2001; Song and Der Kiureghian 2003). The FORM and SORM-based methods become inaccurate when the limit-states are highly nonlinear. Monte Carlo sampling (MCS) approaches can be accurate but computationally expensive. To reduce the computational effort, surrogate-based reliability analysis methods have been developed (Hu and Du 2015; Wang and Wang 2015; Youn and Choi 2004). For illustration, consider a performance function represented as $g(\mathbf{X})$ and let $g(\mathbf{X}) = 0$ represent the limit state. If the computational model for $g(\mathbf{X})$ is expensive, two categories of surrogates have been used in the reliability analysis literature – surrogates that estimate the output for any given input i.e. $g(\mathbf{X})$, and surrogates that particularly model the failure limit state i.e. $g(\mathbf{X}) = 0$. For convenience, the surrogates that replace $g(\mathbf{X})$ and $g(\mathbf{X})$ - $g_o = 0$ are termed in this paper as general-purpose surrogates and limit state surrogates, respectively.

In the context of general purpose surrogate (commonly known as response surface), Faravelli (1989), Choi et al. (2004) used a polynomial expansion response surface model, Papadrakikis et al. (1996) used a neural network based surrogate, Dubourg et al. (2013), Kaymaz (2005) used a Kriging (or Gaussian process) surrogate. In the case of limit state surrogates (which are basically classifiers), Bichon et al. (2008, 2011) and Echard et al. (2011) used a Gaussian process (GP) surrogate while Song et al. (2013) used a support vector machine (SVM)-based surrogate. In this paper, the limit state surrogates (classifiers) are considered. The surrogate-based reliability methods have mainly considered only aleatory uncertainty (natural variability) so far. However, practical reliability analysis is affected by many sources of epistemic uncertainty (lack of knowledge); therefore, this paper investigates approaches to include such uncertainty in surrogate-based reliability analysis.

Epistemic uncertainty can be divided into two categories – statistical uncertainty and model uncertainty. Statistical uncertainty stems from inadequacies in the available data (e.g., sparse, imprecise, qualitative, missing, or erroneous) (Sankararaman and Mahadevan 2011; Wang et al. 2009; Youn and Wang 2006) which results in uncertainty regarding the probability distributions of the input random variables. Model uncertainty is due to uncertainty in model parameters, numerical solution errors, and model form error. The model form error and numerical solution errors together are referred to in this paper as model discrepancy. In recent years, efforts have been made to account for both aleatory and epistemic uncertainty within reliability analysis, such as the auxiliary variable approach (Der Kiureghian and Ditlevsen 2009; Nannapaneni and Mahadevan 2015; Sankararaman and Mahadevan 2013a), the conditional reliability index method (Hu et al. 2015), and the Bayesian network approach (Mahadevan et al. 2001; Straub and Der Kiureghian 2010). These methods have only concentrated on component reliability analysis (i.e. a single limit state) considering a few sources of epistemic uncertainty.

Uncertainty due to the use of a surrogate, also called surrogate uncertainty, is an important source of uncertainty when using a surrogate for reliability analysis. In this paper, we are using the term surrogate uncertainty to represent the variance in the surrogate prediction and not the bias. For instance, when a GP surrogate is used, the prediction at any input is a Gaussian

distribution. Another source of uncertainty that is often ignored is the Monte Carlo sampling (MCS) error, which is the uncertainty due to the use of limited Monte Carlo samples in reliability analysis. As the number of Monte Carlo samples increases, the MCS error decreases and as the number of samples reaches infinity, the MCS error tends to zero. Since an infinite number of samples is not possible, there exists some residual MCS error that needs to be quantified and included in the reliability analysis. Techniques for systematic incorporation of various epistemic uncertainty sources, along with surrogate uncertainty and MCS error in limit state surrogate-based reliability analysis are not available yet. Therefore, this paper seeks to incorporate all the above stated sources of uncertainty in limit state surrogate-based reliability analysis. Here, "all" refers to all the uncertainties discussed up to and including that paragraph, which include aleatory and statistical uncertainty in input random variables (distribution parameter, distribution type), uncertain model parameters, model discrepancy, surrogate uncertainty and Monte Carlo sampling errors. Both component and system reliability analyses are considered. For the sake of convenience, reliability analysis based on a limit state surrogate is termed as surrogate-based reliability analysis in the rest of the paper.

For the construction of limit state surrogate, the physics-based simulation model is assumed to be available. Along with the simulation model, the model discrepancy associated with it is assumed to be quantified using calibration experiments and available. If the Kennedy and O'Hagan calibration framework (Kennedy and O'Hagan 2001) is used, then the model discrepancy is represented using a GP model. Kennedy and O'Hagan, in their paper (Kennedy and O'Hagan 2001) have proposed two approaches for model calibration – (1) fully Bayesian, and (2) modular Bayesian. A good overview of both fully and marginal Bayesian approaches is available in (Arendt et al. 2012). In the fully Bayesian approach, the model parameters of simulation model and the hyper parameters are calibrated together by obtaining joint posterior distributions. Using the joint posterior distributions, marginal distributions can be obtained.

The model calibration using the modular Bayesian approach can be summarized in the following steps (Arendt et al. 2012): (1) Estimation of MLE (maximum likelihood estimates) of the hyper parameters of the simulation model GP surrogate using the simulation data, (2) Estimation of MLE of hyper parameters of the model discrepancy GP surrogate using the experimental data, simulation data, and hyper parameters of the simulation model GP surrogate, (3) Calibration of model parameters based on the estimated hyper parameters of the GP surrogates, and (4) Prediction, where the overall prediction is marginalized over the model parameters. The model prediction at any input, when conditioned on the model parameters and MLE of hyper parameters of GP surrogates, follows a Gaussian distribution. However, the unconditional prediction at any input is obtained by marginalizing over the posterior distributions of the model parameters; in this case, the model prediction is not Gaussian. When the overall prediction is marginalized over the model parameters, an expected value for the reliability estimate is obtained. The above marginalization results in the overall prediction to not being Gaussian. However, since the goal in this paper is to quantify the uncertainty in the reliability estimate (as opposed to an expected value), marginalization over model parameters is not performed and the model parameters (with their posterior distributions) are treated just like any stochastic inputs. In this case, the prediction (for a given realization of system input and model parameter) will be Gaussian. As stated in (Arendt et al. 2012), the separation of both the GP models for calibration is intuitive and therefore, the modular Bayesian approach for calibration is adopted.

The outputs of the model calibration analysis using KOH framework are – (1) data on the inputs, simulation output and observations, which are used for model calibration, (2) a GP model for the simulation model, (3) a GP model for the model discrepancy, and (4) Posterior distribution of the model parameters. The above four elements are later used in the construction of a limit-state surrogate for reliability analysis.

Earlier studies such as (Dubourg et al. 2013; Kaymaz 2005; Bichon et al. 2008, 2011; Echard et al. 2011) considered construction of the limit state surrogate for the physics-based simulation model without any model discrepancy. In this paper, the model discrepancy is also considered in constructing the limit state surrogate. However, construction of any surrogate requires point values of paired input–output data whereas in the presence of model discrepancy, the output at any input is a PDF. This problem is overcome in this paper by using an auxiliary approach, which allows for a one-to-one relationship between input and output.

Several types of surrogates are available in the literature such as polynomial chaos, Gaussian process (GP), support vector machines, neural networks etc. This paper uses a GP surrogate to *model the limit state* following the Efficient Global Reliability Analysis (EGRA) approach proposed by Bichon et al. (2008, 2011). However, the techniques proposed in this paper are not limited to GP models and can be extended to any surrogate. The GP surrogate of the simulation model obtained from the KOH framework-based calibration analysis is further refined to construct the limit state surrogate by adding more training points close to the limit state. The training points are adaptively selected by maximizing a learning function called the Expected Feasibility Function (EFF). More details about the EFF and selection of training points are provided in Section 2.1.2.

In the presence of statistical uncertainty, an input variable can be represented through a family of PDFs, where each PDF corresponds to a realization of the distribution parameters and a distribution type. The sampling of the input variable can be done through a nested sampling approach where the distribution type and parameters are sampled in the outer loop, and the samples of the input variable are generated in the inner loop. This nested-loop procedure is computationally expensive; therefore a faster single loop sampling approach using the probability integral transform (Sankararaman and Mahadevan 2011) is used here.

The major contribution through this paper is a unified framework connecting the model calibration analysis to constructing a limit state surrogate and estimating the uncertainty in reliability analysis by incorporating different sources of epistemic uncertainty. The basic contributions include: (1) The use of the auxiliary variable approach to represent the model discrepancy for its inclusion in limit state surrogate refinement, and (2) Quantification of different types of epistemic uncertainty and their incorporation in reliability analysis to quantify the uncertainty in the reliability estimate.

The remainder of this paper is organized as follows: In Section 2, brief backgrounds on component and system reliability analysis in the presence of aleatory uncertainty, and on the quantification of different sources of epistemic uncertainty are provided. Section 3 presents the proposed methodology for component and system reliability analysis in the presence of various sources of epistemic uncertainty. Two numerical examples are used to demonstrate the proposed methodology in Section 4. Concluding remarks are provided in Section 5.

2 Background

2.1 Reliability analysis with a limit state surrogate

Here, the definitions for component and system reliability, and an introduction to Efficient Global Reliability Analysis (EGRA) are provided.

2.1.1 Component and system reliability estimation using a surrogate

Component reliability Consider a single limit state $g(\mathbf{X})=0$, where $\mathbf{X}=[X_1,\ X_2,\ \cdots,\ X_n]$ is a vector of input random variables. The component failure probability (p_f^c) is given as

$$p_f^c = \Pr\{g(\mathbf{X})\leq 0\} \tag{1}$$

In surrogate-based reliability analysis using MCS, the component failure probability can be calculated as

$$p_f^c = \sum_{j=1}^{n_{MCS}} I\left(\hat{g}\left(\mathbf{x}^{(j)}\right)\leq 0\right)/n_{MCS} \tag{2}$$

where $\mathbf{x}$ is a realization of random variables $\mathbf{X}$, n_{MCS} is the number of MCS samples, $\hat{g}(\mathbf{x})$ is the surrogate, $\hat{g}(\mathbf{x}^{(j)})$ is the surrogate prediction at the j^{th} sampling point $\mathbf{x}^{(j)}$, and $I(\hat{g}_i(\mathbf{x}^{(j)})\leq 0)$ is a failure indicator function defined as

$$I(\text{event}) = \begin{cases} 1, \text{if an event is true} \\ 0, \text{otherwise} \end{cases} \tag{3}$$

System reliability System failure events may be defined through a series, parallel or a mixed series/parallel combination of component failures (Wang et al. 2011). Let $g_i(\mathbf{X})$, $i=1,\ 2,\ \cdots,\ m$ be the limit-state functions. The failure probability of a series combination of individual failures is given by

$$p_f^s = \Pr\left\{\bigcup_{i=1}^{m} g_i(\mathbf{X})\leq 0\right\} \tag{4}$$

where p_f^s is the system failure probability, $\Pr\{\cdot\}$ is probability, and "$\cup$" is union. The system failure probability, using a surrogate and MCS, can be calculated as

$$p_f^s = \sum_{j=1}^{n_{MCS}} I\left(\bigcup_{i=1}^{m} \hat{g}_i\left(\mathbf{x}^{(j)}\right)\leq 0\right)/n_{MCS} \tag{5}$$

where $\hat{g}_i(\mathbf{x}^{(j)})$ is the surrogate prediction of the i^{th} limit-state function at the j^{th} sampling point $\mathbf{x}^{(j)}$. Equation (5) can be rewritten as

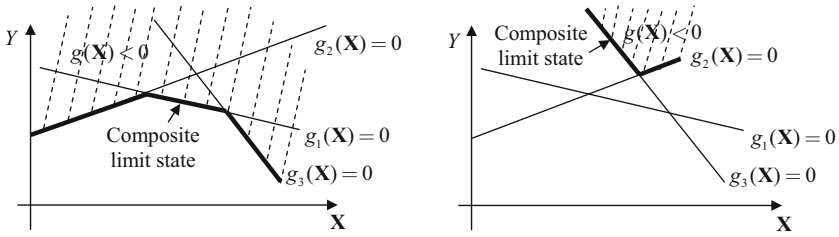
$$\begin{aligned} p_f^s &= \sum_{j=1}^{n_{MCS}} I\left(\min_{i\in m}\left\{\hat{g}_i\left(\mathbf{x}^{(j)}\right)\right\}\leq 0\right)/n_{MCS} \\ &= \sum_{j=1}^{n_{MCS}} I\left(\hat{g}_{\min}\left(\mathbf{x}^{(j)}\right)\leq 0\right)/n_{MCS} \end{aligned} \tag{6}$$

The failure probability of a parallel combination of individual failures is given by

$$p_f^s = \Pr\left\{\bigcap_{i=1}^{m} g_i(\mathbf{X})\leq 0\right\} \tag{7}$$

where "$\cap$" is intersection. Similar to the series system, the failure probability of a parallel system can be calculated as

$$p_f^s = \sum_{j=1}^{n_{MCS}} I\left(\bigcap_{i=1}^{m} \hat{g}_i\left(\mathbf{x}^{(j)}\right)\leq 0\right)/n_{MCS} \tag{8}$$



(a) Series system (b) Parallel system

Fig. 1 Composite limit state of three individual limit states

The above equation can be written as

$$p_f^s = \sum_{j=1}^{n_{MCS}} I\left(\max_{i\in m}\left\{\hat{g}_i\left(\mathbf{x}^{(j)}\right)\right\}\le 0\right)/n_{MCS}$$
$$= \sum_{j=1}^{n_{MCS}} I\left(\hat{g}_{\max}\left(\mathbf{x}^{(j)}\right)\le 0\right)/n_{MCS} \quad (9)$$

Figure 1 provides graphical illustrations of composite limit states for series and parallel combinations of three individual limit states $g_i(\mathbf{X})=0$, $i=1, 2, 3$. When a system failure is defined through a mixture of series and parallel combination of component failures, the system reliability can be estimated by the following steps (Haldar and Mahadevan 2000b): (1) decompose the combined system into a set of mutually exclusive series combinations, (2) compute the reliability of each series combination, and (3) compute the system reliability using the reliabilities of each of the series combinations.

2.1.2 Efficient global reliability analysis (EGRA)

In EGRA, a GP surrogate is constructed to approximate limit state function $g_{ext}(\mathbf{X})=0$, i.e., we are building a surrogate for the classifier or boundary between failure and safety. (Note that $g_{ext}(\mathbf{X})=g(\mathbf{X})$ for a component, $g_{ext}(\mathbf{X})=g_{\min}(\mathbf{X})$ for a series system; and $g_{ext}(\mathbf{X})=g_{\max}(\mathbf{X})$ for a parallel system). Using the idea of the Expected Improvement (EI) (Jones et al. 1998), EGRA adaptively selects training points close to the limit state to accurately model $g_{ext}(\mathbf{X})$. The selection of training points is based on a learning function called the Expected Feasibility Function (EFF) defined as (Bichon et al. 2008)

$$EFF(\mathbf{x}) = \left(\mu_g(\mathbf{x})-e\right)\left[2\Phi\left(\frac{e-\mu_g(\mathbf{x})}{\sigma_g(\mathbf{x})}\right)-\Phi\left(\frac{e^L-\mu_g(\mathbf{x})}{\sigma_g(\mathbf{x})}\right)-\Phi\left(\frac{e^U-\mu_g(\mathbf{x})}{\sigma_g(\mathbf{x})}\right)\right]$$
$$-\sigma_g(\mathbf{x})\left[2\phi\left(\frac{e-\mu_g(\mathbf{x})}{\sigma_g(\mathbf{x})}\right)-\phi\left(\frac{e^L-\mu_g(\mathbf{x})}{\sigma_g(\mathbf{x})}\right)-\phi\left(\frac{e^U-\mu_g(\mathbf{x})}{\sigma_g(\mathbf{x})}\right)\right]$$
$$-\left[\Phi\left(\frac{e^L-\mu_g(\mathbf{x})}{\sigma_g(\mathbf{x})}\right)-\Phi\left(\frac{e^U-\mu_g(\mathbf{x})}{\sigma_g(\mathbf{x})}\right)\right] \quad (10)$$

where e is the failure threshold ($e=0$ in this paper), $e^U=e+\varepsilon$, $e^L=e-\varepsilon$, $\mu_g(\mathbf{x})$, $\sigma_g(\mathbf{x})$ are the mean and standard deviation of the GP prediction at point $\mathbf{X}=\mathbf{x}$, ε is usually chosen as $\varepsilon=2\sigma_g(\mathbf{x})$ (Bichon et al. 2008), and $\Phi(\cdot)$ and $\phi(\cdot)$ are the CDF and PDF of a standard Gaussian variable, respectively. In EGRA, a new training point is identified by maximizing the *EFF* as $\mathbf{x}^* = \text{argmax}_{\mathbf{x}}\{EFF(\mathbf{x})\}$. More details about EGRA are available in (Bichon et al. 2011). EGRA and other similar methods that focus on limit state surrogates such as META-IS (Dubourg et al. 2013) and AK-MCS (Echard et al. 2011) have so far concentrated on reliability analysis with only aleatory uncertainty. In practical applications, several sources of epistemic uncertainty may be involved in the reliability analysis. The next section discusses several sources of epistemic uncertainty and their quantification.

2.2 Quantification of epistemic uncertainty

2.2.1 Statistical uncertainty regarding the inputs

In practical applications, it is common to have only data on an input variable X, which could be sparse, and/or interval data causing uncertainty in its PDF. Both parametric and non-parametric approaches have been developed to address the issue of statistical uncertainty.

(a) *Parametric approach*

In a parametric approach, an input variable is represented using a distribution type and distribution parameters. The presence of limited data causes uncertainty regarding the distribution type and its parameters, which are expressed as probability distributions in a Bayesian approach.

Distribution parameter uncertainty Let a dataset **D** for a variable X consist of n point data p_i ($i=1$ to n) and m interval data $[a_j, b_j]$ ($j=1$ to m). The likelihood function for the distribution parameters θ can be constructed as

$$L(\theta) = \prod_{i=1}^{n} f_X\left(X=p_i\middle|\theta\right)\prod_{j=1}^{m}\left[F_X\left(X=b_j\middle|\theta\right)-F_X\left(X=a_j\middle|\theta\right)\right] \quad (11)$$

where $f_X(x)$ and $F_X(x)$ represent the PDF and CDF of X respectively. From the likelihood function, the PDFs of the distribution parameters can be obtained using Bayes' theorem.

Distribution type uncertainty Two approaches are available to handle distribution type uncertainty - (1) Composite distribution of possible distribution types using Bayesian Model Averaging (BMA) (Hoeting et al. 1999), or (2) Single distribution type that best describes the data using Bayesian Hypothesis Testing (BHT) (Bernardo and Rueda 2002; Ling and Mahadevan 2013). BMA is used here and the composite distribution is formulated as $f_X(x|\boldsymbol{\theta}) = \sum_{k=1}^{N} w_k f_X(x|\boldsymbol{\theta}_k)$ where $f_X(x|\boldsymbol{\theta}_k)$ is the k^{th} distribution type with w_k and $\boldsymbol{\theta}_k$ representing its weight and distribution parameters. The weights can be computed by comparing the posterior probabilities of the distribution types as

$$\frac{\Pr(f_X(x|\boldsymbol{\theta}_c)|\mathbf{D})}{\Pr(f_X(x|\boldsymbol{\theta}_d)|\mathbf{D})} = \frac{\Pr(\mathbf{D}|f_X(x|\boldsymbol{\theta}_c))}{\Pr(\mathbf{D}|f_X(x|\boldsymbol{\theta}_d))} \frac{\Pr(f_X(x|\boldsymbol{\theta}_c))}{\Pr(f_X(x|\boldsymbol{\theta}_d))} \tag{12}$$

where $\Pr(f_X(x|\boldsymbol{\theta}_c))$ and $\Pr(f_X(x|\boldsymbol{\theta}_d))$ are the prior probabilities of the two distribution types and $\Pr(f_X(x|\boldsymbol{\theta}_c)|\mathbf{D})/\Pr(f_X(x|\boldsymbol{\theta}_d)|\mathbf{D})$ is the ratio of their likelihoods. In the presence of multiple plausible distributions, weights for all the distributions can be computed with respect to a particular distribution. In the presence of multiple plausible distributions, weights for all the distributions can be computed with respect to a particular distribution. If $M_1, M_2 \ldots M_n$ are plausible distributions, then ratios of posterior weights can be computed between (M_1, M_n), $(M_2, M_n) \ldots (M_{n-1}, M_n)$. After obtaining the ratio of weights computed using Eq. (12), weights can be normalized. Using the normalized weights and the plausible distribution types, the composite distribution can be constructed. Please refer to (Sankararaman and Mahadevan 2013b) for more details.

(b) *Non-parametric approach*

As opposed to the parametric approach, the non-parametric approach does not assume any particular distribution type or distribution parameters but the PDF is constructed using interpolation techniques.

Let a dataset $\mathbf{D}$ for a variable X consist of n point data $p_i (i = 1 \text{ to } n)$ and m interval data $[a_j, b_j]$ $(j = 1 \text{ to } m)$; the domain of X is discretized into Q points to model the non-parametric distribution. Let the PDF values at these discretized points be equal to q_i $(i = 1, 2, \cdots, Q)$. Since $\mathbf{q} = q_i$ $(i = 1, 2, \cdots, Q)$ is unknown, they can be estimated by solving the following optimization problem:

$$\begin{aligned} &\max_{\mathbf{q}} L(\mathbf{q}) = \prod_{i=1}^{n} f_X\left(X = p_i \middle| \mathbf{q}\right) \prod_{j=1}^{m} \left[F_X\left(X = b_j \middle| \mathbf{q}\right) - F_X\left(X = a_j \middle| \mathbf{q}\right)\right] \\ &s.t.\ \ \mathbf{q} \geq 0;\ \ f_X(x) \geq 0;\ \ \int f_X(x)\, \mathrm{d}x = 1 \end{aligned} \tag{13}$$

After obtaining the PDF values at these discretized points, interpolation techniques are used to estimate the PDF values at any other input values. More details on the likelihood-based non-parametric method are provided in (Sankararaman and Mahadevan 2011).

2.2.2 Model uncertainty

This section discusses the quantification of different types of model uncertainty such as model parameter uncertainty and model discrepancy. In addition, surrogate uncertainty and MCS error are discussed in the context of surrogate-based reliability analysis.

(a) *Model parameter uncertainty*

Model parameter uncertainty represents the uncertainty in the model parameters due to either natural variability or limited data or both. The three possible scenarios of model parameter uncertainty are – (1) model parameter is deterministic but unknown (epistemic uncertainty), (2) model parameter is stochastic with known distribution parameters (aleatory uncertainty), and (3) model parameter is stochastic with unknown distribution parameters (aleatory and epistemic uncertainty). If a model parameter is deterministic but unknown, it can be estimated using available data using least squares, maximum likelihood or Bayesian calibration.

Model parameters ($\boldsymbol{\Psi}$) that are associated with aleatory uncertainty (probability distributions) and with fixed distribution parameters, can be treated similar to input variables for reliability analysis and the techniques used for quantification of uncertainty in the inputs (parametric and non-parametric approaches, described in Section 2.2.1) can also be used for model parameters. If the distribution parameters of $\boldsymbol{\Psi}$ are unknown (both aleatory and epistemic uncertainty), then one of the aforementioned calibration techniques can be used to estimate the distribution parameters using available data. Here, Bayesian calibration is used to explicitly quantify the uncertainty in the model parameters or their distribution parameters.

(b) *Model discrepancy*

Model discrepancy, in this discussion, represents the combined error introduced due to the assumptions and simplifications made in building a model (model form error) as well as the errors that arise in the methodology adopted in solving the model equations (numerical solution errors). Different types of numerical solution errors exist such as discretization error, round-off error, and truncation error. Suppose $g_{obs}(\mathbf{X})$, $g_{\text{mod}\,el}(\mathbf{X})$, and $\delta(\mathbf{X})$ represent the observations, simulation model prediction and model discrepancy respectively. For a given $\mathbf{X} = \mathbf{x}$, the three quantities are related as

$g_{obs}(\mathbf{x}) = g_{\mathrm{mod}\,el}(\mathbf{x}) + \delta(\mathbf{x}) + \varepsilon_{obs}(\mathbf{x})$. Here, $\varepsilon_{obs}(\mathbf{x})$ refers to the observation (or experimental) error. When following the KOH framework, the quantification of the model discrepancy can be performed together with the calibration of model parameters. Note that in this paper, as stated in Section I, we adopt the modular Bayesian approach of the KOH framework.

(c) *Reliability analysis errors*

Different types of errors that arise in carrying out reliability analysis such as surrogate uncertainty and uncertainty quantification (UQ) error are discussed below.

Surrogate uncertainty In this paper, the variance (and not bias) associated with the prediction of a surrogate is called surrogate uncertainty. For example, the prediction of a GP model is a Gaussian distribution with parameters dependent on the input. Since a surrogate (GP model) is employed in this paper for reliability analysis, the details about surrogate uncertainty and its impact on the reliability estimate are discussed later, in Section. 3.2.2.

Monte Carlo Simulation (MCS) error MCS error represents the error due to the use of limited number of Monte Carlo samples for uncertainty propagation. The MCS error is quantified as the difference between the empirical CDF (constructed using Monte Carlo samples after uncertainty propagation) and the true CDF of the output quantity of interest (Sankararaman and Mahadevan 2013b).

Consider the model $Y = g(\mathbf{X})$ and let N_s samples be generated from the input and propagated through the model to obtain the samples of the output Y. Let $\overline{F}_Y(y)$ represent the empirical CDF constructed from N_s samples and $F_Y(y)$ be the true CDF. If n_y samples are less than a value of $Y = y^*$, then n_y follows a binomial distribution $n_y \sim B(N_s, F_Y(y))$ considering the value of each sample as the result of a Bernoulli trial (Haldar and Mahadevan 2000a). As N_s becomes larger, the binomial distribution can be approximated with a Gaussian distribution. A review of several empirical rules that have been proposed in the literature for approximating a binomial distribution with a Gaussian distribution is provided in (Emura and Lin 2015). The most commonly used rule, as stated in (Emura and Lin 2015), is $N_s F_Y(y) > 5$ and $N_s(1 - F_Y(y)) > 5$. Let us consider this case. A large of Monte Carlo samples are generally used in surrogate-based methods since it is computationally inexpensive. Even if we use about 100,000 samples (which is a common number in surrogate-based methods), the threshold $F_Y(y)$ value is $\frac{5}{100000} = 0.00005$, which covers more than 99.99 % of the domain.

Since the CDF value from the Monte Carlo output samples is given by $\overline{F}_Y(y) = n_y/N_s$, we have

$$\overline{F}_Y(y) \tilde{} N\left(F_Y(y),\ \sqrt{F_Y(y)(1-F_Y(y))/N_s}\right) \tag{14}$$

Therefore, the MCS error associated with $\overline{F}_Y(y)$ can be expressed as a Gaussian random variable with mean and standard deviation as given in Eq. (14). Since the true CDF $F_Y(y)$ is unknown, Eq. (14) cannot be used directly. However, confidence intervals for $F_Y(y)$ can be estimated given the empirical CDF $\overline{F}_Y(y)$, the number of samples used N_s and the degree of accuracy $1-\gamma$ as $\left(\overline{F}_Y(y) + \frac{z^2_{\gamma/2}}{2N_s} \pm z_{\gamma/2}\sqrt{\frac{\overline{F}_Y(y)\left(1-\overline{F}_Y(y)\right)}{N_s} + \frac{z^2_{\gamma/2}}{4N_s^2}}\right)\frac{1}{1+\frac{z^2_{\gamma/2}}{N_s}}$, where $z_{\gamma/2}$ refers to the $1-\frac{\gamma}{2}$ quantile of the standard normal distribution (Agresti and Coull 1998). It should be noted that the true CDF $F_Y(y)$ is a fixed quantity but unknown; therefore, it is an epistemic source of uncertainty and quantified using confidence intervals. For given values of $\overline{F}_Y(y)$ and N_s, we can estimate the percentile values of $F_Y(y)$ by varying the accuracy parameter γ. From the percentile values, the entire CDF can be numerically constructed which can then be used to obtain a PDF.

For illustration, let l, u represent the lower and upper bounds of confidence intervals corresponding to accuracy parameter γ. Therefore, $\Pr(F_Y(y) < u) = 1-\frac{\gamma}{2}$ and $\Pr(F_Y(y) < l) = \frac{\gamma}{2}$. Hence, the CDF values at $F_Y(y) = l, u$ are $\frac{\gamma}{2}$ and $1-\frac{\gamma}{2}$ respectively. Following the same procedure at multiple values of γ, the CDF values at the corresponding lower and upper bounds of the confidence intervals can be obtained which can be used to construct the CDF of $F_Y(y)$.

The next section presents techniques for the inclusion of different types of epistemic uncertainty discussed in this section in surrogate-based reliability analysis.

3 Proposed methodology

An overview of the proposed methodology for surrogate-based reliability estimation including different types of epistemic uncertainty is presented in Fig. 2.

The overall approach can be divided into two stages as shown in Fig. 2 – (1) construction of a surrogate, and (2) use of the surrogate for reliability analysis. Uncertainty sources such as statistical uncertainty in the inputs, uncertain model parameters, and model discrepancy influence the surrogate construction, thereby affecting the reliability estimate, whereas the surrogate uncertainty and MCS error do not influence the surrogate construction but only affect the reliability estimate. Details regarding the handling of different sources of epistemic uncertainty in these two stages are discussed below.

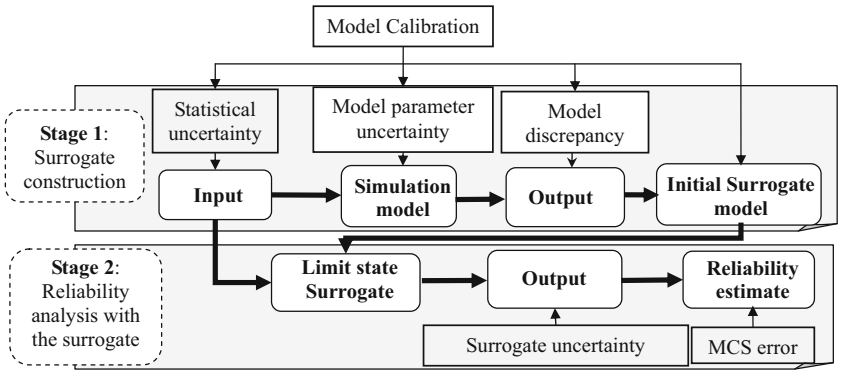


Fig. 2 Flow chart of proposed method for reliability analysis under uncertainty

3.1 Stage 1: surrogate construction including epistemic uncertainty

First, the different inputs to be included in the surrogate are discussed which is followed by the surrogate construction.

3.1.1 Inputs to the surrogate

In this paper, the inputs for the surrogate model include the inputs for the simulation model ($\mathbf{X}$), the uncertain model parameters ($\mathbf{\Psi}$) and the model discrepancy. Consider a system with m limit state functions given by $g_i(\mathbf{X}, \mathbf{\Psi}) = g_{model,\ i}(\mathbf{X}, \mathbf{\Psi}) + \hat{\delta}_i(\mathbf{X}), \quad i = 1,\ 2,\ \cdots,\ m$ where $g_{model,\ i}(\mathbf{X}, \mathbf{\Psi})$ and $\hat{\delta}_i(\mathbf{X})$ are the simulation model and the model discrepancy of the i^{th} limit state function respectively. Since KOH calibration framework is used, the model discrepancy is not modeled as a function of model parameters. However, including the model parameters in the model discrepancy term might be a more rigorous. The quantification of the model discrepancy is by using the simulation model. The performance of the overall model (simulation model and model discrepancy) might be satisfactory over the entire range of the inputs but the main idea of the limit state surrogate is to model the limit state perfectly and is not concerned about the its performance in the interior domain of the inputs (away from limit state). Since the limit state model governs the reliability estimate (and its uncertainty), we wish to model it as precisely as possible. In system reliability, it is possible in some cases that we may have some calibration data on a few limit states but not a lot of data on the remaining limit states. Therefore, the construction of the limit state surrogate refines the system limit state by obtaining training points considering all the individual limit states. The focus of this work is to quantify the uncertainty in the reliability prediction by investigating various sources of epistemic uncertainty. It is based on the assumption that the model calibration has already been performed. We obtain a GP surrogate for the model discrepancy from calibration analysis.

The key idea is to include the model discrepancy terms in the construction of the limit state surrogate and then use the surrogate for reliability analysis. The limit-state function $g_{ext}(\mathbf{X}, \mathbf{\Psi})$ for surrogate-based reliability analysis is formulated as

$$g_{ext}(\mathbf{X}, \mathbf{\Psi}) = \begin{cases} g_{model}(\mathbf{X}, \mathbf{\Psi}) + \hat{\delta}(\mathbf{X}), & \text{for a component} \\ \min\limits_{i=1,\ 2,\ \cdots,\ m} \left\{ g_{model,\ i}(\mathbf{X}, \mathbf{\Psi}) + \hat{\delta}_i(\mathbf{X}) \right\}, & \text{for a series system} \\ \max\limits_{i=1,2,\ \cdots,\ m} \left\{ g_{model,\ i}(\mathbf{X}, \mathbf{\Psi}) + \hat{\delta}_i(\mathbf{X}) \right\}, & \text{for a parallel system} \end{cases} \tag{15}$$

Note that the model discrepancy, $\hat{\delta}_i(\mathbf{X})$, $i = 1,\ 2,\ \cdots,\ m$, is random at any given point $\mathbf{X} = \mathbf{x}$, which results in uncertainty in the response $g_{ext}(\mathbf{X}, \mathbf{\Psi})$. Directly constructing a surrogate for the implicit response given in Eq. (15) is not practical due to randomness in $\hat{\delta}_i(\mathbf{X})$. An explicit representation of variability in $\hat{\delta}_i(\mathbf{X})$ is required to formulate Eq. (15) as a deterministic function; this challenge is addressed using the auxiliary variable method.

The "auxiliary variable" was introduced by Sankararaman and Mahadevan (2013a, b) in the context of random variables whose distribution types and parameters are uncertain. Such variables are associated with both aleatory and epistemic uncertainty. The auxiliary variable is used to represent the aleatory part of the overall uncertainty. Mathematically, an auxiliary variable u, based on the probability integral transform, is defined as $u = F_X(x|\Theta = \theta) = \int_{-\infty}^{x} f_X(w|\Theta = \theta)\, dw$, where θ is a realization of the distribution parameter Θ, w is a dummy variable used for integration, and $F_X(x|\Theta = \theta)$ denotes the CDF value of variable X for a realization of Θ. For a given realization of Θ, and a realization of the auxiliary variable u gives a unique value of X. This gives a deterministic relationship between (Θ, u) and X. In this work, we have extended the concept of the auxiliary variable further, to represent the epistemic uncertainty in the prediction of model discrepancy and in the reliability estimate. In general, whenever there is a stochastic mapping,

i.e., mapping of a single value to a probability distribution, the auxiliary variable can be used to convert the stochastic mapping to a deterministic mapping.

An auxiliary variable is used to represent the variability in the prediction of model discrepancy in order to build the limit state surrogate. Since a GP is used to model $\hat{\delta}_i(\mathbf{X})$, for any given $\mathbf{X}=\mathbf{x}$, the model discrepancy follows a Gaussian distribution given by $\hat{\delta}_i(\mathbf{x}) \sim N\left(\mu_{\delta_i(\mathbf{x})},\ \sigma^2_{\delta_i(\mathbf{x})}\right)$. The PDF of $\hat{\delta}_i(\mathbf{x})$ is $f_\delta\left(\delta_i(\mathbf{x})|\mu=\mu_{\delta_i(\mathbf{x})},\ \sigma=\sigma_{\delta_i(\mathbf{x})}\right)$ and the CDF value is given by

$$u_{MD,i} = F_\delta\left(\delta_i(\mathbf{x},\psi)|\mu=\mu_{\delta_i(\mathbf{x},\psi)},\ \sigma=\sigma_{\delta_i(\mathbf{x})}\right) = \int_{-\infty}^{\delta_x} f_\delta\left(w|\mu=\mu_{\delta_i(\mathbf{x})},\sigma=\sigma_{\delta_i(\mathbf{x})}\right)dw \tag{16}$$

where w is a dummy variable for integration. The auxiliary variable follows a uniform distribution between 0 and 1 (since the CDF values range from 0 to 1). Using the auxiliary variable, a realization of $\hat{\delta}_i(\mathbf{x},\psi)$ for a given $U_{MD,i}=u_{MD,i}$, $\boldsymbol{\mu}=\mu_{\delta_i(\mathbf{x})}$, and $\boldsymbol{\sigma}=\sigma_{\delta_i(\mathbf{x})}$ can be computed by

$$\delta_i(\mathbf{x}) = F^{-1}_{\delta_i(\mathbf{x})}\left(U_{MD,i}=u_{MD,i}|\boldsymbol{\mu}=\mu_{\delta_i(\mathbf{x})},\ \boldsymbol{\sigma}=\sigma_{\delta_i(\mathbf{x})}\right) \tag{17}$$

where $F^{-1}_{\delta_i(\mathbf{x})}(\cdot)$ is the inverse Gaussian CDF of $\hat{\delta}_i(\mathbf{x})$. More details about the auxiliary variable approach are available in (Song et al. 2013). Combining Eqs. (15) and (17), we have

$$g_{ext}(\mathbf{X},\boldsymbol{\Psi},\ \mathbf{U}) = \begin{cases} g_{\text{model}}(\mathbf{X},\boldsymbol{\Psi}) + F^{-1}_{\delta(\mathbf{x})}\left(U_{MD}|\mu=\mu_{\delta(\mathbf{x})},\ \sigma=\sigma_{\delta(\mathbf{x})}\right), & \text{for a component} \\ \min\limits_{i=1,\ 2,\ \cdots,\ m}\left\{g_{\text{model},\ i}(\mathbf{X},\boldsymbol{\Psi}) + F^{-1}_{\delta_i(\mathbf{x})}\left(U_{MD,i}|\mu=\mu_{\delta_i(\mathbf{x})},\ \sigma=\sigma_{\delta_i(\mathbf{x})}\right)\right\}, & \text{for a series system} \\ \max\limits_{i=1,\ 2,\ \cdots,\ m}\left\{g_{\text{model},\ i}(\mathbf{X},\boldsymbol{\Psi}) + F^{-1}_{\delta_i(\mathbf{x})}\left(U_{MD,i}|\mu=\mu_{\delta_i(\mathbf{x})},\ \sigma=\sigma_{\delta_i(\mathbf{x})}\right)\right\}, & \text{for a parallel system} \end{cases} \tag{18}$$

Thus, the original model with stochastic output is mapped to a deterministic model, which can then be used to build a surrogate for reliability analysis. Initial GP surrogates (for simulation model and model discrepancy) are obtained from the KOH calibration framework as shown in Stage 1 of Fig. 2. These surrogates can directly be used for reliability analysis. However, if better accuracy is desired, then the GP surrogate of the simulation model can be further refined around the limit state (Stage 2 in Fig. 2). This local refinement requires additional training points and these are obtained by evaluating the simulation model Eq. (18). After constructing the limit state surrogate, it is directly used for reliability analysis without any further runs of the original model.

3.1.2 Including epistemic uncertainty in surrogate construction

As stated in Section 2.1.2, the training points for the construction of the limit state surrogate are adaptively selected by maximizing a learning function called the Expected Feasibility Function (EFF). Several optimization techniques (gradient-based, sampling-based) are available for maximizing the EFF. In this paper, a Monte Carlo sampling-based technique is implemented, as demonstrated in the Adaptive-Kriging Monte Carlo Simulation (AK-MCS) method (Echard et al. 2011). The key idea in MCS-based optimization is to generate a pool of samples of the inputs from their corresponding PDFs and choose the sample from the pool that maximizes the EFF. The epistemic uncertainty in the inputs of the limit state surrogate can be included in the surrogate modeling construction through the sampling of input variables, which is discussed below.

Inputs to the simulation model (X) Quantification of uncertainty in the distribution parameters and distribution type of random input variables was discussed in Section 2.2. For a realization of distribution parameters and distribution type, the input variable is represented by a PDF; therefore, for multiple realizations of distribution parameters and type, the input variable is represented through a family of PDFs. The traditional approach for sampling of an input variable with uncertain distribution parameters and distribution type is through a nested double-loop procedure where the distribution parameters and distribution type are sampled in the outer loop and samples of the input random variable are generated in the inner loop. The double-loop sampling procedure is computationally expensive; therefore, a single-loop sampling procedure using an auxiliary variable based on the probability integral transform is used. An auxiliary variable U_X is defined to represent the aleatory uncertainty in a random variable and represented using the probability integral transform as

$$u_X = F_X\left(x\middle|\boldsymbol{\Theta}_X=\boldsymbol{\theta}_X, d_X=d_X^*\right) = \int_{-\infty}^{x} f_X\left(w\middle|\boldsymbol{\Theta}_X=\boldsymbol{\theta}_X, d_X=d_X^*\right)dw \tag{19}$$

where $\boldsymbol{\Theta}_X, d_X$ represent the uncertain distribution parameters and distribution type; their realizations are represented by $\boldsymbol{\theta}_X$,

d_X^*, w is a dummy variable for integration, $F_X(x)$ and $f_X(x)$ represent the CDF and PDF of X respectively. Thus, for a realization of auxiliary variable $U_X = u_X$, distribution parameters θ_X and distribution type d_X^* generated from their corresponding PDFs, one realization of the input variable can be obtained as $x = F_X^{-1}(u_X | \Theta_X = \theta_X, d_X = d_X^*)$ using the inverse CDF method. Following this procedure, several realizations of the input can be obtained.

If a non-parametric approach is used for the representation of statistical uncertainty (i.e. no distribution type or parameters), its CDF is constructed using numerical techniques which is then used for generating samples using inverse CDF technique. If an input variable is not associated with any statistical uncertainty (i.e. only known distribution type and distribution parameters), then the conventional inverse CDF technique (Haldar and Mahadevan 2000b) can be used for generating samples.

To perform sampling of correlated variables, the correlated variables are first transformed into an uncorrelated space using orthogonal transformation or Cholesky decomposition. Samples are then generated individually in the uncorrelated space and transformed into the correlated space. Please refer to Chapter 9 in (Haldar and Mahadevan 2000b) for more details.

Uncertain model parameters ($\boldsymbol{\Psi}$) Since we are adopting the modular Bayesian approach, the obtained posterior distributions of the model parameters are conditioned on the MLE of hyper parameters, and correlations are not calculated between the model parameters and the model discrepancy. Correlations would be calculated when the fully Bayesian approach of KOH framework is implemented. As stated in Section 2.2.2, the uncertain model parameters ($\boldsymbol{\Psi}$) can be treated similar to the inputs; therefore the sampling techniques presented above for the inputs can also be used for sampling the uncertain model parameters.

Model discrepancy ($\mathbf{U_{MD}}$) Section 3.1.1 proposed an auxiliary variable approach to explicitly represent the uncertainty in the model discrepancy. Since the auxiliary variables follow a uniform distribution between 0 and 1, their sampling can be carried out through random uniform sampling between 0 and 1.

From the model calibration analysis, as stated in Section 1, we have data on the inputs, model parameters and simulation model output that was used for calibration. If we have k samples, then k random samples of the model discrepancy ($\mathbf{U_{MD}}$) are generated and these data can be used as initial training points to construct the limit state surrogate. Then, more training points are adaptively added by maximizing the EFF until a convergence criterion is satisfied. In this paper, convergence is assumed to be achieved when the maximum EFF is less than a threshold value ($\text{EFF}^{\max} < \text{EFF}^*$).

3.2 Stage 2: uncertainty quantification in reliability analysis

3.2.1 Reliability analysis

In Stage 1, several Monte Carlo samples of the inputs ($\mathbf{X}$, $\boldsymbol{\Psi}$, $\mathbf{U_{MD}}$) were generated to carry out the maximization of EFF for adaptive selection of training points for limit state surrogate construction. Since these samples are generated from their corresponding PDFs, these samples can also be used to carry out reliability analysis. Using these samples, the failure probability can be calculated as

$$p_f^s = \sum_{j=1}^{n_{MCS}} I\left(\hat{g}_{ext}\left(\mathbf{x}^{(j)},\ \mathbf{u_{MD}}^{(j)},\ \boldsymbol{\Psi}^{(j)}\right) \le 0\right) / n_{MCS} \qquad (20)$$

where $\mathbf{x}^{(j)}$, $\mathbf{u_{MD}}^{(j)}$, $\boldsymbol{\Psi}^{(j)}$ are the j-th sample of $\mathbf{X}$, $\mathbf{UMD}$, $\boldsymbol{\Psi}$, which are input variables to the simulation model, auxiliary variables for model discrepancy and uncertain model parameters, respectively. $\hat{g}_{ext}(.)$ represents the surrogate built to approximate the true limit state $g_{ext}(.)$ and $I(.)$ is the failure indicator function as defined in Section 2.1.1

3.2.2 Inclusion of surrogate uncertainty in the reliability estimate

Since a GP surrogate is used, the prediction at any input is a Gaussian distribution with parameters dependent on the input. In most cases, only the mean predictions $\mu_{\hat{g}_{ext}}\left(\mathbf{x}^{(j)},\ \mathbf{u_{MD}}^{(j)},\ \boldsymbol{\Psi}^{(j)}\right)$ are used to estimate $\hat{g}_{ext}(\mathbf{x}^{(j)},\ \mathbf{u_{MD}}^{(j)},\ \boldsymbol{\Psi}^{(j)})$. The failure probability is therefore estimated as

$$\hat{p}_f^s = \sum_{j=1}^{n_{MCS}} I\left(\mu_{\hat{g}_{ext}}\left(\mathbf{x}^{(j)},\ \mathbf{u_{MD}}^{(j)},\ \boldsymbol{\Psi}^{(j)}\right) \le 0\right) / n_{MCS} \qquad (21)$$

where $\hat{p}_f^s$ is the estimate of the failure probability obtained by using the mean predictions. If our purpose is only to estimate the expected failure probability, we may ignore the correlation between the uncertain responses (Jiang et al. 2013). However, the focus of this paper is to quantify the uncertainty in the reliability estimate. Therefore, the uncertainty due the surrogate prediction is also considered to quantify the overall uncertainty in reliability estimate. When the accuracy of the surrogate is high (i.e. the uncertainty of prediction is low), the above treatment of using the mean predictions works well, and results in a single value of the reliability estimate. If the accuracy of the model prediction is low, it becomes necessary to also include the prediction uncertainty for reliability estimation. In order to quantify the effects of surrogate uncertainty on reliability analysis, an uncertainty quantification problem is therefore formulated as shown in Fig. 3.

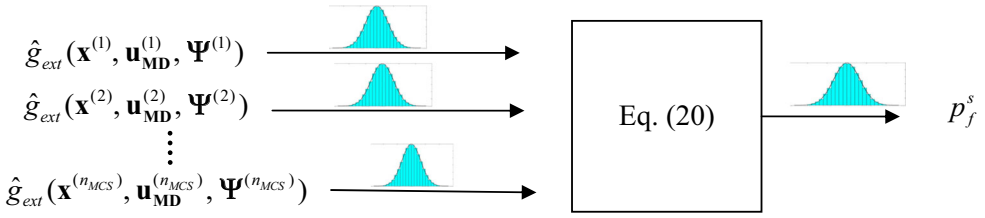


Fig. 3 Effects of surrogate uncertainty on reliability analysis

(a) *Correlation analysis of surrogate predictions*

For any input $\boldsymbol{\varphi}^{(j)} = [\mathbf{x}^{(j)}, \mathbf{u}_{\mathbf{MD}}{}^{(j)}, \mathbf{\Psi}^{(j)}]$, the prediction from the surrogate follows a Gaussian distribution given by $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)}) \sim N\left(\mu_{\hat{g}_{ext}}(\boldsymbol{\varphi}^{(j)}), \sigma^2_{\hat{g}_{ext}}(\boldsymbol{\varphi}^{(j)})\right)$. Also, $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $j = 1, 2, \cdots, n_{MCS}$, are correlated due to the covariance function assumed in a GP.

As indicated in Fig. 3, the uncertainty of the system failure probability estimate (the unconditional failure probability estimate given in Eq. (20)) due to surrogate uncertainty can be quantified by propagating the uncertainty in $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $j = 1, 2, \cdots, n_{MCS}$, through Eq. (20). Since $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $j = 1, 2, \cdots, n_{MCS}$, are n_{MCS} correlated random variables, their correlations are analyzed based on which a sampling-based method is developed for uncertainty quantification.

In MCS-based reliability analysis, n_{MCS} is usually large. Performing the correlation analysis for the n_{MCS} random variables is computationally expensive. In order to reduce the number of random variables ($\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $j = 1, 2, \cdots, n_{MCS}$), $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $j = 1, 2, \cdots, n_{MCS}$ are partitioned into two groups based on the probability that they may result in the error of failure probability estimate. The first group includes responses $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $j = 1, 2, \cdots, n_{g1}$, for which the probability of making an error in the sign of $I(\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)}))$ is very low (i.e., 0.001). Therefore, the mean predictions $\mu_{g_{ext}}(\boldsymbol{\varphi}^{(j)})$ are used to substitute for $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$ in Eq. (20). The remaining responses in $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $j = 1, 2, \cdots, n_{MCS}$ form the second group, which are treated as random variables. The system failure probability given in Eq. (21) then becomes (Hu and Mahadevan 2015)

$$p_f^s = \left(\sum_{i=1}^{n_{g1}} I\left(\mu_{g_{ext}}(\boldsymbol{\varphi}^{(j)})\right) + \sum_{i=1}^{n_{g2}} I\left(\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})\right)\right) \Big/ n_{MCS} \tag{22}$$

where n_{g1} and n_{g2} are the number of samples in the first and second groups respectively. The partition of $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $j = 1, 2, \cdots, n_{MCS}$ is achieved based on the following function (Echard et al. 2011)

$$U_{AK}(\boldsymbol{\varphi}^{(j)}) = \left|\mu_{\hat{g}_{ext}}(\boldsymbol{\varphi}^{(j)})\right| \Big/ \sigma_{\hat{g}_{ext}}(\boldsymbol{\varphi}^{(j)}) \tag{23}$$

$U_{AK}(\boldsymbol{\varphi}^{(j)})$ represents the coefficient of variation of the model prediction, which can be used to estimate the probability of making an error in the sign of $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$. The first group of responses correspond to $U_{AK}(\boldsymbol{\varphi}^{(j)}) \geq 3.1$ and the rest of the responses fall into the second group. Defining the training points in the current GP model as $\boldsymbol{\varphi}^s$ and $g_{ext}(\boldsymbol{\varphi}^s)$, the covariance matrix of $\hat{g}_{ext}(\boldsymbol{\varphi}^{(k)})$, $k = 1, 2, \cdots, n_{g2}$ conditioned on the training points is given by

$$\Sigma_{p|t} = \Sigma_{pp} - \Sigma_{pt}\Sigma_{tt}^{-1}\Sigma_{pt}^T \tag{24}$$

where Σ_{pp}, Σ_{pt}, and Σ_{tt} are the covariance matrixes between $\hat{g}_{ext}(\boldsymbol{\varphi}^{(k)})$ and $\hat{g}_{ext}(\boldsymbol{\varphi}^{(k)})$, $\hat{g}_{ext}(\boldsymbol{\varphi}^{(k)})$ and $g_{ext}(\boldsymbol{\varphi}^s)$, $g_{ext}(\boldsymbol{\varphi}^s)$ and $g_{ext}(\boldsymbol{\varphi}^s)$, respectively. Based on the covariance matrix $\Sigma_{p|t}$, the conditional correlation matrix $\boldsymbol{\rho}_{p|t}$ of $\hat{g}_{ext}(\boldsymbol{\varphi}^{(k)})$, $k = 1, 2, \cdots$, n_{g2} is equal to ρ_{ij}, which represents the correlation between $\hat{g}_{ext}(\boldsymbol{\varphi}^{(i)})$ and $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$, $i, j = 1, 2, \cdots, n_{g2}$, conditioned on current training points. Thus, the correlation matrix is an $n_{g2} \times n_{g2}$ square matrix with diagonal elements as 1.

(b) *Propagation of surrogate prediction uncertainty*

After obtaining the correlation matrix, the sampling-based method can be used to propagate the surrogate prediction uncertainty of $\hat{g}_{ext}(\boldsymbol{\varphi}^{(j)})$ to the uncertainty in p_f^s based on Eq. (22). To do this, samples of $\hat{g}_{ext}(\boldsymbol{\varphi}^{(k)})$, $k = 1, 2, \cdots, n_{g2}$, are generated using the following expression (Hu and Mahadevan 2015):

$$\hat{g}_{ext}(\boldsymbol{\varphi}^{(k)}) = \mu_{\hat{g}_{ext}}(\boldsymbol{\varphi}^{(k)}) + \sigma_{\hat{g}_{ext}}(\boldsymbol{\varphi}^{(k)}) \sum_{j=1}^{n_{g2}} \xi_j \boldsymbol{\Phi}_j^T \boldsymbol{\rho}_{:\,i} \Big/ \sqrt{\eta_j} \tag{25}$$

where ξ_j, $j = 1, 2, \cdots, n_{g2}$, are independent standard Gaussian variables; and η_i and $\boldsymbol{\Phi}_i^T$ are the eigenvalues and eigenvectors of $\boldsymbol{\rho}_{p|t}$ and $\boldsymbol{\rho}_{:\,i} = \left[\rho_{i1}, \rho_{i2}, \cdots, \rho_{in_{g2}}\right]^T$. Since the surrogate prediction at each input is a random variable, N_{sim} samples are generated for each $\hat{g}_{ext}(\mathbf{x}^{(k)})$, $k = 1, 2, \cdots, n_{g2}$, resulting in the following sampling matrix:

$$\mathbf{g}_{N_{sim} \times n_{g2}} = \{g_{ext}(i,j)\}, \forall i = 1, \cdots, N_{sim}; j = 1, \cdots, n_{g_2} \tag{26}$$

Using the samples of $\hat{g}_{ext}(\mathbf{x}^{(k)})$, $k = 1, 2, \cdots, n_{g2}$ and Eq. (22), samples of p_f^s are obtained as

$$p_f^s(j) = \left\{\sum_{i=1}^{n_{g1}} I\left(\mu_{g_{ext}}(\boldsymbol{\varphi}^{(i)})\right) + \sum_{i=1}^{n_{g2}} I(g_{ext}(i,j))\right\} \Big/ n_{MCS}, j = 1, 2, \cdots, N_{sim} \tag{27}$$

where $p_f^s(j)$ is the j^{th} sample of p_f^s due to the surrogate prediction uncertainty. Using the samples of $p_f^s(j)$, $j = 1, 2, \cdots$,

N_{sim}, a PDF can be constructed for the system failure probability p_f^s. This distribution represents the uncertainty in p_f^s due to surrogate uncertainty.

3.2.3 Inclusion of MCS error in reliability estimate

Using Eq. (27), several samples of p_f^s are obtained through correlated sampling of the model predictions at several inputs. As discussed in Sec. 2.2.2 (c), there exists an uncertainty in the estimation of each failure probability sample, $p_f^s(j), j=1, 2, \cdots, N_{sim}$ due to the limited number of Monte Carlo samples (referred to here as MCS error), which results in each $p_f^s(j)$ being a random variable. To avoid any confusion in the notation, the failure probability when MCS error is also considered is denoted as $p_f^{MCS}(j)$. Following the discussion on MCS error in Section 2.2.2 (c), we can construct the PDF of $p_f^{MCS}(j)$ by estimating its quantiles using $p_f^s(j)$, number of samples n_{MCS} and degree of accuracy $1-\gamma$ as

$$\left(p_f^s(j)+\frac{z_{\gamma/2}^2}{2n_{MCS}} \pm z_{\gamma/2}\sqrt{\frac{p_f^s(j)\left(1-p_f^s(j)\right)}{N_s}+\frac{z_{\gamma/2}^2}{4n_{MCS}^2}}\right)\frac{1}{1+\frac{z_{\gamma/2}^2}{n_{MCS}}}, \tag{28}$$

where $z_{\gamma/2}$ represents the $1-\frac{\gamma}{2}$ quantile of a standard normal distribution. After including surrogate uncertainty and MCS error, p_f^s is represented by a family of PDFs (Fig. 4b).

After the inclusion of surrogate uncertainty, the failure probability is represented using a PDF as shown in Fig. 4a. p_f^s is a sample from this PDF (shown as a red dot in Fig. 4a). When MCS error is also considered, p_f^s is not a single value but a PDF denoted as p_f^{MCS} (shown as a continuous red curve in Fig. 4b). Similarly, each sample from the PDF in Fig. 4a corresponds to a different PDF in Fig. 4b. Thus, the failure probability is represented as a family of PDFs in the presence of surrogate uncertainty and MCS error. This family of PDFs can then be integrated to an unconditional PDF (bold broken red curve in Fig. 4b) using the auxiliary variable approach described in Section 3.1.

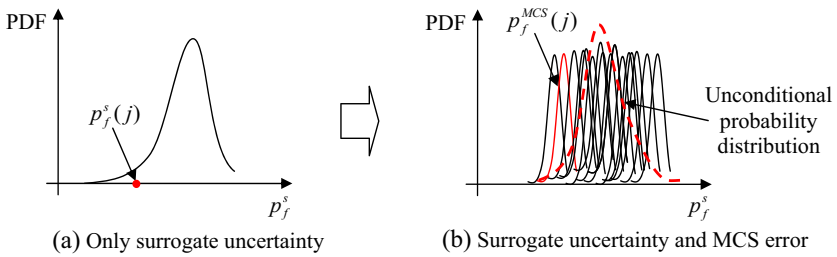
The family of PDFs for the failure probability estimate (Fig. 4b) can be treated similar to the family of PDFs for an input variable and uncertain distribution parameters. Note that in the case of an input variable, the auxiliary variable represents its aleatory uncertainty whereas in the case of the failure probability estimate, it represents the epistemic uncertainty due to limited Monte Carlo samples. The single loop sampling approach used for sampling the input variables can also be used for sampling the failure probability estimates.

Assume that there are N_{sim} PDFs of the failure probability (i.e., N_{sim} samples of p_f^s), N_{sim} samples of the auxiliary variable U_{MCS} are generated in the interval [0, 1]. Here, the auxiliary variable is used to represent the contribution of epistemic uncertainty of MCS error to the overall uncertainty in the reliability estimate. The reliability estimate without the inclusion of surrogate uncertainty and MCS errors is a point-value (single-value). However, when surrogate uncertainty and MCS errors are included, the reliability estimate is represented through a family of PDFs as shown in Fig. 4b. As discussed in Section 2.2.2 (c), the reliability estimate is represented through a PDF with some distribution parameters. Here, these parameters are calculated after incorporating aleatory uncertainty, statistical uncertainty in input variables, uncertain model parameters, model discrepancy and surrogate uncertainty. Therefore, for each realization of distribution parameters, the reliability estimate follows a PDF.

For one realization of the auxiliary variable and p_f^s, one realization of the failure probability estimate from its unconditional PDF is obtained. For a given value of p_f^s, there exists an associated PDF due to MCS error which is numerically estimated after obtaining the quantiles as stated above. Denoting the generated samples as $u_{MCS}^{(1)}, \cdots, u_{MCS}^{(N_{sim})}$, the samples of unconditional failure probability estimate can be obtained as

$$p_f^{new}(j) = F_{MCS}^{-1}\left(u_{MCS}^{(j)} \middle| p_f^s(j)\right), j = 1, 2, \cdots, N_{sim} \tag{29}$$

where $F_{MCS}^{-1}(\cdot)$ is the inverse CDF of the MCS error. Based on the generated samples of $p_f^{new}(j)$, the unconditional PDF of the failure probability can be estimated. This PDF represents the uncertainty in the failure probability estimate due to both surrogate prediction uncertainty and MCS error.



(a) Only surrogate uncertainty (b) Surrogate uncertainty and MCS error

Fig. 4 Uncertainty in failure probability due to surrogate uncertainty and MCS error

3.3 Summary of the proposed methodology

To summarize, the first stage consists of the construction of a limit state surrogate to approximate the true but unknown limit state (component or system). From the model calibration analysis using KOH framework, a general-purpose GP surrogate of the simulation model is available, which is further refined by adding more training points close to the limit state through maximization of the Expected Feasibility Function (EFF), to obtain the limit state surrogate. Along with the original input variables to the simulation model and uncertain model parameters, model discrepancy is also considered as an input by explicitly representing its variability through an auxiliary variable. Monte Carlo sampling-based maximization of the EFF is carried out which involves generating a large pool of samples of the inputs from their corresponding PDFs and selecting a sample, which maximizes the EFF value as a new training point.

In the second stage, reliability analysis is carried out using the GP surrogate through Monte Carlo sampling. The Monte Carlo samples generated during the first stage for EFF maximization are propagated through the GP surrogate for reliability analysis. The prediction from a GP follows a Gaussian distribution, and outputs at different inputs are correlated; this surrogate uncertainty is included in the failure probability estimate through correlated sampling of the model predictions at several inputs. Since correlated sampling of all inputs is computationally expensive, the inputs are divided into two groups depending on the probability that they affect the failure probability estimate; this probability is analyzed by obtaining the coefficient of variation (COV) of model prediction. Only the mean values of model predictions are used if COV ≥ 3.1, whereas the randomness in model prediction is considered for all other inputs. Several realizations of the model predictions are obtained through correlated sampling resulting in several samples of failure probability estimate, one for each realization. After consideration of surrogate uncertainty, the failure probability estimate is represented using a PDF. The consideration of MCS error (due to limited Monte Carlo samples for uncertainty propagation) results in each of the failure probability samples to be represented by a PDF; this results in the failure probability estimate to be represented using a family of PDFs. If necessary, the family of PDFs can be integrated to form an unconditional PDF, which can be reported as the failure probability estimate after considering all sources of epistemic uncertainty.

4 Numerical examples

4.1 A single component with multiple limit states

Consider a short cantilever beam subjected to a point load at its free end as shown in Fig. 5. Two limit states – maximum deflection and maximum stress are considered for the failure

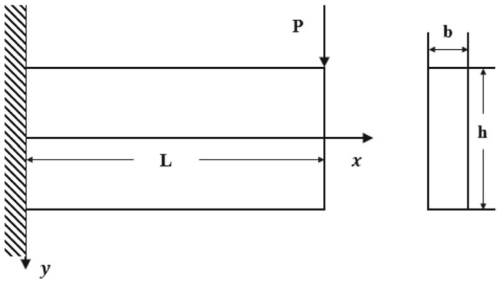


Fig. 5 A cantilever beam with point-load at the free end

of the beam. The goal in this problem is to compute the reliability with respect to each of the individual limit states and the system reliability, considering both the limit states.

Assume that the free end deflection of the beam due to the point load is modeled (according to the Euler-Bernoulli beam theory) as $d(P)=PL^3/3EI$. Here, E and I represent the Young's modulus and moment of inertia respectively. Since the beam is short, there exists a discrepancy between the experimental observations and predicted deflections computed (because the Euler-Bernoulli model is not accurate for short beams). This model discrepancy is calibrated using the KOH framework. A total of 30 points were generated through Latin Hypercube Sampling (LHS) across the domain of inputs, at which both simulation and experimental data are available. Among them 25 points were used for calibration (model parameters and model discrepancy using the KOH framework) and 5 points were used for validation. The maximum COV (coefficient of variation) among the testing points is 0.1134. Therefore, the deflection of the beam, after accounting for model discrepancy $\delta(P)$, is given as

$$d(P) = PL^3/3EI + \delta(P) \tag{30}$$

The expression for the computation of maximum stress is given as

$$s(P) = PLh/2I \tag{31}$$

The load P is assumed to be aleatory with uncertain distribution type and distribution parameters. E and L are aleatory following Gaussian distributions with known distribution parameters. The cross-section parameters (b, h) are also assumed to be aleatory with known parameters due to geometric variations in the manufacturing process. Table 1 shows the random variables

Table 1 Example 1. Variables and their statistics

Parameters		Distribution	Mean	Standard deviation
P	μ_P ($\times 10^5$ N)	Normal	35	0.3
	σ_P ($\times 10^5$ N)	Lognormal	4	0.1
E ($\times 10^9$ N/m^2)		Normal	210	10
L (m)		Normal	2	0.01
b (m)		Normal	0.18	0.005
h (m)		Normal	0.75	0.005

Table 2 Distribution types and their probabilities for Load P

Distribution type	Normal	Type 1 EVD
Probability	0.2	0.8

and their statistics used in this example, and Table 2 gives two candidate distribution types of P and their corresponding probabilities.

The two limit state functions corresponding to deflection and stress are given as

$$g_d(P) = d_0 - d(P) \tag{32}$$

$$g_s(P) = s_0 - s(P) \tag{33}$$

where $g_d(P) < 0$ and $g_s(P) < 0$ indicate failure, and s_0, d_0 are the limiting values of stress and deflection. In this example, the threshold values for stress and deflection are assumed deterministic (d_0 = 0.01 m, s_0 = 500 MPa). Let the models in Eqs. (30, 31) be known as simulation models for deflection and stress respectively. The system fails when either $g_d(P) < 0$ or $g_s(P) < 0$, and the system failure probability is given by

$$p_f^s = \Pr\{g_d(P) \le 0 \cup g_s(P) \le 0\} \tag{34}$$

(a) *Component reliability analysis*

The reliability analysis with respect to each of the individual limit states is first performed using the proposed method. Since the deflection limit state is associated with model discrepancy, the reliability analysis is performed with and without discrepancy, in order to investigate the effectiveness of the proposed method in handling model discrepancy during surrogate construction. Tables 3 and 4 show the reliability analysis results, along with the number of function evaluations (NOF), without considering the surrogate uncertainty and MCS error, in order to compare with the simulation model estimate. In this example, the threshold EFF value for selection of training points in surrogate construction is assumed to be 0.002. The results show that the proposed method can comprehensively estimate the component reliability in the presence of various sources of epistemic uncertainty. The results illustrate that the reliability estimate can be improved by considering the model discrepancy (Table 3). Figure 6 shows the comparison between the simulation model limit state, and the limit-state from surrogate with and without consideration of model discrepancy. Since it is not possible to show the limit state contours with all the random variables, Fig. 6 shows the contours between length and load on the X and Y axes respectively. The other random variables E, b, h are conditioned at their mean values. It shows that the surrogate constructed taking model discrepancy into consideration is closer to the true limit state than the surrogate without considering the model discrepancy. Figure 7 provide the failure probability distribution with respect to each limit state after considering both the surrogate uncertainty and MCS error. The figures illustrate that the proposed method can effectively quantify the uncertainty in the reliability analysis results.

Table 3 Example 1. Component failure probability results (Deflection)

	p_f	NOF	ε (%)
Simulation model	0.0249	1×10^6	–
Limit state surrogate without model discrepancy	0.0057	25 + 0	77.11
Limit state surrogate with model discrepancy	0.0248	25 + 2	0.4

The NOF include the initial number of samples from calibration (25) and the number of added new samples

Table 4 Example 1. Component failure probability results (Stress)

	p_f	NOF	ε (%)
Simulation model	0.0605	1×10^6	–
Limit state surrogate model	0.0633	32	4.63

(b) *System reliability analysis*

A surrogate is constructed considering both the limit states for system reliability analysis using the method discussed in Sec. 3.2. Figure 8 gives the PDF of failure probability after considering both the deflection and stress limit states. It shows that the proposed method can effectively quantify the uncertainty in the system failure probability estimate (Table 5).

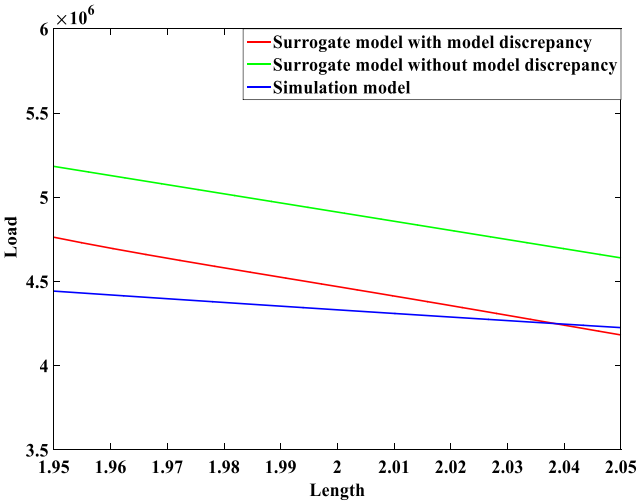


Fig. 6 Comparison of limit states between the simulation model and surrogate, with and without considering model discrepancy

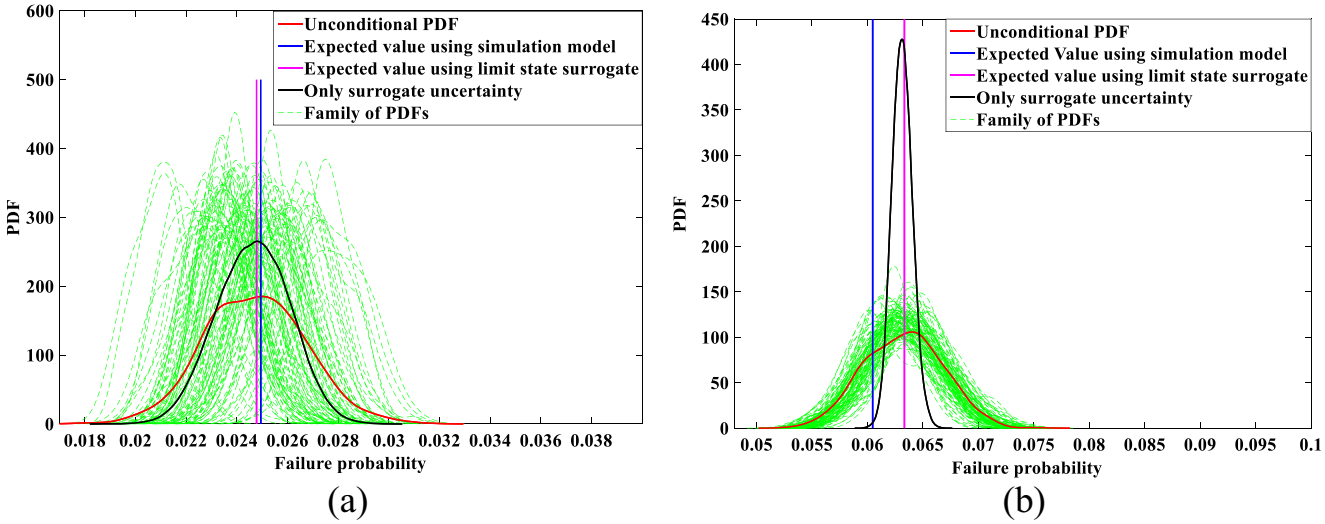


Fig. 7 Failure probability with respect to (**a**) deflection, and (**b**) stress

4.2 A system of multiple components each with a limit state

Example 1 demonstrated the proposed method for component reliability analysis and reliability analysis of a series system. In this example, reliability analysis of a parallel system is demonstrated. Consider the two-bar system shown in Fig. 9. The system consists of two bars supporting a panel on which a load P is applied at its center.

The two bars are assumed to be made of different materials with different failure stress characteristics. The Young's moduli of the materials (E_A, E_B) are assumed to be aleatory variables with known distribution types and distribution parameters. For illustration, the bars are assumed to be of equal length (L = 1 m) and the area of cross-section of the two bars are assumed to be deterministic ($A_A = 0.04$ m^2, $A_B = 0.0625$ m^2). The forces in the bars are given by

$$F_i = P(E_i A_i/(E_A A_A + E_B A_B)), i = A, B \tag{35}$$

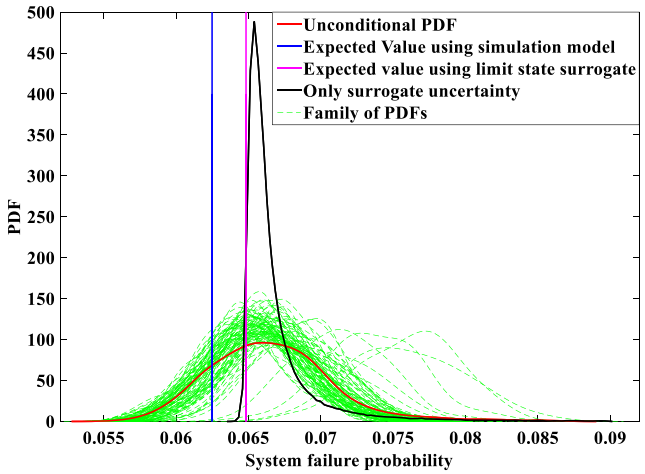


Fig. 8 System failure probability with respect to both deflection and stress limit states

Let σ_A^0 and σ_B^0 represent the failure stresses of the bars; the limit state functions are given by

$$g_i(\mathbf{X}) = A_i\sigma_i^0 - P(E_i A_i/(E_A A_A + E_B A_B)), i = A, B \tag{36}$$

The applied load P is assumed to aleatory, but with uncertain distribution type and uncertain distribution parameters. The failure stress is assumed to be deterministic but not known precisely. Some point and interval data are assumed to be available from material testing. Table 6 provides the list of variables used in this example and their statistics.

Using the available data, non-parametric distributions are constructed for both the threshold stresses using the likelihood approach discussed in Section 2, and spline-based interpolation is employed for the PDF modeling. The failure stress of a material is not calibrated. We assume that data about failure stress is directly available from material testing and we construct a non-parametric distribution to represent the uncertainty. After constructing the PDF, it is treated as any other input random variable for reliability analysis.

For illustration purposes, the same candidate distribution types for load P and their probabilities are used as in Table 2. The system described here is a parallel system and therefore failure occurs when both the bars fail i.e. $g_A(\mathbf{X}) < 0$ and $g_B(\mathbf{X}) < 0$. The system failure probability is given by

$$p_f^s = \Pr\{g_A(\mathbf{X}) < 0 \cap g_B(\mathbf{X}) < 0\} \tag{37}$$

Table 5 Example 1. Failure probability estimates using MCS and proposed method

	p_f^s	NOF	ε (%)
Simulation model	0.0625	1×10^6	–
Limit state surrogate model	0.0648	276	3.68

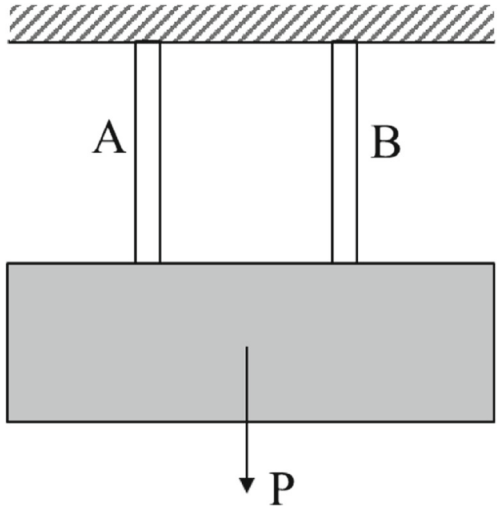


Fig. 9 A two-bar system

Table 7 gives the system failure probability analysis results after consideration of statistical uncertainty, model discrepancy and uncertain model parameters, and without considering surrogate uncertainty and MCS error. Similar to the previous example, the threshold EFF value for selection of training points in surrogate construction is assumed to be 0.002. Figure 10 shows the system failure probability estimates without considering surrogate uncertainty and MCS error, considering only surrogate uncertainty and considering surrogate both uncertainty and MCS error. It can be seen that the surrogate uncertainty and MCS error result in larger uncertainty in the reliability estimate compared to considering only surrogate uncertainty. In order to reduce the uncertainty in the estimate, the surrogate needs to be further refined by adding more training points and the number of Monte Carlo samples need to be increased to reduce the MCS error.

Table 7 System Failure probability estimates using simulation model and limit state surrogate

	p_f^s	NOF	Error (%)
Simulation model (Euler model + model discrepancy)	0.0048	6×10^5	–
Limit state surrogate model	0.0049	302	2.08

5 Conclusion

This paper proposed a unified framework connecting the model calibration analysis to the construction of a limit state surrogate and estimating the uncertainty in reliability estimate due to the incorporation of different types of epistemic uncertainty. Epistemic uncertainty due to data (statistical uncertainty) and model (model parameter uncertainty, model discrepancy, surrogate uncertainty, Monte Carlo sampling error) are considered. A parametric approach is proposed for quantification of input variables with statistical uncertainty. Non-parametric distributions are used to quantify uncertain model parameters with statistical uncertainty. First, model calibration is carried out using the KOH framework, whose output include GP surrogates for simulation model and model discrepancy. The general purpose (global) GP surrogate of simulation model is further refined by adding more training points close to the limit state (local refinement) to obtain the limit state surrogate for reliability analysis; this is referred to as a hybrid approach (local refinement of the general purpose GP surrogate). An auxiliary variable is used to represent the model discrepancy, which enables the construction of a limit state surrogate that includes the model discrepancy. The inputs for the limit state surrogate are the original input variables along with the uncertain model parameters and the model discrepancy. The selection of training points is carried out using a learning function

Table 6 Example 2. Variables and their statistics

Variable		Distribution	Mean	Standard deviation
Load P	μ_P ($\times 10^6$ N)	Normal	19	0.3
	σ_P ($\times 10^6$ N)	Lognormal	1.6	0.1
Young's Modulus E_A ($\times 10^9$ Pa)		Normal	210	20
Young's Modulus E_B ($\times 10^9$ Pa)		Normal	180	15
Failure stress of bar A, σ_A^0 ($\times 10^7$ Pa)		Point data – [24.7, 24.95,25.1,25.3] Interval data – [(24.6, 24.62), (24.8, 24.84), (25.1, 25.15)]		
Failure stress of bar B, σ_B^0 ($\times 10^7$ Pa)		Point data – [21.8, 21.92,22.08,22.15] Interval data – [(21.85, 21.88), (22.06, 22.08), (22.13, 22.16)]		

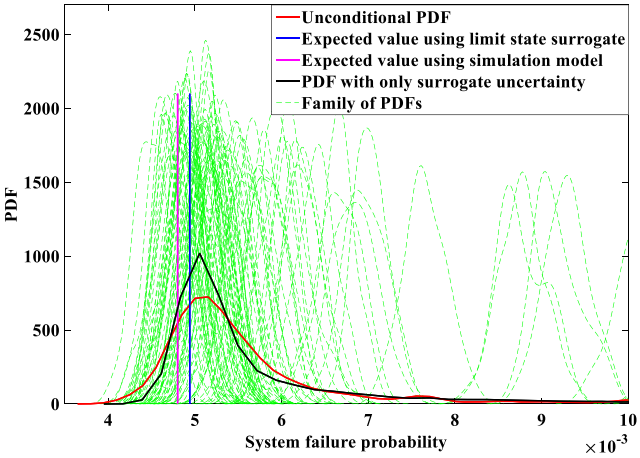


Fig. 10 System failure probability estimate using simulation model and limit state surrogate

called the Expected Feasibility Function (EFF). The GP model is then used to carry out reliability analysis using Monte Carlo sampling. A single-loop sampling approach using an auxiliary variable is used for the sampling of input variables with statistical uncertainty and samples of the inputs are generated and passed through the surrogate to obtain the reliability estimate. Note that this results in a single value of the failure probability.

The model prediction using a GP model is a Gaussian distribution with the parameters dependent on the input. In addition, the model predictions at different inputs are correlated due to the covariance function in the GP model. To account for the surrogate uncertainty (i.e., the variability in the prediction), correlated sampling of model predictions at several inputs is carried out and used for reliability analysis. Inclusion of surrogate uncertainty results in a probability distribution (PDF) for the failure probability. When MCS error is incorporated, each sample in the failure probability distribution is represented by a PDF, which results in a family of PDFs for the failure probability.

Two examples demonstrated the effectiveness of the proposed method. The proposed method is able to not only address heterogeneous sources of epistemic uncertainty during system reliability analysis, but also provide the uncertainty associated with the reliability estimate.

This works considered only time-independent reliability analysis. Future work should consider time-dependent reliability analysis. Quantities with spatial variability (and associated epistemic uncertainty) which further complicate the problem also need to be considered.

Ultimately, the purpose of uncertainty quantification is to inform decision making, such as resource allocation or design optimization. The inclusion of epistemic uncertainty in reliability-based design optimization (RBDO) is still a topic of active research. The framework developed in this paper provides a systematic approach to quantifying and including several sources of epistemic uncertainty within reliability analysis, and therefore facilitates their subsequent inclusion in design optimization. Future work may also consider the quantification of relative contributions of different sources of uncertainty on the reliability estimate, which will facilitate resource allocation for epistemic uncertainty reduction (e.g., additional data collection vs model refinement).

Acknowledgments The research reported in this paper was supported by funds from the Air Force Office of Scientific Research (Grant No. FA9550-15-1-0018, Technical Monitor: Dr. David Stargel). The support is gratefully acknowledged. The authors also thank the reviewers for their helpful comments.

References

Agresti A, Coull BA (1998) Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat 52(2):119–126

Arendt PD, Apley DW, Chen W (2012) Quantification of model uncertainty: Calibration, model discrepancy, and identifiability. J Mech Des 134(10):100908

Bernardo JM, Rueda R (2002) Bayesian hypothesis testing: a reference approach. Int Stat Rev 70(3):351–372

Bichon BJ, Eldred MS, Swiler LP, Mahadevan S, McFarland JM (2008) Efficient global reliability analysis for nonlinear implicit performance functions. AIAA J 46(10):2459–2468

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

Choi SK, Grandhi RV, Canfield RA (2004) Structural reliability under non-Gaussian stochastic behavior. Comput Struct 82(13):1113–1121

Der Kiureghian A, Ditlevsen O (2009) Aleatory or epistemic? Does it matter? Struct Saf 31(2):105–112

Du X (2010) System reliability analysis with saddlepoint approximation. Struct Multidiscip Optim 42(2):193–208

Dubourg V, Sudret B, Deheeger F (2013) Metamodel-based importance sampling for structural reliability analysis. Probab Eng Mech 33:47–57

Echard B, Gayton N, Lemaire M (2011) AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. Struct Saf 33(2):145–154

Emura T, Lin YS (2015) A comparison of normal approximation rules for attribute control charts. Qual Reliab Eng Int 31(3):411–418

Faravelli L (1989) Response-surface approach for reliability analysis. J Eng Mech 115(12):2763–2781

Haldar A, Mahadevan S (2000a) Reliability assessment using stochastic finite element analysis. John Wiley & Sons, New York

Haldar A, Mahadevan S (2000b) Probability, reliability, and statistical methods in engineering design. John Wiley & Sons, New York

Hoeting JA, Madigan D, Raftery AE, Volinsky CT (1999) Bayesian model averaging: a tutorial. Stat Sci 14(4):382–401

Hu Z, Du X (2015) Mixed efficient global optimization for time-dependent reliability analysis. J Mech Des 137(5):051401

Hu Z, Mahadevan S, (2015) Global sensitivity analysis-enhanced surrogate (GSAS) modeling for reliability analysis. Struct Multidiscip Optim :1–21

Hu Z, Mahadevan S, Du X (2015) Uncertainty quantification in time-dependent reliability analysis method. The ASME 2015 international design engineering technical conferences (IDETC) and computers and information in engineering conference (CIE), August 2–7, 2015 in Boston, MA

Jiang Z, Chen W, Fu Y, Yang R (2013) Reliability-based design optimization with model bias and data uncertainty. SAE Int J Mater Manuf 6(3):502–516

Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. J Glob Optim 13(4):455–492

Kaymaz I (2005) Application of kriging method to structural reliability problems. Struct Saf 27(2):133–151

Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models. J R Stat Soc Ser B Stat Methodol 63(3):425–464

Ling Y, Mahadevan S (2013) Quantitative model validation techniques: New insights. Reliab Eng Syst Saf 111:217–231

Mahadevan S, Zhang R, Smith N (2001) Bayesian networks for system reliability reassessment. Struct Saf 23(3):231–251

Nannapaneni S, Mahadevan S (2015) Model and data uncertainty effects on reliability estimation. 17th AIAA non-deterministic approaches conference. AIAA SciTech, Florida

Papadrakakis M, Papadopoulos V, Lagaros ND (1996) Structural reliability analysis of elastic–plastic structures using neural networks and Monte Carlo simulation. Comput Methods Appl Mech Eng 136(1): 145–163

Sankararaman S, Mahadevan S (2011) Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data. Reliab Eng Syst Saf 96(7):814–824

Sankararaman S, Mahadevan S (2013a) Separating the contributions of variability and parameter uncertainty in probability distributions. Reliab Eng Syst Saf 112:187–199

Sankararaman S, Mahadevan S (2013b) Distribution type uncertainty due to sparse and imprecise data. Mech Syst Signal Process 37(1):182–198

Song J, Der Kiureghian A (2003) Bounds on system reliability by linear programming. J Eng Mech 129(6):627–636

Song H, Choi KK, Lee I, Zhao L, Lamb D (2013) Adaptive virtual support vector machine for reliability analysis of high-dimensional problems. Struct Multidiscip Optim 47(4):479–491

Straub D, Der Kiureghian A (2010) Bayesian network enhanced with structural reliability methods: methodology. J Eng Mech 136(10): 1248–1258

Wang Z, Wang P (2015) A double-loop adaptive sampling approach for sensitivity-free dynamic reliability analysis. Reliab Eng Syst Saf 142:346–356

Wang P, Youn BD, Xi Z, Kloess A (2009) Bayesian reliability analysis with evolving, insufficient, and subjective data sets. J Mech Des 131(11):111008

Wang P, Hu C, Youn BD (2011) A generalized complementary intersection method (GCIM) for system reliability analysis. J Mech Des 133(7):071003

Youn BD, Choi KK (2004) A new response surface methodology for reliability-based design optimization. Comput Struct 82(2):241–256

Youn BD, Wang P (2006) Bayesian reliability based design optimization under both aleatory and epistemic uncertainties, 11th AIAA/ISSMO multidisciplinary analysis and optimization conference, 6–8 September. Portsmouth, Virginia