Accelerating subset simulation with surrogate model ensuring unbiased failure probability estimation

OPEN ACCESS
IOP Publishing | ISECAEP
Journal of Reliability Science and Engineering
J. Reliab. Sci. Eng. **1** (2025) 045201 (19pp)
https://doi.org/10.1088/3050-2454/ae202a

Accelerating subset simulation with surrogate model ensuring unbiased failure probability estimation

Yuanzhuo Ma[1,2], Chuang Li[2] and Binbin Li[3,*]

[1] School of Renewable Energy, Hohai University, Changzhou, People's Republic of China
[2] School of Electrical and Power Engineering, Hohai University, Nanjing, People's Republic of China
[3] State Key Laboratory of Biobased Transportation Fuel Technology, ZJU-UIUC Institute, Zhejiang University, Haining, People's Republic of China

E-mail: binbinli@intl.zju.edu.cn

Received 23 July 2025, revised 17 October 2025
Accepted for publication 17 November 2025
Published 8 December 2025

Abstract

Surrogate models can accelerate the failure probability estimation, but at the risk of biasedness when the surrogate does not capture the real failure surface, which is typical for real engineering problems with high dimensionality and/or high nonlinearity. Following the idea of delayed acceptance (DA) in Markov Chain Monte Carlo (MCMC), a new strategy of combining subset simulation (SuS) and surrogate model is developed in this paper, named as DA-SuS. It decomposes the acceptance process in MCMC into three steps, in which the candidate samples are first checked by the surrogate model. If rejected by the surrogate model, the sample is no more considered. This DA does not destroy the detailed balance of MCMC, i.e., the stationary distribution will always be the targeted one, no matter how bad the surrogate model is. Consequently, the asymptotic unbiasedness of MCMC estimator is preserved. Although the statistical efficiency deteriorates slightly, the DA brings computational savings because only those candidates with high probability in the failure domain are eventually evaluated by the true model, thus increasing the overall computational efficiency. Three improvement strategies are further introduced, including adaptive training of Kriging model, misjudgment error-guided acceptance and chain-wise updating scheme. The performance of DA-SuS algorithm is demonstrated via three benchmark examples, and compared with conventional SuS and its variants equipped with surrogate models. The proposed DA-SuS algorithm yields an unbiased estimation of the failure probability, even in the case with high dimensionality and nonlinearity, although its statistical and computational efficiency depend on the quality of the surrogate.

Keywords: subset simulation, surrogate model, unbiasedness, delayed-acceptance, Kriging

* Author to whom any correspondence should be addressed.

1. Introduction

For a specific failure mode in structural reliability estimation, the corresponding failure probability p_{f} can be evaluated as [1]

$$p_{\mathrm{f}}=\int_{\Omega}\pi\left(\boldsymbol{x}\right)\mathrm{d}\boldsymbol{x}=\int_{\Re^{n}}\pi\left(\boldsymbol{x}\right)I_{\Omega}\left(\boldsymbol{x}\right)\mathrm{d}\boldsymbol{x} \tag{1}$$

where $X=[X_1,X_2,\cdots,X_n]^{\mathrm{T}}\in\Re^{n}$ is an n-dimensional input random vector. $\pi\left(\boldsymbol{x}\right)$ is the joint probability density function (PDF) of $\boldsymbol{X}$. The failure domain is represented as $\Omega=\left\{x\in\Re^{n}\left|g\left(\boldsymbol{x}\right)\leqslant 0\right.\right\}$, where $g\left(\boldsymbol{x}\right)$ is the limit state function (LSF) representing the failure mode. $I_{\Omega}\left(\boldsymbol{x}\right)$ is an indicator function, which is equal to 1 if $\boldsymbol{x}\in\Omega$ and 0 otherwise. Note that $\boldsymbol{x}$ is a realization of random vector $\boldsymbol{X}$. Without loss of generality, the integral in equation (1) is more convenient to work in the standard normal space by transforming the random vector $\boldsymbol{X}$ into the standard normal random vector $\boldsymbol{U}$ using Rosenblatt transformation [2] or Nataf transformation [3]. More specifically, the failure probability can be expressed in the standard normal space as

$$p_{\mathrm{f}}=\int_{F}\varphi_{n}\left(\boldsymbol{u}\right)\mathrm{d}\boldsymbol{u}=\int_{\Re^{n}}\varphi_{n}\left(\boldsymbol{u}\right)I_{F}\left(\boldsymbol{u}\right)\mathrm{d}\boldsymbol{u} \tag{2}$$

where $\varphi\left(\cdot\right)$ is the PDF of standard normal random variable. The transformed LSF $g\left(\boldsymbol{u}\right)=g\left(T^{-1}\left(\boldsymbol{u}\right)\right)$ defines the failure domain $F=\left\{\boldsymbol{u}\in\Re^{n}\left|g\left(\boldsymbol{u}\right)\leq 0\right.\right\}$ in the standard normal space.

Conventional direct numerical integration is infeasible for solving this integral due to the typically large dimension or complex failure domain in practical problems. Analytical approximate methods, such as the first order reliability method [4] and the second order reliability method [5], are incapable of solving the nonlinear and high dimensional problem as well because they are the linear or quadratic approximation of the LSF. The direct Monte Carlo simulation (MCS) [6] can theoretically yield an unbiased estimate of the failure probability because of the law of large numbers, while huge computational burden is usually required to achieve a desired precision. To address this, variance reduction-based simulation techniques such as important sampling (IS) [7], directional simulation (DS) [8–10], line sampling (LS) [11, 12] and subset simulation (SuS) [13] were developed. Among them, SuS is most widely used due to its good performance in high dimensional problems, which has been extended to multiple failure modes [14], Bayesian inference [15, 16], stochastic optimization [17–21], and sensitivity analysis [22, 23]. To further improve its sampling efficiency, several Markov Chain Monte Carlo (MCMC) sampling strategies have been introduced, e.g. the adaptive conditional sampling (aCS) [24], Hamiltonian Monte Carlo [25], elliptical slice sampling [26], asymptotic SuS [27] and a 'relaxed' version of SuS (Re-SuS) [28]. The up-to-date work [29] also provided an adaptive scheme with an optimal setting of correlation parameters in aCS for effectively suppressing the systematic growth of candidate rejection and correlation along Markov chains. A Bayesian perspective on the failure probability integral estimation was proposed as well to quantify, propagate and reduce numerical uncertainty behind the failure probability due to discretization error [30]. Another branch of methods were deduced from the law of probability conservation, such as the probability density evolution method [31] and the direct probability integral method [32]. These methods are applicable for both static and dynamic reliability problems, while the computation is still expensive for many engineering applications.

To further reduce the computational burden in repeated evaluation of LSF, surrogate models have been developed for failure probability estimation, such as response surface method [33], polynomial chaos expansion [34], Kriging model [35–40], support vector machine [41, 42], Bayesian networks [43–45] and convolutional neural network [46]. Among them, Kriging is widely used, since it not only provides the mean estimate of the LSF but also the mean squared error of the estimation. The adaptive learning schemes were introduced to form the so called adaptive Kriging (AK)-based methods, such as AK-MCS [47], AK-IS [41, 48–50], AK-DS [51], AK-LS [52], AK-distance-based subdomain [53], AK-improved weighted sampling [54], and AK-SuS [39, 42, 55–59]. Several typical learning functions were developed, such as U-learning [47], expected feasibility function [60], H-learning [61], the least improvement function [62], the confidence interval squeezing function [63], and the expected system improvement function [64] (to name a few). Beyond a single LSF, the AK-based failure probability estimation has been extended to multiple responses [65, 66]. No matter how adaptive schemes have enhanced efficiency and accuracy, the surrogate models inevitably introduce estimation errors, inducing potential bias in the estimation of failure probabilities. In addition, delicate schemes of surrogate models are not able to depict the real failure surface in real engineering problems with high dimension ($>$20) or high complexity.

This paper proposes a new way to accelerate SuS with surrogate models that ensures the unbiasedness of the estimation, even if the surrogate model does not fit the LSF well. That is, we adopt the Kriging model to screen the candidate samples first in MCMC, and only evaluate the value of LSF at candidate samples accepted by the Kriging model. By doing this, the calculation of the real LSF at rejected samples can be avoided without introducing bias in the failure probability estimation. Since the real LSF at all accepted candidate samples in MCMC are still calculated, an unbiased estimation is guaranteed, including those problems with high dimension ($>$20) and/or high complexity. The proposed method is based on the strategy of delayed acceptance [67] (DA) in MCMC sampling, so that we name our method as DA-SuS. To enhance its computational efficiency, three improvement strategies are further proposed. The first improvement (denoted as DA-SuS-1) lies in the adaptive training of the Kriging model in each simulation level, to increase the prediction accuracy and thus reduce the unnecessary LSF evaluation. The second strategy (named as DA-SuS-2) further reduces the number of LSF evaluation by directly accepting candidate samples with small

misjudgment error in the Kriging model, sacrificing accuracy for efficiency. Besides the above accepting strategy, the continuous update of Kriging model (named as DA-SuS-3) is adopted once a Markov chain is completed to further enhance the computational efficiency.

This paper is organized into five main sections. A brief review of SuS and Kriging model are provided in section 2. Section 3 gives the general principle of DA-SuS. Three improvement strategies are further provided in section 4. Finally, three empirical studies are used in section 5 to demonstrate the performance of the proposed method.

2. Backgrounds and related techniques

A brief review of SuS and Kriging model are given in this section. It provides the readers the necessary background to understand the proposed methods.

2.1. Brief introduction of SuS

SuS is an efficient algorithm to estimate the failure probability by converting a small failure probability into a product of a sequence of relatively large conditional probabilities by introducing intermediate failure events adaptively. Based on the conditional probability rule, one can have

$$p_{\mathrm{f}} = \Pr(F) = \Pr\left(\bigcap\nolimits_{i=1}^{m} F_i\right) = \Pr(F_1)\prod\nolimits_{i=2}^{m}\Pr(F_i|F_{i-1}) \tag{3}$$

where $F_i = \{\boldsymbol{u} \in \Re^n | g(\boldsymbol{u}) \le l_i\}$ is the intermediate failure domain satisfying $F_1 \supset F_2 \supset \cdots \supset F_m \supset F$ and l_i is the intermediate failure threshold satisfying $l_1 > l_2 > \cdots > l_m = 0$. $\Pr(F_i|F_{i-1})$ is the conditional probability or the so-called 'level probability'. The intermediate sampling density (ISD) is chosen as the truncated multivariate normal distribution

$$\varphi_n(\boldsymbol{u}|F_{i-1}) = \frac{\varphi_n(u) I_{F_i}(u)}{\Pr(F_i)}. \tag{4}$$

The first level of SuS is a direct MC, which yields the first failure domain $F_1 = \{\boldsymbol{u} \in \Re^n | g(\boldsymbol{u}) \le l_1\}$. Given one realization $\left\{\boldsymbol{u}^0_{(1)}, \boldsymbol{u}^0_{(2)}, \cdots, \boldsymbol{u}^0_{(N)}\right\}$ of N independent and identically distributed (i.i.d.) samples $\left\{\boldsymbol{U}^0_{(1)}, \boldsymbol{U}^0_{(2)}, \cdots, \boldsymbol{U}^0_{(N)}\right\}$ from the standard normal distribution $\varphi_n(\boldsymbol{u})$, they are sorted in ascending order according to their LSF values, denoted as $\left\{\boldsymbol{u}^0_{(1)}, \boldsymbol{u}^0_{(2)}, \cdots, \boldsymbol{u}^0_{(N)}\right\}$. The first threshold is then determined as $l_1 = g\left(\boldsymbol{u}^0_{(N_{\mathrm{c}})}\right)$, where $N_{\mathrm{c}} = N_{\mathrm{p0}}$ is selected to be an integer. The estimator of the corresponding level probability is

$$\Pr(F_1) \leftarrow \hat{p}_1 = \frac{1}{N}\sum_{k=1}^{N} I_{\Omega}\left(U^0_{(k)}\right). \tag{5}$$

Thereafter, regarding $\left\{\boldsymbol{u}^{i-1}_{(1)}, \boldsymbol{u}^{i-1}_{(2)}, \cdots, \boldsymbol{u}^{i-1}_{(N_{\mathrm{c}})}\right\}$ as N_{c} 'seeds'. Generate $N_{\mathrm{s}} = (N\text{–}N_{\mathrm{c}})/N_{\mathrm{c}}$ correlated random samples $\left\{\boldsymbol{u}^i_{(k,1)}, \boldsymbol{u}^i_{(k,2)}, \cdots, \boldsymbol{u}^i_{(k,N_{\mathrm{s}})}\right\}$ from each kth seed $\boldsymbol{u}^{i-1}_{(k)}$ $(k = 1, \ldots, N_{\mathrm{c}})$ with the target distribution $\varphi_n(\boldsymbol{u}|F_i)$. It can totally generate a of size N–N_c samples in the $i+1$-th $(i = 1,\ldots, m-1)$ simulation level so that the sample size of each level keeps to be a constant N. Sort again the realization $\left\{\left\{u^{i-1}_{(1)}, u^{i-1}_{(2)}, \cdots, u^{i-1}_{(N_{\mathrm{c}})}\right\}, \left\{u^i_{(1,1)}, u^i_{(1,2)}, \cdots, u^i_{(1,N_{\mathrm{s}})}\right\}, \cdots, \left\{u^i_{(N_{\mathrm{c}},1)}, u^i_{(N_{\mathrm{c}},2)}, \cdots, u^i_{(N_{\mathrm{c}},N_{\mathrm{s}})}\right\}\right\}$ in ascending order as $\left\{\boldsymbol{u}^i_{(1)}, \boldsymbol{u}^i_{(2)}, \cdots, \boldsymbol{u}^i_{(N)}\right\}$ according to their LSF values, threshold $l_{i+1} = g\left(u^i_{(N_{\mathrm{c}})}\right)$ of the $i+1$-th intermediate failure domain $F_{i+1} = \{\boldsymbol{u} \in \Re^n | g(\boldsymbol{u}) \le l_{i+1}\}$ is obtained. The MCMC estimator for the level probability then reads

$$\Pr(F_i|F_{i-1}) \leftarrow \hat{p}_i = \frac{1}{N}\sum_{j=1}^{N_{\mathrm{c}}}\sum_{k=1}^{N_{\mathrm{s}}} I_{F_i}\left(U^{i-1}_{(j,k)}\right), i = 2, \cdots, m. \tag{6}$$

Collecting all the level probability estimators, it finally yields the SuS estimator of the failure probability as

$$p_{\mathrm{f}} = \Pr(F) \leftarrow \hat{p}_{\mathrm{f}} = \prod_{i=1}^{m}\hat{p}_i. \tag{7}$$

For completeness, the pseudocode of the SuS algorithm is provided in table 1 as a reference in subsequent sections.

2.2. Basics of Kriging model

The predicted value of Kriging can be expressed as a sum of a deterministic regression model $R_e(\boldsymbol{\beta}, \mathbf{x})$ and a random process $Z(\mathbf{x})$, i.e.,

$$G(\mathbf{x}) = R_e(\boldsymbol{\beta}, \mathbf{x}) + Z(\mathbf{x}) = \mathbf{h}(\mathbf{x})\boldsymbol{\beta} + Z(\mathbf{x}), \tag{8}$$

where $\mathbf{h}(\mathbf{x})$ is the basis function, determining the regression form. The correlation coefficient $\boldsymbol{\beta}$ is a regression parameter. The random process $Z(\mathbf{x})$ is assumed to be normally distributed with zero-mean and covariance matrix as

$$\mathrm{COV}(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) = {\sigma_Z}^2 R(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta}) \tag{9}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ represent the value of test samples, and ${\sigma_Z}^2$ is the variance parameter. The versatility of Kriging model is determined by the correlation function $R(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta})$ between $\mathbf{x}_i$ and $\mathbf{x}_j$. For simplicity, the squared exponential function is adopted in this paper as

$$R(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta}) = \exp\left(-\sum_{k=1}^{n}\theta_k\left(x_{j(k)} - x_{i(k)}\right)^2\right) \tag{10}$$

where n is the dimension of random variables. Here, $x_{i(k)}$, $x_{j(k)}$ and θ_k are the kth component of $\mathbf{x}_i$, $\mathbf{x}_j$, and $\boldsymbol{\theta}$, respectively.

Table 1. Pseudocode of SuS.

Algorithm 1. SuS.

1.*Input*
Level probability $p_0 = 0.1$, number of samples per level N.
2.*Change of Variable*
Rosenblatt or Nataf transformation $\boldsymbol{R} = \boldsymbol{T}(\boldsymbol{u})$. *% $\boldsymbol{R}$ is standard multivariate normal RV*
3.*Direct MC simulation*
Generate initial set of samples $\left\{\boldsymbol{u}_{(1)}^0, \boldsymbol{u}_{(2)}^0, \cdots, \boldsymbol{u}_{(N)}^0\right\}$, calculate the real responses $\boldsymbol{g}$.
Sort $\left\{\boldsymbol{u}_{(1)}^0, \boldsymbol{u}_{(2)}^0, \cdots, \boldsymbol{u}_{(N)}^0\right\}$ in ascending order according to $\boldsymbol{g}$ and select the top Np_0 samples as seed samples $\left\{\boldsymbol{u}_{(1)}^0, \boldsymbol{u}_{(2)}^0, \cdots, \boldsymbol{u}_{(N_c)}^0\right\}$, determine threshold $l_1 = g(\boldsymbol{u}_{(N_c)}^0)$, $\widehat{p}_1 = p_0$.
4. *Initialization*
Level index $i = 1$and stopping criterion *stopflag* $= 0$.
5. *Iterations*
While *stopflag* $= 0$
(1) Run component-wise Metropolis-Hastings with seed samples $\left\{\boldsymbol{u}_{(1)}^{i-1}, \boldsymbol{u}_{(2)}^{i-1}, \cdots, \boldsymbol{u}_{(N_c)}^{i-1}\right\}$. Accept or reject candidate samples:
For $k = 1$ to N_c *% N_c is the number of seed samples*
% N_s is the number of candidate samples generated from each seed sample
For $j = 1$ to N_s
Generate the candidate sample $\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}$ from seed sample $\boldsymbol{u}_{(k)}^{i-1}$.
Calculate the real response $g(\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}})$ based on $\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}$.
If $g(v_{(k,j)}^{\boldsymbol{i}}) < l_i$
Accept the candidate sample $\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}$, $\boldsymbol{u}_{(k,j)}^{\boldsymbol{i}} = \boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}$, $g_{(k,j)}^{\boldsymbol{i}} = g(\boldsymbol{v}_{(k,j)}^{i})$.
else
Reject the candidate sample $v_{(k,j)}^{\boldsymbol{i}}$, $\boldsymbol{u}_{(k,j)}^{\boldsymbol{i}} = \boldsymbol{u}_{(k)}^{i-1}$, $g_{(k,j)}^{\boldsymbol{i}} = g_{(k)}^{i-1}$.
End
Update seed sample to current state sample $\boldsymbol{u}_{(k)}^{i-1} = \boldsymbol{u}_{(k,j)}^{i}$.
End
End
(2) Sort $\left\{\left\{\boldsymbol{u}_{(1)}^{i-1}, \boldsymbol{u}_{(2)}^{i-1}, \cdots, \boldsymbol{u}_{(N_c)}^{i-1}\right\}, \left\{\boldsymbol{u}_{(1,1)}^{i}, \boldsymbol{u}_{(1,2)}^{i}, \cdots, u_{(1,N_s)}^{i}\right\}, \cdots, \left\{\boldsymbol{u}_{(N_c,1)}^{i}, u_{(N_c,2)}^{i}, \cdots, u_{(N_c,N_s)}^{i}\right\}\right\}$ in ascending order as $\left\{\boldsymbol{u}_{(1)}^{i}, \boldsymbol{u}_{(2)}^{i}, \cdots, \boldsymbol{u}_{(N)}^{i}\right\}$ according to $\boldsymbol{g}$ and update threshold $l_{i+1} = g(\boldsymbol{u}_{(N_c)}^{i})$.
(3) Update stopping criterion *stopflag* and estimate the level probability $\widehat{p}_{i+1}$:
If $l_{i+1} < 0$
stopflag $= 1$.
$p_{i+1} = (\text{size}(\text{find}(\boldsymbol{g} < 0), 2)/N)$.
else
stopflag $= 0$.
$p_{i+1} = p_0$.
Select the top N_c samples as seed samples.
End
$i = i+1$.
End
6. *Output*
Calculate the estimated failure probability $p_f = \prod_{i=1}^{m} p_i$.

Defining $\mathbf{R}$ as the correlation matrix of t samples, its (i, j)-the element can be calculated as

$$R_{ij} = R(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta}),\ i,j = 1,\ldots,t. \tag{11}$$

The unknown parameter θ can be obtained by the maximum likelihood estimation method as

$$\boldsymbol{\theta}^* = \arg\min_{\theta}\left\{|\mathbf{R}|^{\frac{1}{t}}\sigma_Z{}^2\right\}. \tag{12}$$

When θ is determined, β and $\sigma_Z{}^2$ can then be calculated as

$$\boldsymbol{\beta}^* = \left(\boldsymbol{R}_e{}^{\mathrm{T}}\mathbf{R}^{-1}\boldsymbol{R}_e\right)^{-1}\boldsymbol{R}_e{}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Y} \tag{13}$$

$$\sigma_Z{}^2 = \frac{1}{t}(\mathbf{Y} - \boldsymbol{R}_e\boldsymbol{\beta}^*)^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{Y} - \boldsymbol{R}_e\boldsymbol{\beta}^*) \tag{14}$$

where $\mathbf{Y}$ and $\mathbf{R}_e$ are the real response set and regression function value of the training sample set, respectively. Taking advantage of the fact that the conditional distribution in a multivariate normal distribution is still normal, the predicted response of the Kriging model satisfies $G(\mathbf{x}) \sim$

$N\left(\mu_{\hat{G}}(\mathbf{x}), {\sigma_{\hat{G}}}^2(\mathbf{x})\right)$ at an arbitrary testing point $\mathbf{x}$, where the mean and variance are given as:

$$\mu_{\hat{G}}(\mathbf{x}) = \boldsymbol{h}(\mathbf{x})^{\mathrm{T}}\boldsymbol{\beta}^* + \mathbf{r}(\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{Y} - \boldsymbol{R}_e\boldsymbol{\beta}^*\right) \tag{15}$$

$$\sigma_{\hat{G}}{}^2(\mathbf{x}) = \sigma_Z{}^2\left(1 + \mathbf{u}^{\mathrm{T}}\left(\boldsymbol{R}_e{}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\mathbf{u} - \mathbf{r}(\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r}(\mathbf{x})\right). \tag{16}$$

Here, $\mathbf{u} = \boldsymbol{R}_e{}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r}(\mathbf{x}) - \boldsymbol{h}(\mathbf{x})$ and $\mathbf{r}$ is the correlation coefficient matrix between the testing point and the training sample set.

3. SuS with DA

In this section, we propose a new way of combing SuS with surrogate model, which ensures the unbiasedness of the failure probability estimation, even if the surrogate model does not really capture the failure surface. We name it as DA-SuS, because it is based on the strategy of DA in the MCMC sampling [67]. The idea behind DA-SuS is to decompose the acceptance ratio, when sampling from the intermediate ISD, into a product of sub-ratios and then sequentially compare them with independent uniform variates to stop earlier once one term is below the corresponding uniform. Although various surrogate models can be integrated into this framework, this paper takes the most widely used Kriging model as an example.

3.1. Main algorithm of DA-SuS

The basic procedure of DA-SuS is the same as that in section 2.1, except that the MCMC sampling is revised to combine the Kriging model. Note that the proposed DA strategy can be implemented for most MCMC sampling algorithms with acceptance-rejection scheme, e.g. component-wise Metropolis–Hastings (CW-MH) sampling and aCS [24]. For simplicity, we explain the general idea with the standard MH algorithm in this paper.

Consider the standard MH algorithm for sampling from the ISD $\varphi_n(\boldsymbol{u}|F_i)$ using the transition PDF $p(\boldsymbol{v}|\boldsymbol{u})$ to generate a new sample $\boldsymbol{v}$ with

$$p(\boldsymbol{v}|\boldsymbol{u}) = a(\boldsymbol{u},\boldsymbol{v})\,q(\boldsymbol{v}|\boldsymbol{u}) + (1 - r(\boldsymbol{u}))\,\delta_u(\boldsymbol{v}) \tag{17}$$

where $q(\boldsymbol{v}|\boldsymbol{u})$ is the proposal PDF and $r(u) = \int a(\boldsymbol{u},\boldsymbol{v})\,q(\boldsymbol{v}|\boldsymbol{u})\,\mathrm{d}\boldsymbol{v}$. Here, $\delta_u(\cdot)$ is the Dirac delta function at $\mathbf{u}$, and $a(\boldsymbol{u}, \boldsymbol{v})$ is the acceptance probability with the following form

$$a(\boldsymbol{u},\boldsymbol{v}) = \min\{1, \kappa(\boldsymbol{u},\boldsymbol{v})\} = \min\left\{1, \frac{\varphi_n(\boldsymbol{v})\,q(u|\boldsymbol{v})\,I_{F_i}(\boldsymbol{v})}{\varphi_n(\boldsymbol{u})\,q(\boldsymbol{v}|\boldsymbol{u})\,I_{F_i}(\boldsymbol{u})}\right\} \tag{18}$$

where $\kappa(\boldsymbol{u},\boldsymbol{v})$ is called the acceptance ratio. The acceptance probability $a(\boldsymbol{u},\boldsymbol{v})$ is compared with a unit(0, 1) variate to decide whether or not the Markov chain switches from the current value $\boldsymbol{u}$ to the candidate value $\boldsymbol{v}$.

Applying the DA strategy, one can first decompose the acceptance ratio $\kappa(\boldsymbol{u}, \boldsymbol{v})$ as a triple product

$$\begin{aligned}\kappa(\boldsymbol{u},\boldsymbol{v}) &= \kappa_1(\boldsymbol{u},\boldsymbol{v})\,\kappa_2(\boldsymbol{u},\boldsymbol{v})\,\kappa_3(\boldsymbol{u},\boldsymbol{v}) \\ &= \frac{\varphi_n(\boldsymbol{v})\,q(\boldsymbol{u}|\boldsymbol{v})}{\varphi_n(u)\,q(\boldsymbol{v}|\boldsymbol{u})}\frac{\Pr_{\mathrm{K}}(\boldsymbol{v}\in F_i)}{\Pr_{\mathrm{K}}(\boldsymbol{u}\in F_i)}\frac{I_{F_i}(\boldsymbol{v})/\Pr_{\mathrm{K}}(\boldsymbol{v}\in F_i)}{I_{F_i}(\boldsymbol{u})/\Pr_{\mathrm{K}}(\boldsymbol{u}\in F_i)}\end{aligned} \tag{19}$$

where $\Pr_{\mathrm{K}}(\boldsymbol{v}\in F_i)$ represents the probability of $\boldsymbol{v}$ belonging to the intermediate failure domain F_i predicted by the Kriging model. Since the LSF $G(\boldsymbol{v})$ is approximated by a Kriging model, $\Pr_{\mathrm{K}}(\boldsymbol{v}\in F_i)$ becomes a natural surrogate of the indicator function $I_{F_i}(\boldsymbol{v})$. Given the conditional distribution $G(\boldsymbol{v}) \sim N(\mu(\boldsymbol{v}), \sigma^2(v))$, one has

$$\Pr_{\mathrm{K}}(\boldsymbol{v}\in F_i) = \Pr_{\mathrm{K}}(G(\boldsymbol{v}) \le y_i) = \Phi\left(\frac{y_i - \mu(\boldsymbol{v})}{\sigma(\boldsymbol{v})}\right) \tag{20}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal variate.

Each sub-ratio $\kappa_i(\boldsymbol{u}, \boldsymbol{v})$ in equation (19) is nonnegative and satisfies the balance condition $\kappa_i(\boldsymbol{u}, \boldsymbol{v}) = \kappa_i(\boldsymbol{u}, \boldsymbol{v})^{-1}$ if one defines $1/0 = \infty$ for mathematical completeness, because the indicator function $I_{F_i}(\boldsymbol{v})$ may be equal to zero. Correspondingly, one can accept the transition from $\boldsymbol{u}$ to $\boldsymbol{v}$ with the probability

$$\begin{aligned}\tilde{a}(\boldsymbol{u},\boldsymbol{v}) &= \tilde{a}_1(\boldsymbol{u},\boldsymbol{v})\,\tilde{a}_2(\boldsymbol{u},\boldsymbol{v})\,\tilde{a}_3(\boldsymbol{u},\boldsymbol{v}) \\ &= \min\{1, \kappa_1(\boldsymbol{u},\boldsymbol{v})\}\min\{1, \kappa_2(\boldsymbol{u},\boldsymbol{v})\}\min\{1, \kappa_3(\boldsymbol{u},\boldsymbol{v})\}\end{aligned} \tag{21}$$

i.e. by successively comparing random uniform variates u_i to the terms $\tilde{a}_i(u, v)$ and the first rejection signaling that the proposed value should not be considered any further.

To verify the validity of the DA strategy, i.e. the ISD $\varphi_n(u|F_i)$ is the stationary distribution of the resulting Markov chain, the detailed balance condition is proved as follows:

$$\begin{aligned}&\varphi_n(u|F_i)\,q(\boldsymbol{v}|\boldsymbol{u})\,\tilde{a}(\boldsymbol{u},\boldsymbol{v}) \\ &\quad= \varphi_n(u|F_i)\,q(\boldsymbol{v}|\boldsymbol{u})\min\{1, \kappa_1(\boldsymbol{u},\boldsymbol{v})\} \\ &\qquad\times \min\{1, \kappa_2(\boldsymbol{u},\boldsymbol{v})\}\min\{1, \kappa_3(\boldsymbol{u},\boldsymbol{v})\} \\ &\quad= \varphi_n(u|F_i)\,q(\boldsymbol{v}|\boldsymbol{u})\,\kappa_1(\boldsymbol{u},\boldsymbol{v})\min\{1, \kappa_1(v,u)\}\,\kappa_2(\boldsymbol{u},\boldsymbol{v}) \\ &\qquad\times \min\{1, \kappa_2(\boldsymbol{v},\boldsymbol{u})\} \\ &\qquad\times \kappa_3(\boldsymbol{u},\boldsymbol{v})\min\{1, \kappa_3(\boldsymbol{v},\boldsymbol{u})\} \\ &\quad= \varphi_n(\boldsymbol{u}|F_i)\,q(\boldsymbol{v}|\boldsymbol{u})\,\kappa(\boldsymbol{u},\boldsymbol{v})\,\tilde{a}(\boldsymbol{v},\boldsymbol{u}) \\ &\quad= \varphi_n(\boldsymbol{v}|F_i)\,q(\boldsymbol{u}|\boldsymbol{v})\,\tilde{a}(\boldsymbol{v},\boldsymbol{u})\end{aligned} \tag{22}$$

where, we have used the property $\min\{1, x\} = x\min\{1, 1/x\}$ for $x \geqslant 0$ in the second equation, and in the last equation one can verify that $\varphi_n(\boldsymbol{u}|F_i)\,q(\boldsymbol{v}|\boldsymbol{u})\,\kappa(\boldsymbol{u},\boldsymbol{v}) = \varphi_n(\boldsymbol{v}|F_i)\,q(\boldsymbol{u}|\boldsymbol{v})$. Equation (22) shows that the DA strategy satisfies the detailed balance, so that it can converge to the same stationary distribution as the standard M–H algorithm.

One implementation of the proposed DA-SuS strategy is provided in table 2 for the CW-MHs sampling. We can see that the real calculation of LSFs of all rejected samples is replaced with Kriging model, avoiding the time-consuming evaluation of true LSF.

Table 2. Pseudocode of DA-SuS.

Algorithm 2. DA-SuS.

%Candidate samples are accepted/rejected based on predicted responses. Those %accepted then undergo real calculation for further validation of acceptance
For $k = 1$ to N_c
  For $j = 1$ to N_s
    For $h = 1$ to n *% n is the dimension of v (Take CW M-H as an example)*
      Select a one-dimensional proposal PDF $q(v|u)$ that meets the requirement of symmetry, and simulate the h-th component of the candidate sample ξ_h from $q_h\left(\cdot|\boldsymbol{u}_{(k)}^{i-1}(h)\right)$.
      Calculate the ratio $\kappa_h = q_h(\xi_h)/q_h\left(\boldsymbol{u}_{(k)}^{i-1}(h)\right)$.
      Accept or reject ξ_h at the candidate sample as:
$$\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}(h) = \begin{cases} \xi_h, & \min\{1,\kappa_h\} > \text{random}[0,1] \\ u_{(k)}^{i-1}(h), & \min\{1,\kappa_h\} \leqslant \text{random}[0,1] \end{cases}$$
    End *% Conditional sampling can be used as alternatives*
    Calculate the predicted response $\hat{g}(\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}})$ based on $\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}$.
    If $\hat{g}(\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}) < l_i$
      Calculate the real response $g(\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}})$ based on $\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}$.
      Accept or reject the candidate sample $v\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}$ based on $g(\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}})$, similar to SuS.
    else
      Reject the candidate sample $\boldsymbol{v}_{(k,j)}^{\boldsymbol{i}}$, $\boldsymbol{u}_{(k,j)}^{\boldsymbol{i}} = \boldsymbol{u}_{(k)}^{i-1}$, $g_{(k,j)}^{i} = g_{(k)}^{i-1}$.
    End
    Update seed sample to current state sample $\boldsymbol{u}_{(k)}^{i-1} = \boldsymbol{u}_{(k,j)}^{\boldsymbol{i}}$.
  End
End

The drawback of the DA strategy is that the acceptance probability $\tilde{a}(\boldsymbol{u}, \boldsymbol{v})$ is always smaller than $a(\boldsymbol{u}, \boldsymbol{v})$ of the standard M–H algorithm

$$\begin{aligned}\tilde{a}(\boldsymbol{u},\boldsymbol{v}) &= \min\{1,\kappa_1(\boldsymbol{u},\boldsymbol{v})\}\min\{1,\kappa_2(\boldsymbol{u},\boldsymbol{v})\}\min\{1,\kappa_3(\boldsymbol{u},\boldsymbol{v})\} \\ &\leq \min\{1,\kappa_1(\boldsymbol{u},\boldsymbol{v})\kappa_2(\boldsymbol{u},\boldsymbol{v})\kappa_3(\boldsymbol{u},\boldsymbol{v})\} = a(\boldsymbol{u},\boldsymbol{v}) \end{aligned} \tag{23}$$

since $\min\{1, x\}\min\{1, y\} \leqslant \min\{1, xy\}$ for $x, y \geqslant 0$. Consequently, one can conclude that the asymptotic variance of sample averages calculated using the DA strategy is greater than or equal to that calculated using the standard M–H algorithm based on the Peskun ordering [68]. Therefore, in general, the DA strategy will be less statistically efficient than the standard M-H algorithm (20), (23).

The statistical efficiency of the DA-SuS algorithm depends on the quality of the Kriging surrogate. Consider the ideal case where the Kriging model perfectly matches the LSF with negligible uncertainty. In that case, all samples not in the failure domain F_i are rejected by the Kriging model and $\tilde{a}_3(\boldsymbol{u}, \boldsymbol{v}) = 1$ for all preconditioned samples. As a consequence, the speedup of the DA-SuS algorithm over the standard M–H algorithm is the inverse of the average acceptance rate and the statistical efficiency is not sacrificed. It follows that the DA-SuS algorithm is more beneficial when the rejection rate is high. However, incorrect classification of samples by Kriging would deteriorate the performance of the DA-SuS algorithm. More specifically, the false negative error leads to lower statistical efficiency because the Kriging model rejects the sample lying in the failure domain; meanwhile, the false positive error incurs unnecessary evaluation of the LSF. Reasoning of the misjudgment error of Kriging model in DA-SuS can be referred to the appendix, and its impact will be discussed in the following illustrative examples.

Unlike the usual way of using the surrogate model in an approximation method, the DA-SuS algorithm yields an unbiased estimator of the failure probability because the DA-SuS algorithm naturally belongs to simulation methods. The error introduced by Kriging only affects the efficiency of the estimator but not its unbiasedness.

3.2. Illustrative example

To demonstrate the unbiasedness of the DA-SuS algorithm, an illustrative example is considered in this section by comparing with SuS and AK-SuS. In order to analyze the impact of the quality of the surrogate model on the acceptance/rejection rate and misjudgment error in DA-SuS, the AK-MCS (using U-learning and error-based stop criterion (ESC) [69]) is applied to construct the Kriging models with different error conditions. To be specific, the initial Kriging models with error of 0.1 and 0.01 are considered. The number of initial training samples N_{initial} is set as 3. Number of both test samples N_{test} and augmented test samples N_{au} are set to be 10^4.

Table 3. Results of reliability estimation by AK-MCS.

	Error	p_f	N_{call}
MCS		2.23×10^{-3}	1×10^6
AK-MCS-initial	1	NA	3
AK-MCS	0.1	2.02×10^{-3}	$41(3 + 38)$
AK-MCS	0.01	2.16×10^{-3}	$57(3 + 54)$

Note: 'N_{call}' denotes the size of calculations of real LSF. '3 + 38' in Row 4, Column 4 represents 3 initial training samples +38 adaptive enrichment samples.

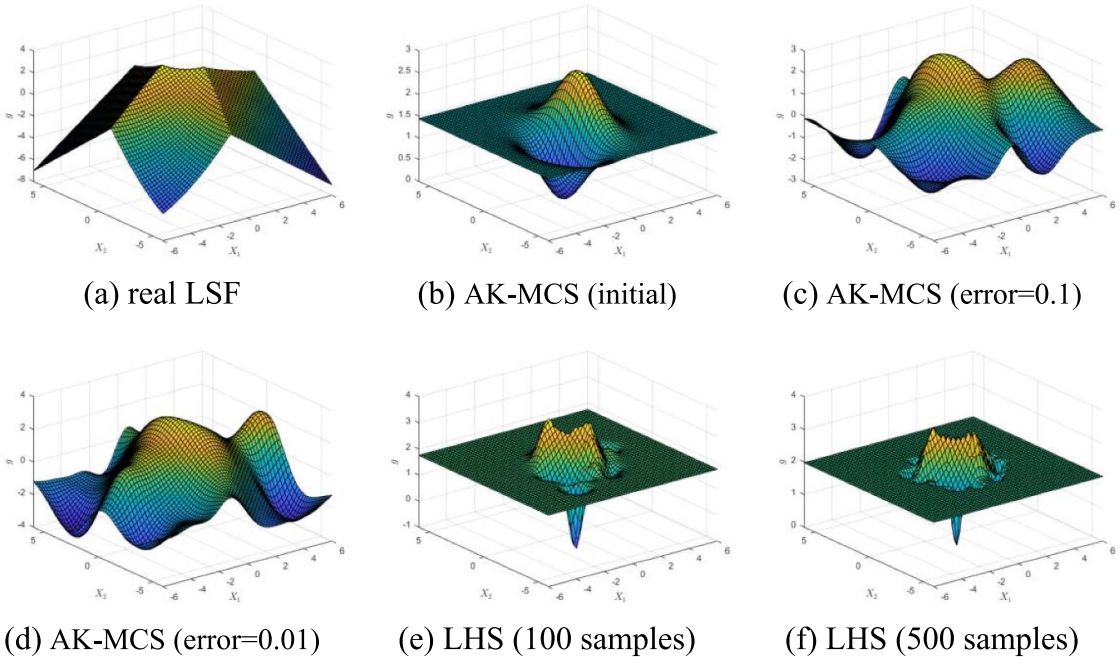


Figure 1. Comparison diagrams of approximate LSFs and real LSF.

We consider a benchmark example for reliability estimation, i.e. four-boundary series system model [47] with LSF expressed as

$$\begin{cases} g_1 = -(X_1 + X_2)\Big/\sqrt{2} + 0.1(X_1 - X_2)^2 + 3, \\ g_2 = X_1 - X_2 + 7\Big/\sqrt{2}, \\ g_3 = (X_1 + X_2)\Big/\sqrt{2} + 0.1(X_1 - X_2)^2 + 3, \\ g_4 = X_2 - X_1 + 7\Big/\sqrt{2}, \\ g = \min(g_1, g_2, g_3, g_4) \end{cases} \tag{24}$$

where X_1 and X_2 are independent standard normal random variables. The results in table 3 show that the number of required training samples increases as the model accuracy improves (error decreases), and the estimated failure probabilities become increasingly closer to that by MCS. Note that the initial Kriging model did not yield the failure probability, due to its relatively low accuracy, denoted as 'NA' in the table.

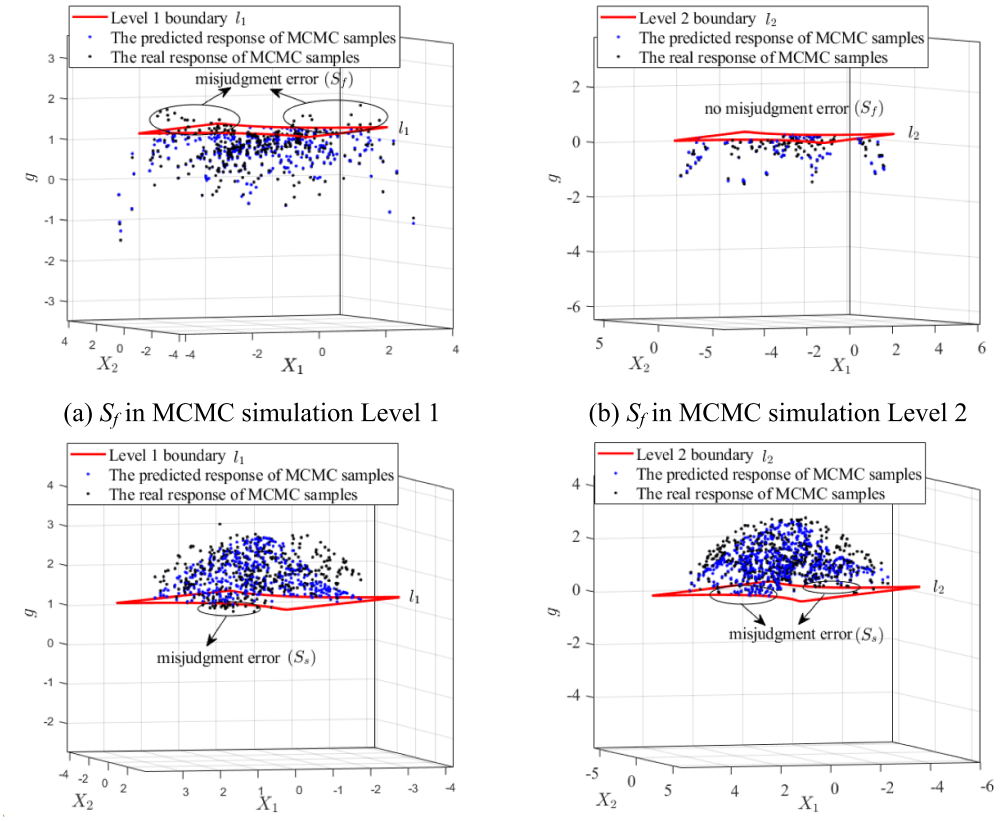
In order to better reflect the quality of the surrogate models constructed using the AK-MCS, Latin hypercube sampling (LHS) was used to construct two direct Kriging models (100 and 500 training samples) as a reference. The approximated LSFs of all surrogate models were then compared with the real LSF, as shown in figure 1. It can be observed that the approximated LSFs by the surrogate models constructed using the LHS present a significant discrepancy from the real LSF in the failure domain. In contrast, the surrogate models built using AK-MCS achieved higher accuracy. Besides, the approximated LSF becomes closer to the real LSF as the error controlled by ESC decreases. Although there is also a certain difference between the approximated LSF of the initial Kriging model and the real LSF (it explains again why no feasible estimator can be obtained by the initial model), it was retained to compare with the other surrogate models obtained using AK-MCS, which helps illustrate the impact of the quality of the surrogate on the performance of the SS-DA algorithm.

The rejection rate and misjudgment error at each MCMC simulation level are presented in table 4. For the convenience of replicating the same results, the seed in the pseudo-random number generator in MATLAB is fixed as 'seed = 15; rand('twister', seed); randn('state', seed + 1)'. Note that it is

Table 4. Results of reliability estimation for illustrating example.

	Error	p_f ($\times 10^{-3}$)	N_{call}		MCMC simulation level 1	MCMC simulation level 2	Deviation
SuS		2.57 ± 0.45	2800	Rejection rate	0.518	0.656	0
AK-SuS	Initial	NA	3	$\widehat{S}_s$	73	123	NA
				S_s	4	19	
				$\widehat{S}_f$	219	211	
				S_f	447	341	
				Rejection rate (surrogate model)	0.199	0.290	
DA-SuS	Initial	1.89 ± 1.02	1756	$\widehat{S}_s$	205	138	26.5%
				S_s	239	54	
				$\widehat{S}_f$	72	133	
				S_f	101	125	
				Rejection rate (surrogate model)	0.669	0.496	
				Rejection rate (real function)	0.339	0.275	
AK-SuS	0.1	2.01 ± 0.51	41 (3 + 38)	$\widehat{S}_s$	33	8	27.5%
				S_s	20	3	
				$\widehat{S}_f$	29	5	
				S_f	89	4	
				Rejection rate (surrogate model)	0.512	0.694	
DA-SuS	0.1	2.30 ± 0.49	1794	$\widehat{S}_s$	42	11	18.0%
				S_s	39	17	
				$\widehat{S}_f$	24	5	
				S_f	63	2	
				Rejection rate (surrogate model)	0.484	0.679	
				Rejection rate (real function)	0.136	0.007	
AK-SuS	0.01	2.32 ± 0.58	57 (3 + 54)	$\widehat{S}_s$	25	2	10.2%
				S_s	10	1	
				$\widehat{S}_f$	29	6	
				S_f	99	0	
				Rejection rate (surrogate model)	0.494	0.642	
DA-SuS	0.01	2.62 ± 0.67	1823	$\widehat{S}_s$	34	9	1.1%
				S_s	23	13	
				$\widehat{S}_f$	24	5	
				S_f	60	0	
				Rejection rate (surrogate model)	0.470	0.679	
				Rejection rate (real function)	0.126	0	

Note: 'N_{call}' denotes the size of calculations of real LSF. '3 + 38' represents 3 initial training samples + 38 adaptive enrichment samples. Feasible estimator cannot be obtained by the initial Kriging model, due to its relatively low accuracy, which is denoted as 'NA'. 'Deviation' denotes the percentage deviation of the results of all algorithms relative to the SuS result (with SuS result as the reference).



(a) S_f in MCMC simulation Level 1

(b) S_f in MCMC simulation Level 2

(c) S_s in MCMC simulation Level 1

(d) S_s in MCMC simulation Level 2

Figure 2. The misjudgment errors during each MCMC simulation in DA-SuS.

unnecessary to fix the seed of pseudo-random number generator in real practice.

The impact of the misjudgment error of Kriging model in DA-SuS is illustrated in this example as well, details of which can be found in the appendix. To be specific, $\hat{S}_s$ and $\hat{S}_f$ in table 4 represent the maximum errors obtained from the upper bound calculated using equation (42) in the appendix, while S_s and S_f refer to the true misjudgment error at each MCMC simulation level. As the accuracy of the Kriging model improves, both $\hat{S}_s$ and $\hat{S}_f$, as well as S_s and S_f, show an overall decreasing trend. However, there are errors between $\hat{S}_s$ and S_s, as well as $\hat{S}_f$ and S_f, due to the error between the real probability of sum of dependent Bernoulli trials for each MCMC simulation and the Poisson approximation mentioned in the appendix

For S_f obtained by the AK-SuS and DA-SuS algorithms based on the same accuracy of surrogate model, S_f in DA-SuS is smaller than that in AK-SuS, resulting in a lower misjudgment error for the accepted part of the surrogate model in DA-SuS and reducing unnecessary evaluation of the LSF. S_s represents the misjudgment error for the rejected part of the surrogate model, which leads to a reduction in statistical efficiency due to the erroneous rejection of samples located within the failure domain. During the implementation of the AK-SuS and DA-SuS algorithms, each approximated LSF becomes closer to the real LSF as it approached to the failure boundary when MCMC simulation progressed, which leads to a reduction in both S_s and S_f.

As a reference, samples corresponding to the misjudgment error at each MCMC simulation level in the DA-SuS algorithm are shown in figure 2. It can be observed that the misjudgment error in MCMC simulation level 2 is lower than that in level 1 because each approximated LSF at the level 2 boundary (l_2) is closer to the real LSF than at the level 1 boundary (l_1). Meanwhile, the improvement in model accuracy allows DA-SuS to achieve results closer to SuS compared to AK-SuS. When the error is 0.01, the result of DA-SuS is the closest to SuS, showing nearly no bias. Although DA-SuS requires more samples compared to AK-SuS, it can provide an unbiased estimator corresponding to that in conventional SuS meanwhile save a proportion of calculation (the rejected parts in MCMC). It does provide a new perspective of an unbiased surrogated-based method for conventional SuS and its variant with any type as well. Furthermore, AK-SuS fails to yield valid results when the accuracy of surrogate models is extremely low (initial Kriging with 3 samples), while DA-SuS could still produce a solution close to that of SuS despite certain deviation. This phenomenon demonstrates that the introduction of the DA strategy provides a solution for scenarios where the accuracy of the surrogate model cannot be guaranteed, and the corresponding statistical characteristics will be analyzed

Table 5. Pseudocode of enrichment strategy in DA-SuS-1.

Algorithm 3. Enrichment strategy in DA-SuS-1.

% Update Kriging model before accepting or rejecting the candidate samples
If $i = 1$
  Use the initial sample set $\boldsymbol{u}^0$ that has already calculated the real responses to update Kriging model.
else
  Use all samples with real responses in each simulation level to update Kriging model.
End

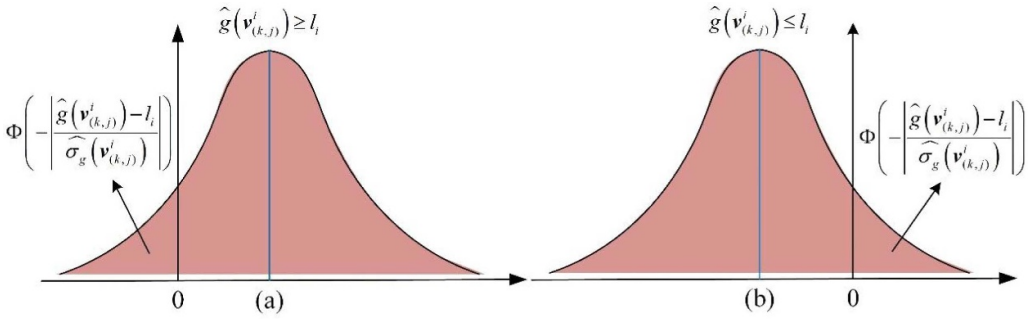


Figure 3. An illustration of the probability of wrong estimation in equation (38) considering $\hat{g}\left(\boldsymbol{v}^i_{(k,j)}\right) \geq l_i$ (a) or $\hat{g}\left(\boldsymbol{v}^i_{(k,j)}\right) \leq l_i$ (b).

in detail in section 5. Kriging model from AK-MCS with error equals to 0.01 is then adopted in section 5 as well.

4. Improvement strategies

To enhance the statistical and computational efficiency of the proposed DA-SuS, three improvement strategies are incorporated in this section. For brevity, the proposed DA-SuS in section 3 is noted as the DA-SuS-0.

4.1. Adaptive training of Kriging model

The first improvement (denoted as DA-SuS-1) is to adaptively update the Kriging model using all samples with real responses in each simulation level. We provide the corresponding pseudocode in table 5 for completeness, although it should be self-evident.

4.2. Misjudgment error-guided acceptance

The second improvement comes from the early stopping in the sample generation. It is guided from the calculated misjudgment error of DA-SuS as shown in the appendix). The real LSF will not be calculated in DA if the misjudgment error is considered to be small. We denote the improved algorithm with early stopping as DA-SuS-2 for simplicity.

Take figure 3 as an example. For the jth candidate sample $\boldsymbol{v}^i_{(k,j)}$ generated from the k-th seed sample $\boldsymbol{u}^{i-1}_{(k)}$ in the $i+1$-th simulation level with the threshold l_i of the corresponding intermediate failure event, the probability of wrong estimation of the predictor $\hat{g}\left(\boldsymbol{v}^i_{(k,j)}\right)$ from Kriging in equation (38) considering $\hat{g}\left(\boldsymbol{v}^i_{(k,j)}\right) \geq l_i$ or $\hat{g}\left(\boldsymbol{v}^i_{(k,j)}\right) \leq l_i$ can be quantified by

$$\Phi\left(U^i_{(k,j)}\right) = \Phi\left(\left|\frac{\hat{g}\left(\boldsymbol{v}^i_{(k,j)}\right) - l_i}{\hat{\sigma}_g\left(\boldsymbol{v}^i_{(k,j)}\right)}\right|\right) \tag{25}$$

According to U-learning in the conventional AK-MCS [47], the misjudgment error of DA-SuS when accepting/rejecting candidate samples in each MCMC simulation is controlled by $U>2$ for each candidate sample, which corresponds to the misjudgment error is lower than $1-\Phi(2) \approx 0.023$. If the predictor $\hat{g}\left(\boldsymbol{v}^i_{(k,j)}\right)$ and $\hat{\sigma}_g\left(\boldsymbol{v}^i_{(k,j)}\right)$ of each accepted candidate $\boldsymbol{v}^i_{(k,j)}$ by Kriging satisfies

$$U^i_{(k,j)} = \left|\frac{\hat{g}\left(\boldsymbol{v}^i_{(k,j)}\right) - l_i}{\hat{\sigma}_g\left(\boldsymbol{v}^i_{(k,j)}\right)}\right| > 2 \tag{26}$$

it can be directly accepted, rather than further determining whether it is accepted or not based on its real response of the real LSF. This thereby avoids the calculation of the real LSF for candidate samples that satisfy a scaled misjudgment error and saves part of the computational burden of DA-SuS-1. The corresponding pseudocode is provided in table 6.

4.3. Chain-wise updating of Kriging model

To further enhance the computational efficiency, an AK model updating in each Markov Chain is adopted, and denote this

Table 6. Pseudocode of DA-SuS-2.

Algorithm 4. DA-SuS-2.

% For each candidate sample, first use a scaled misjudgment error (i.e. $U > 2$)
% to decide whether it is necessary to run a DA process
For $k = 1$ to N_c
  For $j = 1$ to N_s
  Generate the candidate sample $\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}$ from seed sample $u^{i-1}_{(k)}$.
  Calculate the predicted response and standard deviation $\left[\hat{g}\left(\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}\right), \widehat{\sigma_g}\left(\boldsymbol{v}^{i}_{(k,j)}\right)\right]$ based on $\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}$.
  If $\hat{g}(\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}) < l_i$
    Calculate $\boldsymbol{U}^{\boldsymbol{i}}_{(k,j)}$ based on the new U-learning scheme, $\boldsymbol{U}^{\boldsymbol{i}}_{(k,j)} = \left|\frac{\hat{g}\left(\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}\right) - l_i}{\widehat{\sigma_g}\left(\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}\right)}\right|$.
    If $\boldsymbol{U}^{\boldsymbol{i}}_{(k,j)} > 2$
      Accept the candidate sample $\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}$, $\boldsymbol{U}^{i}_{(k,j)} = \boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}$, $g^{i}_{(k,j)} = \hat{g}(v^{i}_{(k,j)})$.
    else
      Similar to DA-SuS, after calculating the real response of $\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}$, accept or reject the candidate sample $\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}$.
    End
  else
      Reject the candidate sample $\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}$, $\boldsymbol{u}^{\boldsymbol{i}}_{(k,j)} = \boldsymbol{u}^{i-1}_{(k)}$, $g^{\boldsymbol{i}}_{(k,j)} = g^{i-1}_{(k)}$.
    End
    Update seed sample to current state sample $\boldsymbol{U}^{i-1}_{(k)} = \boldsymbol{U}^{\boldsymbol{i}}_{(k,j)}$.
  End
End
% The scheme for updating Kriging model is consistent with that in DA-SuS-1.

Table 7. Pseudocode of the chain-wise updating scheme in DA-SuS-3.

Algorithm 4. Chain-wise updating in DA-SuS-3.

% Update Kriging model in each Markov Chain corresponds to each seed sample
If $i = 1$
  Consistent with DA-SuS-1, update Kriging model using the initial sample set X, where the real responses have already been calculated.
else
  For $k = 1$ to N_c *% in k-th Markov Chain*
    For $j = 1$ to N_s
      Generate the candidate sample $\boldsymbol{v}^{\boldsymbol{i}}_{(k,j)}$ from seed sample $\boldsymbol{u}^{i-1}_{(k)}$.
      Accept or reject the candidate sample the same as DA-SuS-2.
    End
    Update Kriging model using all the candidate samples with real responses in the *k-th* Markov Chain.
  End
End

revision as DA-SuS-3. In DA-SuS-1 and DA-SuS-2, Kriging model is updated once a simulation level is finished. Although it improves the quality of learned Kriging model, it falls behind the sampling process, thus does not provide much local information to guide the exploration of samples. A more elaborate scheme is proposed by refining the Kriging model on the fly once a Markov Chain is completed. Together with the misjudgment error-guided acceptance, it can further reduce the computational burden. The corresponding pseudocode is provided in table 7.

5. Empirical studies

In this section, a high-dimensional linear example, a highly nonlinear one, a high-dimensional nonlinear one and an engineering application are selected to further demonstrate the performance of the proposed DA-SuS algorithm and its three improvement strategies. Meanwhile, SuS, AK-SuS, AK-MCS and AK-IS are implemented as well to provide a contrast with the results of DA-SuS and its three improvement strategies. The Gaussian function is selected as the correlation function

for all examples. The linear regression function is used for each Kriging model in all examples. During each implementation of all algorithms, size of samples N and the conditional probability p_0 are set as 10^3 and 0.1.

5.1. Numerical examples

Example A: *High-dimensional linear LSF*
The LSF $g(\boldsymbol{X})$ shown in equation (27) represents an n-dimensional hyperplane. It has also been studied in [64].

$$g(\boldsymbol{X}) = \beta n^{-1/2} - \sum_{i=1}^{n} X_i \tag{27}$$

where X_i $(i = 1, 2 \cdots n)$ are independent standard normal distributed variables; β is set to be 3, corresponding to a failure probability $p_f = 1.35 \times 10^{-3}$, irrespective of the number of random variables. Since the LSF employed is linear, the analytical solution can be readily obtained by FORM. In this paper, n is set as 20.

Example B: *Highly nonlinear LSF*
This example comprises a two-degree-of-freedom primary-secondary system featuring an indeterminately damped oscillator subjected to white noise excitation, as elaborated in [65]. This problem is marked by an eight-dimensional LSF and highly nonlinear in each dimension. The LSF $g(\boldsymbol{X})$ is expressed as

$$g(\boldsymbol{X}) = F_s - 3k_s \times \sqrt{\frac{\pi S_0}{4\xi_s \omega_s^3}\left[\frac{\xi_a \xi_s}{\xi_p \xi_s (4\xi_a^2 + \theta^2) + \gamma \xi_a^2}\frac{\left(\xi_p \omega_p^3 + \xi_s \omega_s^3\right)\omega_p}{4\xi_a \omega_a^4}\right]} \tag{28}$$

where $\omega_p = \sqrt{k_p/m_p}$, $\omega_s = \sqrt{k_s/m_s}$, $\omega_a = \frac{\omega_p + \omega_s}{2}$, $\xi_a = \frac{\xi_p + \xi_s}{2}$, $\gamma = \frac{m_s}{m_p}$, $\theta = (\omega_p - \omega_s)/\omega_a$. All random parameters are assumed to follow a lognormal distribution, with the mean and c.o.v. shown in table 8.

Example C: *High-dimensional nonlinear LSF*
The final example also encompasses a high-dimensional nonlinear LSF, extracted from [66] as

$$g(\boldsymbol{X}) = \frac{2}{n-1}\sum_{i=1}^{n-1} X_i^2 - X_n \tag{29}$$

where X_1, X_2, …, $X_{n\text{-}1}$ are independent standard normal random variables, and X_n is normally distributed with mean of 1 and standard derivation of 0.2. The dimension n is set to be 101, 150 and 200, respectively, corresponding to the failure probability of 1×10^{-3}, 2.99×10^{-4} and 1.23×10^{-4}. In this paper, n is set as 101.

Table 8. Distribution of random parameters.

Variable	m_p	m_s	k_p	k_s	ζ_p	ζ_s	F_s	S_0
Mean	1.5	0.01	1	0.01	0.05	0.02	27.5	100
c.o.v.	0.1	0.1	0.2	0.2	0.4	0.5	0.1	0.1

AK-MCS mentioned in section 3.2 is used to construct surrogate models with an error of 0.01 for the three examples. In examples A, the number of initial training samples N_{initial} is set as 21, while both the test samples N_{test} and the augmented test samples N_{au} are set as 10^5. 62 more samples are then generated to enrich the Kriging model. That is 83 samples (21 + 62) for example A in table 9. Due to the high dimension or highly nonlinearity in example B and example C, more samples are used to construct Kriging by AK-MCS. Those are 160 (10 + 150) and 302 (102 + 200) for Example B and Example C, respectively.

For a thorough comparison considering underlying randomness, both algorithms are run repeatedly for 100 times to provide the corresponding statistics, which can be referred to table 9. To be specific, the estimated failure probabilities of these algorithms are represented by the sample mean ± sample standard deviation, with the deviation calculated from 100 independent runs. The results of MCS are also listed for reference, along with the reliability evaluation results of the surrogate model constructed using the AK-MCS method.

AK-MCS with ESC (an error of 0.01) is capable of giving a high quality of Kriging model for Example A, which can be inferred from that result of AK-MCS is close to that of MCS, so as that of AK-SuS to that of SuS. However, it still remains visible biases compared with the original method for AK-MCS and AK-SuS. Similarly, the result of AK-IS for Example A also exhibits a visible bias when compared to the MCS result. On the contrary, the comparison between the results of SuS and DA-based methods reveals that DA-based methods can yield relatively unbiased results. To be specific, result of the original DA-SuS gets closer to that of SuS than that of AK-SuS, and compared with AK-MCS and AK-IS, result of the original DA-SuS is closer to the reference value (the result of MCS). By enriching Kriging using all samples with real LSFs, the statistical efficiency of the result of DA-SuS-1 is improved. This not only reduces the sample computation load but also results in no bias to the SuS result of Example A. Compared with DA-SuS-1, DA-SuS-2 and DA-SuS-3 further enhance the computational efficiency, with a further reduction in sample computation while almost no sacrifice of statistical efficiency. Indeed, 1083 of samples of real response in DA-SuS-2 and DA-SuS-3 include (83 training sample from AK-MCS and 1000 MC samples in the first simulation level), which indicates (U>2) in the first MCMC is already satisfied and no DA is introduced. That is why DA-SuS-2 and DA-SuS-3 saves more samples (real responses calculated in DA) than

Table 9. Statistical results of reliability estimation.

Example		A(20)	B(8)	C(101)
MCS	p_f	1.38×10^{-3}	3.60×10^{-7}	1.11×10^{-3}
	N_{call}	1×10^5	1×10^8	1×10^5
AK-MCS	p_f	1.61×10^{-3}	1.98×10^{-6}	2.00×10^{-4}
	N_{call}	83 (21 + 62)	160 (10 + 150)	302 (102 + 200)
AK-IS	p_f	$(1.30 \pm 0.42) \times 10^{-3}$	$(2.74 \pm 1.69) \times 10^{-6}$	$(1.75 \pm 0.82) \times 10^{-4}$
	N_{call}	83	160	302
SuS	p_f	$(1.37 \pm 0.36) \times 10^{-3}$	$(3.84 \pm 2.98) \times 10^{-7}$	$(1.26 \pm 0.34) \times 10^{-3}$
	N_{call}	2944	6463	2989
AK-SuS	p_f	$(1.29 \pm 0.34) \times 10^{-3}$	$(1.84 \pm 1.52) \times 10^{-6}$	$(1.03 \pm 1.36) \times 10^{-7}$
	N_{call}	83	160	302
	Deviation	5.83%	379.16%	99.99%
DA-SuS-0	p_f	$(1.34 \pm 0.33) \times 10^{-3}$	$(5.03 \pm 3.74) \times 10^{-7}$	$(1.06 \pm 0.75) \times 10^{-3}$
	N_{call}	2229	3673	1337
	Deviation	2.19%	30.99%	15.87%
DA-SuS-1	p_f	$(1.37 \pm 0.32) \times 10^{-3}$	$(4.59 \pm 3.17) \times 10^{-7}$	$(1.33 \pm 0.84) \times 10^{-3}$
	N_{call}	2200	3352	1444
	Deviation	0%	19.53%	5.56%
DA-SuS-2	p_f	$(1.38 \pm 0.35) \times 10^{-3}$	$(4.82 \pm 3.86) \times 10^{-7}$	$(1.39 \pm 0.89) \times 10^{-3}$
	N_{call}	1083	2872	1427
	Deviation	0.73%	25.52%	10.32%
DA-SuS-3	p_f	$(1.38 \pm 0.34) \times 10^{-3}$	$(5.28 \pm 3.14) \times 10^{-7}$	$(1.41 \pm 0.79) \times 10^{-3}$
	N_{call}	1083	2127	1340
	Deviation	0.73%	37.50%	11.90%

Note: 'A(20)' denotes the dimension of Example A is 20. So as those for B and C. 'Deviation' denotes the percentage deviation of the mean values of the proposed DA-based methods and AK-SuS relative to those of the SuS results (with SuS result as the reference).

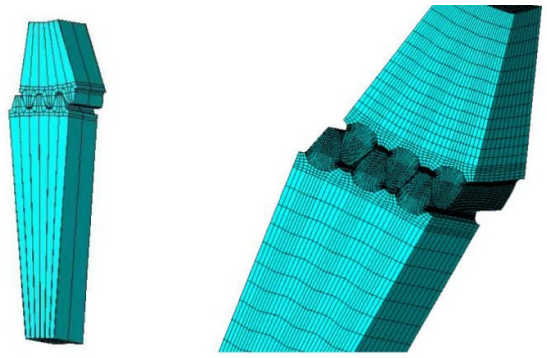
DA-SuS-1 and the improvement of computational efficiency of DA-SuS-2 and DA-SuS-3 is the same.

For Examples B and C, AK-MCS is incapable of constructing an accurate Kriging model for failure probability estimation due to the high nonlinearity or high dimensionality of these two examples, which can be reflected by the results of AK-MCS, AK-SuS and AK-IS. Despite this, the proposed DA-based methods can still provide nearly unbiased results with a slight sacrifice in statistical efficiency (with relatively higher variances). Additionally, DA-SuS-1 shows a marked improvement in statistical efficiency as Kriging is continuously updated using all generated samples with real responses. Based on DA-SuS-1, DA-SuS-2 and DA-SuS-3 further enhance the computational efficiency, with a decreasing trend in sample computation load and a slight sacrifice of statistical efficiency. The more elaborate adaptive scheme in each Markov Chain in DA-SuS-3 saves more computational burden than DA-SuS-2, which only update Kriging in each MCMC simulation level.

All the above demonstrate that the proposed DA strategy for SuS can provide a relatively unbiased estimator for failure probability estimation, especially for highly nonlinear or high dimensional problems which commonly encounter in practice while the state-of-the-art surrogate-assisted methods fails. Moreover, the three improved strategies further enhance the statistical and computational efficiency of the proposed DA strategy.

5.2. *Engineering application - wind turbine gear*

Consider a pair of wind turbine gears in figure 4 as an example, the model of which were established using 75 812 units. During the operation of an offshore wind turbine, the transmission system operates in a harsh environment, and gears face multiple failure modes, including strength, stiffness, fatigue, wear, and vibration. In the reliability assessment of the gear, static performance needs to be evaluated first. Lack of strength can cause gear teeth to break, while insufficient stiffness can cause the material to yield, affecting the wind turbine's operation. Additionally, the contact ratio of the gear significantly affects the smoothness of gear meshing. Therefore, contact ratio and mises strain were selected as the failure modes for the gear in this study.



a) Outline b) Meshing details

Figure 4. Schematic diagram of the gear finite element model.

The contact ratio is an important index for evaluating the continuity and smoothness of gear meshing. The gear transmission in wind turbines usually involves helical gears, and the overall contact ratio of a helical gear is obtained by adding the transverse contact ratio and the normal contact ratio. This value can affect the vibration characteristics of the gear, so it is used as a failure mode in this paper. The calculation equation is as follows:

$$C = \frac{1}{2\pi}\left(z_1\left(\tan\alpha'_1 - \tan\alpha_{t1}\right) + z_2\left(\tan\alpha'_2 - \tan\alpha_{t2}\right)\right) + \frac{b\sin\beta}{\pi m_n} \tag{30}$$

where z_1 and z_2 are the number of teeth of two gears, 134 and 45 respectively. α'_1 and α'_2 are the addendum circle pressure angle at the driving and driven gear, and α_{t1} and α_{t2} are the transverse pressure angle at the driving and driven gear, respectively.

The first LSF is constructed such that when the contact ratio C is lower than the threshold C_0, the gear fails.

$$g_1 = C - C_0. \tag{31}$$

In ANSYS software, the mises strain ε_e can be exported during post-processing. Mises strain ε_e is calculated using equation222, which incorporates the first, second, and third principal strain.

$$\varepsilon_e = \frac{1}{1+v}\sqrt{\frac{1}{2}\left[(\varepsilon_1-\varepsilon_2)^2 + (\varepsilon_2-\varepsilon_3)^2 + (\varepsilon_3-\varepsilon_1)^2\right]} \tag{32}$$

where v is effective Poisson's ratio. When the mises strain ε_e are greater than the allowable strain ε_0, failure of the gear occurs. The second LSF is as follows:

$$g_1 = \varepsilon_0 - \varepsilon_e. \tag{33}$$

The threshold values of the LSFs in this paper were determined based on the allowable contact stress, where $C_0 = 1.85$ and $\varepsilon_0 = 0.005$ were set in this problem.

Table 10. Gear geometric parameters.

	Input gear	Output gear
Number of teeth	134	45
Modulus (mm)	9.5	9.5
Pressure angle (°)	20	20
Spiral angle (°)	20.45	20.45
Teeth thickness (mm)	175	175

Table 11. Random variables.

Parameter	Mean value	Standard deviation
Modulus (mm)	9.5	0.95
Teeth thickness (mm)	175	17.5
Elastic modulus (MPa)	207 000	20 700
Poisson's ratio	0.254	0.0254
Equivalent torque (N * mm)	349 980 000	34 998 000

The specific gear parameters are detailed in table 10. All degrees of freedom of the output gear were constrained, and an equivalent torque of 34 9980 N * m was applied circumferentially to the input gear. Random variables in this paper include modulus, tooth thickness, elastic modulus, Poisson's ratio, and equivalent torque, the standard deviations of random parameters were set to 1/10th of their mean values based on engineering experiences, as shown in table 11. The equivalent torque denotes the output torque of the input gear. This value is related to the wind speed. It was taken as one-tenth in this paper. The range of the truncated normal distribution is selected as $[d - 3\sigma,\ d + 3\sigma]$ according to the rule of 3σ.

Similar to section 5.1, surrogate models were constructed using AK-MCS with an error of 0.01 for the two LSFs. The number of initial training samples N_{initial} is set as 21 for the two LSFs, while both the test samples N_{test} and the augmented test samples N_{au} are set as 10^4. Given that the single calculation time of the real LSFs is relatively long, only 200 more samples are generated to enrich the Kriging models. Those are 221(21 + 200) for the two LSFs.

In this experiment, each algorithm was run independently 10 times. The statistical optimization results from 10 independent runs are shown in table 12. It is evident that the surrogate model constructed with a limited number of samples is not sufficiently accurate, and this can be confirmed by the results of AK-MCS, AK-SuS, and AK-IS. However, the original DA-SuS method not only yields results with the correct order of magnitude, but also these results are highly close to those of SuS, while using fewer samples. Furthermore, the three improved strategies have also significantly enhanced statistical efficiency—this indicates that the proposed DA-SuS method and its three improved strategies can still be applied to complex practical engineering cases with implicit performance functions, even when the surrogate model fails to fit the LSF well.

Table 12. Statistical results of reliability estimation.

Example		First LSF	Second LSF
MCS	p_{f}	4.53×10^{-3}	2.71×10^{-3}
	N_{call}	1×10^{4}	1×10^{4}
AK-MCS	p_{f}	7.35×10^{-2}	6.73×10^{-2}
	N_{call}	221 (21 + 200)	221 (21 + 200)
AK-IS	p_{f}	$(9.37 \pm 0.77) \times 10^{-2}$	$(7.73 \pm 0.65) \times 10^{-2}$
	N_{call}	221	221
SuS	p_{f}	$(3.42 \pm 0.29) \times 10^{-3}$	$(2.42 \pm 0.42) \times 10^{-3}$
	N_{call}	2800	2800
AK-SuS	p_{f}	$(9.82 \pm 0.40) \times 10^{-2}$	$(7.49 \pm 0.92) \times 10^{-2}$
	N_{call}	221	221
	Deviation	2771.35%	2995.04%
DA-SuS-0	p_{f}	$(4.47 \pm 0.31) \times 10^{-3}$	$(1.32 \pm 0.45) \times 10^{-3}$
	N_{call}	2157	2098
	Deviation	30.70%	45.45%
DA-SuS-1	p_{f}	$(4.05 \pm 0.32) \times 10^{-3}$	$(1.67 \pm 0.50) \times 10^{-3}$
	N_{call}	2071	1933
	Deviation	18.42%	30.99%
DA-SuS-2	p_{f}	$(4.23 \pm 0.29) \times 10^{-3}$	$(1.82 \pm 0.47) \times 10^{-3}$
	N_{call}	1987	1871
	Deviation	23.68%	24.79%
DA-SuS-3	p_{f}	$(4.25 \pm 0.38) \times 10^{-3}$	$(2.03 \pm 0.69) \times 10^{-3}$
	N_{call}	1812	1805
	Deviation	24.27%	16.12%

6. Conclusions

The SuS algorithm with DA strategy is proposed in this paper (named as DA-SuS) to accelerate the conventional SuS with a surrogate model. The computational burden of calculating the real LSF of the rejected samples in each MCMC simulation is saved by the surrogate (Kriging model) while real LSF of each accepted candidate sample is still calculated to ensure an unbiased estimation of failure probability. The unbiasedness of the DA-SuS algorithm is not influenced by the quality of the surrogate model, but other performances, e.g. the statistical and computational efficiency, depend on the surrogate model. An illustrative example demonstrates this characteristic of DA, and indicates that the quality of the surrogate model would impact on the misjudgment error for the rejected samples by the surrogate model. This leads to a reduction in statistical efficiency due to the erroneous rejection of samples located within the failure domain.

Three improvement strategies are then introduced into DA-SuS to enhance its statistical and computational efficiency. Empirical studies have indicated that the proposed DA strategy is worthy of reference as an unbiased surrogate-assisted way to accelerate SuS. More importantly, the merit of the proposed DA strategy is significant in real engineering problems with relatively high dimension or high complexity, where the-state-of-art surrogate-assisted methods are incapable of ensuring accuracy. In general, DA-SuS-1 can obtain the best statistical efficiency when the computational resources permit in real practice. DA-SuS-3 gives the best computational efficiency while slightly sacrificing the statistical efficiency.

Note that the proposed DA strategy is illustrated via a Kriging-based MCMC simulation approach. While this concept can be extended to other surrogate-assisted MCMC methods, careful attention must be paid to ensuring the detailed balance condition. A known limitation is that DA-SuS, unlike its theoretical ideal, may still introduce a slight bias due to the limited sample size, particularly when the surrogate model fails to accurately capture the LSF. Furthermore, this work focuses exclusively on problems with a single LSF. Future work can extend the methodology to handle multiple LSFs by leveraging techniques from the generalized SuS [14].

Acknowledgment

This work is supported by the National Natural Science Foundation of China (U23A20662 and 12102125) and the Zhejiang Department of Science and Technology (2024C03255).

Appendix. Misjudgment error of generated MCMC samples

Compared to the conventional SuS, the unbiased Kriging-based method still introduces error when accepting/rejecting candidate samples. Let N_f and N_s denote the real size of samples accepted when they belong to $\Omega_\mathrm{f}=\left\{\boldsymbol{v}^i\in F_j\right\}$ and those belong to $\Omega_\mathrm{s}=\left\{\boldsymbol{v}^i\notin F_j\right\}$ in the $j+1$-th simulation level, respectively. Inspired by ESC for MCS [69], Ω_f and Ω_s can be regarded as the failure and safety region according to y_j in the $j+1$-th simulation level. $\hat{N}_\mathrm{f}$ and $\hat{N}_\mathrm{s}$ denote the corresponding values when the LSF $G(\boldsymbol{v})$ is approximated by a Kriging model, where $\hat{\Omega}_\mathrm{f}$ and $\hat{\Omega}_\mathrm{s}$ are the corresponding two region. Furthermore, let $\hat{S}_\mathrm{f}$ denote the size of candidates in $\hat{\Omega}_\mathrm{f}$ that belong to Ω_s while let $\hat{S}_\mathrm{s}$ denote those in $\hat{\Omega}_\mathrm{s}$ that belong to Ω_f. $\hat{N}_\mathrm{f}$ can then be written as

$$N_\mathrm{f}=\hat{N}_\mathrm{f}+\hat{S}_\mathrm{s}-\hat{S}_\mathrm{f} \tag{34}$$

the following confidence interval of which can be represented with confidence level α as:

$$N_\mathrm{f}\in\left[\hat{N}_\mathrm{f}-\hat{S}_\mathrm{f}^u,\hat{N}_\mathrm{f}+\hat{S}_\mathrm{s}^u\right] \tag{35}$$

where $\hat{S}_\mathrm{f}^u$ and $\hat{S}_\mathrm{s}^u$ are the upper bound of the confidence interval of $\hat{S}_\mathrm{f}$ and $\hat{S}_\mathrm{s}$, respectively. Accordingly, the maximum error can be calculated as:

$$\varepsilon=\left|\frac{\hat{N}_\mathrm{f}}{N_\mathrm{f}}-1\right|\le\max\left(\left|\frac{\hat{N}_\mathrm{f}}{\hat{N}_\mathrm{f}-\hat{S}_\mathrm{f}^u}-1\right|,\left|\frac{\hat{N}_\mathrm{f}}{\hat{N}_\mathrm{f}+\hat{S}_\mathrm{s}^u}-1\right|\right)=\varepsilon_{\max}. \tag{36}$$

Similar to that in ESC for MCS, both $\hat{S}_\mathrm{f}$ and $\hat{S}_\mathrm{s}$ due to the wrong estimation based on Kriging corresponding to the jth threshold of the jth intermediate event can be expressed using an indicator function I as

$$\begin{cases}\hat{S}_\mathrm{f}=\sum\limits_{i=1}^{\hat{N}_\mathrm{f}}I_i,\boldsymbol{v}^i\in\hat{\Omega}_\mathrm{f}\\ \hat{S}_\mathrm{s}=\sum\limits_{i=1}^{\hat{N}_\mathrm{s}}I_i,\boldsymbol{v}^i\in\hat{\Omega}_\mathrm{s}\end{cases}. \tag{37}$$

Inspired by [47], the corresponding probability P_i^{WSE} of wrong estimation satisfies

$$P_i^{\mathrm{WSE}}=P\left(I_i=1\left|\boldsymbol{v}^i\in\hat{\Omega}_\mathrm{f}\cup\hat{\Omega}_\mathrm{s}\right.\right)=\Phi\left(-\left|\frac{\widehat{g}\left(\boldsymbol{v}^i\right)-l_i}{\widehat{\sigma_g}\left(\boldsymbol{v}^i\right)}\right|\right). \tag{38}$$

This derivation can be referred to figure 3. Thus, it is evident that I_i follows a Bernoulli distribution with the following mean and variance:

$$\begin{aligned}&E\left[I_i|\boldsymbol{v}^i\in\hat{\Omega}_\mathrm{f}\cup\hat{\Omega}_\mathrm{s}\right]=P_i^{\mathrm{WSE}}\\&\mathrm{Var}\left[I_i|\boldsymbol{v}^i\in\hat{\Omega}_\mathrm{f}\cup\hat{\Omega}_\mathrm{s}\right]=P_i^{\mathrm{WSE}}\left(1-P_i^{\mathrm{WSE}}\right)\end{aligned}. \tag{39}$$

For independent Bernoulli trials such as that of MCS in ESC, $\hat{S}_\mathrm{f}$ and $\hat{S}_\mathrm{s}$ follow the following Poisson binomial distributions

$$\begin{aligned}&\hat{S}_\mathrm{f}\sim PB\left(\mu_{\hat{S}_\mathrm{f}},\sigma_{\hat{S}_\mathrm{f}}^2\right),v^i\in\hat{\Omega}_\mathrm{f},\mu_{\hat{S}_\mathrm{f}}=\sum_{i=1}^{\hat{N}_\mathrm{f}}P_i^{\mathrm{WSE}},\\&\sigma_{\hat{S}_\mathrm{f}}^2=\sum_{i=1}^{\hat{N}_\mathrm{f}}P_i^{\mathrm{WSE}}\left(1-P_i^{\mathrm{WSE}}\right)\\&\hat{S}_\mathrm{s}\sim PB\left(\mu_{\hat{S}_\mathrm{s}},\sigma_{\hat{S}_\mathrm{s}}^2\right),v^i\in\hat{\Omega}_\mathrm{s},\mu_{\hat{S}_\mathrm{s}}=\sum_{i=1}^{\hat{N}_\mathrm{s}}P_i^{\mathrm{WSE}},\\&\sigma_{\hat{S}_\mathrm{s}}^2=\sum_{i=1}^{\hat{N}_\mathrm{s}}P_i^{\mathrm{WSE}}\left(1-P_i^{\mathrm{WSE}}\right)\end{aligned} \tag{40}$$

which can be numerically approximated by a Poisson distribution with small size of samples while a normal distribution with sufficiently large size of samples as

$$\begin{aligned}&\Pr\left(\hat{S}_\mathrm{f}=k\right)\approx\frac{\mu_{\hat{S}_\mathrm{f}}^k\mathrm{e}^{-\mu_{\hat{S}_\mathrm{f}}}}{k!},k=0,1,\cdots,\hat{N}_\mathrm{f}\\&\hat{S}_\mathrm{s}\sim N\left(\mu_{\hat{S}_\mathrm{s}},\sigma_{\hat{S}_\mathrm{s}}^2\right),v^i\in\hat{\Omega}_\mathrm{s}\end{aligned}. \tag{41}$$

Given the confidence level $\alpha=0.05$, one can obtain the following the confidence intervals

$$\begin{aligned}&\hat{S}_\mathrm{f}\in\left[\boldsymbol{\Gamma}_{\hat{S}_\mathrm{f}}^{-1}\left(\frac{\alpha}{2}\right),\boldsymbol{\Gamma}_{\hat{S}_\mathrm{f}}^{-1}\left(1-\frac{\alpha}{2}\right)\right]\\&\hat{S}_\mathrm{s}\in\left[\mu_{\hat{S}_\mathrm{s}}-1.96\sigma_{\hat{S}_\mathrm{s}},\mu_{\hat{S}_\mathrm{s}}+1.96\sigma_{\hat{S}_\mathrm{s}}\right],\boldsymbol{v}^i\in\hat{\Omega}_\mathrm{s}\end{aligned} \tag{42}$$

where $\boldsymbol{\Gamma}_{\hat{S}_\mathrm{f}}^{-1}(\cdot)$ is the inverse CDF of the Poisson distribution. The maximum error in equation (36) can then be obtained by introducing the upper bounds in equation (42).

For dependent Bernoulli trials such as that of the MCMC samples in each simulation level of SuS, however, the Poisson binomial distribution is impractical as a direct approximation. There are several trials on Poisson approximation for sums of dependent Bernoulli random variables. The following non-uniform bounds by Teerapabolarn and Neammanee [70] is provided here as a reference

$$\begin{aligned}&\left|\Pr\left(W=w_0\right)-\frac{\left(P_i^{\mathrm{WSE}}\right)^{w_0}\mathrm{e}^{-P_i^{\mathrm{WSE}}}}{w_0!}\right|\\&\quad\le\min\left\{\frac{1}{w_0!},\left(P_i^{\mathrm{WSE}}\right)^{-1}\right\}\left(b_1+b_2\right)\end{aligned} \tag{43}$$

where $w_0\in\{1,2,\cdots,|\Gamma|\}$ denotes the number of dependent Bernoulli trials. Suppose that the set $B_\alpha\subsetneq\Gamma$ with $\alpha\in B_\alpha$ is chosen as a neighborhood of α consisting of the set of indices for each $\alpha\in\Gamma$, such that X_α and X_β are dependent (two dependent Bernoulli trials). It satisfies

$$\begin{cases}b_1=\sum\limits_{\alpha\in\Gamma}\sum\limits_{\beta\in B_\alpha}p_\alpha p_\beta\\ b_2=\sum\limits_{\alpha\in\Gamma}\sum\limits_{\beta\in B_\alpha\setminus\{\alpha\}}E\left[X_\alpha X_\beta\right]\end{cases}. \tag{44}$$

It is impractical to direct calculate the value of the right term of equation (43) from MCMC samples in each simulation level of SuS under the definition for dependency in equation (44). However, equation (43) still reveals the error between the real probability of sum of dependent Bernoulli trials for each MCMC simulation and the Poisson approximation.

References

[1] Der Kiureghian A 2022 *Structural and System Reliability* (Cambridge University Press)

[2] Hohenbichler M and Rackwitz R 1981 Non-normal dependent vectors in structural safety *J. Eng. Mech. Div.* **107** 1227–38

[3] Liu P-L and Der Kiureghian A 1986 Multivariate distribution models with prescribed marginals and covariances *Prob. Eng. Mech.* **1** 105–12

[4] Hasofer A M and Lind M C 1974 An exact and invariant first order reliability format *J. Eng. Mech.* **100** 111–21

[5] Zhang J and Du X 2010 A second-order reliability method with first-order efficiency *J. Mech. Des.* **132** 101006

[6] Kroese D P and Rubinstein R Y 2012 Monte Carlo methods *WIREs Comput. Stat.* **4** 48–58

[7] Schuëller G I and Stix R 1987 A critical appraisal of methods to determine failure probabilities *Struct. Saf.* **4** 293–309

[8] Ditlevsen O, Melchers R E and Gluver H 1990 General multi-dimensional probability integration by directional simulation *Comput. Struct.* **36** 355–68

[9] Ditlevsen O, Olesen R and Mohr G 1986 Solution of a class of load combination problems by directional simulation *Struct. Saf.* **4** 95–109

[10] Nie J and Ellingwood B R 2000 Directional methods for structural reliability analysis *Struct. Saf.* **22** 233–49

[11] Schuëller G I, Pradlwarter H J and Koutsourelakis P S 2004 A critical appraisal of reliability estimation procedures for high dimensions *Prob. Eng. Mech.* **19** 463–74

[12] Pradlwarter H J, Pellissetti M F, Schenk C A, Schuëller G I, Kreis A, Fransen S, Calvi A and Klein M 2005 Realistic and efficient reliability estimation for aerospace structures *Comput. Methods Appl. Mech. Eng.* **194** 1597–617

[13] Au S-K and Beck J L 2001 Estimation of small failure probabilities in high dimensions by subset simulation *Prob. Eng. Mech.* **16** 263–77

[14] Li H S, Ma Y Z and Cao Z 2015 A generalized subset simulation approach for estimating small failure probabilities of multiple stochastic responses *Comput. Struct.* **153** 239–51

[15] Betz W, Papaioannou I, Beck J L and Straub D 2018 Bayesian inference with subset simulation: strategies and improvements *Comput. Methods Appl. Mech. Eng.* **331** 72–93

[16] Liao Z, Li B and Wan H-P 2025 From likelihood to limit state: A reliability-inspired framework for Bayesian evidence estimation and high-dimensional sampling *Mech. Syst. Signal Process.* **237** 112969

[17] Li H-S 2011 Subset simulation for unconstrained global optimization *Appl. Math. Model.* **35** 5108–20

[18] Li H-S and Au S-K 2010 Design optimization using subset simulation algorithm *Struct. Saf.* **32** 384–92

[19] Ma Y-Z, Li C-X, Wang Y-Y, Zhang Z-Y, Li H-S, Ding A N and Rui X-T 2024 Robust design optimization of a multi-body system with aleatory and epistemic uncertainty *Reliab. Eng. Syst. Saf.* **245** 110029

[20] Ma Y-Z, Li H-S, Tee K-F and Yao W-X 2018 Combined size and shape optimization of truss structures using subset simulation optimization *Proc. Inst. Mech. Eng. G* **233** 2455–77

[21] Ma Y-Z, Wei J, Liu W-D, Zhi P-P, Zhao Z-Z, Xu C and Li H-S 2025 Design optimization of composite wind turbine blade using complete constrained expected improvement-subset simulation optimization *Renew. Energy* **249** 123187

[22] Song S, Lu Z and Qiao H 2009 Subset simulation for structural reliability sensitivity analysis *Reliab. Eng. Syst. Saf.* **94** 658–65

[23] Ma Y-Z, Jin X-X, Zhao X, Li H-S, Zhao Z-Z and Xu C 2024 Reliability-oriented global sensitivity analysis using subset simulation and space partition *Reliab. Eng. Syst. Saf.* **242** 109794

[24] Papaioannou I, Betz W, Zwirglmaier K and Straub D 2015 MCMC algorithms for subset simulation *Prob. Eng. Mech.* **41** 89–103

[25] Wang Z, Broccardo M and Song J 2019 Hamiltonian Monte Carlo methods for subset simulation in reliability analysis *Struct. Saf.* **76** 51–67

[26] Murray I, Adams R and MacKay D 2010 Elliptical slice sampling *Proc. of the Thirteenth Int. Conf. on Artificial Intelligence and Statistics* PMLR **9** 541–8

[27] Rashki M, Faes M G R, Wei P and Song J 2025 Asymptotic subset simulation: an efficient extrapolation tool for small probabilities approximation *Reliab. Eng. Syst. Saf.* **260** 111034

[28] Li B, Xia W and Liao Z 2025 Relaxed subset simulation for reliability estimation *Reliab. Eng. Syst. Saf.* **264** 111302

[29] Au S-K and Zhou X 2025 Adaptive proposal length scale in subset simulation *Reliab. Eng. Syst. Saf.* **261** 111069

[30] Dang C, Valdebenito M A, Faes M G R, Wei P and Beer M 2022 Structural reliability analysis: a Bayesian perspective *Struct. Saf.* **99** 102259

[31] Li J 2016 Probability density evolution method: background, significance and recent developments *Prob. Eng. Mech.* **44** 111–7

[32] Chen G and Yang D 2021 A unified analysis framework of static and dynamic structural reliabilities based on direct probability integral method *Mech. Syst. Signal Process.* **158** 107783

[33] Bucher C G and Bourgund U 1990 A fast and efficient response surface approach for structural reliability problems *Struct. Saf.* **7** 57–66

[34] Blatman G and Sudret B 2010 An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis *Prob. Eng. Mech.* **25** 183–97

[35] Meng D, Yang H, Yang S, Zhang Y, De Jesus A M, Correia J, Fazeres-Ferradosa T, Macek W, Branco R and Zhu S-P 2024 Kriging-assisted hybrid reliability design and optimization of offshore wind turbine support structure based on a portfolio allocation strategy *Ocean Eng.* **295** 116842

[36] Zhang Y, Song K, Liu D, Xiong C and Chou S 2023 A multi-mode failure boundary exploration and exploitation framework using adaptive Kriging model for system reliability assessment *Prob. Eng. Mech.* **73** 103473

[37] Wang Z, Tu Y, Zhang K, Han Z, Cao Y and Zhou D 2024 An optimization framework for wind farm layout design using CFD-based Kriging model *Ocean Eng.* **293** 116644

[38] Miao X, Zheng Z, Huang X, Ding P and Li S 2023 Reliability analysis and verification of penetration type fatigue crack *Ocean Eng.* **280** 114809

[39] Wang D, Qiu H, Gao L and Jiang C 2021 A single-loop Kriging coupled with subset simulation for time-dependent reliability analysis *Reliab. Eng. Syst. Saf.* **216** 107931

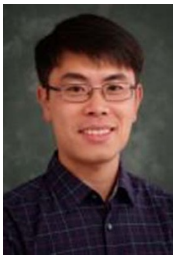
[40] Xin F, Wang P, Wang Q, Li L, Cheng L, Lei H and Ma F 2024 Parallel adaptive ensemble of metamodels combined with hypersphere sampling for rare failure events *Reliab. Eng. Syst. Saf.* **246** 110090

[41] Wang Y, Xie B and Shiyuan E 2022 Adaptive relevance vector machine combined with Markov-chain-based importance

sampling for reliability analysis *Reliab. Eng. Syst. Saf.* **220** 108287
[42] Xie B, Peng C and Wang Y 2023 Combined relevance vector machine technique and subset simulation importance sampling for structural reliability *Appl. Math. Model.* **113** 129–43
[43] Liu Y, Li L and Zhao S 2022 Efficient Bayesian updating with two-step adaptive Kriging *Struct. Saf.* **95** 102172
[44] BahooToroody A, Abaei M M, Banda O V, Montewka J and Kujala P 2022 On reliability assessment of ship machinery system in different autonomy degree; a Bayesian-based approach *Ocean Eng.* **254** 111252
[45] Li H, Soares C G and Huang H-Z 2020 Reliability analysis of a floating offshore wind turbine using Bayesian networks *Ocean Eng.* **217** 107827
[46] Zhang L, Fan Q, Lin J, Zhang Z, Yan X and Li C 2023 A nearly end-to-end deep learning approach to fault diagnosis of wind turbine gearboxes under nonstationary conditions *Eng. Appl. Artif. Intell.* **119** 105735
[47] Echard B, Gayton N and Lemaire M 2011 AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation *Struct. Saf.* **33** 145–54
[48] Yun W, Lu Z, Wang L, Feng K, He P and Dai Y 2021 Error-based stopping criterion for the combined adaptive Kriging and importance sampling method for reliability analysis *Prob. Eng. Mech.* **65** 103131
[49] Cao R, Sun Z, Wang J and Guo F 2022 A single-loop reliability analysis strategy for time-dependent problems with small failure probability *Reliab. Eng. Syst. Saf.* **219** 108230
[50] Zuniga M M, Murangira A and Perdrizet T 2021 Structural reliability assessment through surrogate based importance sampling with dimension reduction *Reliab. Eng. Syst. Saf.* **207** 107289
[51] Zhang X, Lu Z and Cheng K 2021 AK-DS: an adaptive Kriging-based directional sampling method for reliability analysis *Mech. Syst. Signal Process.* **156** 107610
[52] Song J, Wei P, Valdebenito M and Beer M 2021 Active learning line sampling for rare event analysis *Mech. Syst. Signal Process.* **147** 107113
[53] Zhang Y, Dong Y and Xu J 2023 An accelerated active learning Kriging model with the distance-based subdomain and a new stopping criterion for reliability analysis *Reliab. Eng. Syst. Saf.* **231** 109034
[54] Yu S, Wang Z and Li Y 2022 Time and space-variant system reliability analysis through adaptive Kriging and weighted sampling *Mech. Syst. Signal Process.* **166** 108443
[55] Zhao Z, Lu Z-H, Zhang X-Y and Zhao Y-G 2022 A nested single-loop Kriging model coupled with subset simulation for time-dependent system reliability analysis *Reliab. Eng. Syst. Saf.* **228** 108819
[56] Guo H, Dong Y and Gardoni P 2023 Adaptive subset simulation for time-dependent small failure probability incorporating first failure time and single-loop surrogate model *Struct. Saf.* **102** 102327
[57] Chen J, Chen Z, Xu Y and Li H 2021 Efficient reliability analysis combining kriging and subset simulation with two-stage convergence criterion *Reliab. Eng. Syst. Saf.* **214** 107737
[58] Guo H, Dong Y and Gardoni P 2022 Efficient subset simulation for rare-event integrating point-evolution kernel density and adaptive polynomial chaos kriging *Mech. Syst. Signal Process.* **169** 108762
[59] Dhulipala S L N, Shields M D, Chakroborty P, Jiang W, Spencer B W, Hales J D, Labouré V M, Prince Z M, Bolisetti C and Che Y 2022 Reliability estimation of an advanced nuclear fuel using coupled active learning, multifidelity modeling, and subset simulation *Reliab. Eng. Syst. Saf.* **226** 108693
[60] Bichon B J, Eldred M S, Swiler L P, Mahadevan S and McFarland J M 2008 Efficient global reliability analysis for nonlinear implicit performance functions *AIAA J.* **46** 2459–68
[61] Lv Z, Lu Z and Wang P 2015 A new learning function for Kriging and its applications to solve reliability problems in engineering *Comput. Math. Appl.* **70** 1182–97
[62] Sun Z, Wang J, Li R and Tong C 2017 LIF: a new Kriging based learning function and its application to structural reliability analysis *Reliab. Eng. Syst. Saf.* **157** 152–65
[63] Chen Z, Li G, He J, Yang Z and Wang J 2022 Adaptive structural reliability analysis method based on confidence interval squeezing *Reliab. Eng. Syst. Saf.* **225** 108639
[64] Yang S, Jo H, Lee K and Lee I 2022 Expected system improvement (ESI): a new learning function for system reliability analysis *Reliab. Eng. Syst. Saf.* **222** 108449
[65] Ma Y-Z, Zhu Y-C, Li H-S, Nan H, Zhao Z-Z and Jin X-X 2022 Adaptive Kriging-based failure probability estimation for multiple responses *Reliab. Eng. Syst. Saf.* **228** 108771
[66] Ma Y-Z, Li C, Li C-X, Xu B-F, Ding A N, Li H-S and Zhao Z-Z 2025 Reliability assessment of offshore wind turbine gears under multiple failure modes using deeply coupled adaptive Kriging based-generalized subset simulation *Ocean Eng.* **336** 121868
[67] Christen J A, Fox C J J O C and Statistics G 2005 Markov chain Monte Carlo using an approximation *J. Comput. Graph. Stat.* **14** 795–810
[68] Peskun P H J B 1973 Optimum Monte-Carlo sampling using Markov chains *Biometrika* **60** 607–12
[69] Wang Z and Shafieezadeh A 2019 ESC: an efficient error-based stopping criterion for kriging-based reliability analysis methods *Struct. Multidiscip. Opt.* **59** 1621–37
[70] Teerapabolarn K and Neammanee K 2005 A non-uniform bound on Poisson approximation in somatic cell hybrid model *Math. Biosci.* **195** 56–64

Yuanzhuo Ma is an associate professor at the School of Renewable Energy, Hohai University. He is engaged in research in the fields of uncertainty quantification, structural reliability, and robust design optimization, the research results of which have been applied to the reliability, robustness design, evaluation, and ensurance of large-scale wind turbines, thermal protection systems for aerospace aircraft, aviation turbine engines, multi-body systems for weapons, and electronic devices such as mobile phones. He has presided over 7 projects including the National Natural Science Foundation of China, the Rapid Support Project for Equipment Development, the Natural Science Foundation of Jiangsu Province, and the Postdoctoral Foundation. More than 40 academic papers have been published in well-known domestic and international journal, and one Chinese academic monograph has been published in the field of structural reliability and optimization. Previously worked as a reliability engineer at Shanghai Huawei Technologies Co., LTD Youth editorial board member of JRSE, reviewer for RESS and other journals.

Chuang Li is a postgraduate student at the School of Electrical and Power Engineering, Hohai University. He specializes in structural optimization of fluid machinery, with core research focusing on the development of novel subset simulation optimization methods and their practical application in fluid machinery structures. His research objects mainly include wind turbine blades and gear structures. He has published one SCI paper in the related research field, and remains dedicated to advancing efficient and reliable optimization methodologies for fluid machinery structural design.



Binbin Li is a tenured associate professor (2025-) in the Zhejiang University/University of Illinois at Urbana-Champaign Institute (ZJUI) at the Zhejiang University, International Campus. He obtained his PhD in Civil Engineering from the University of California-Berkeley in 2016. Before joining ZJU, he worked as a research associate at the University of Liverpool from 2016–2018.

Dr Li's research focuses on developing innovative statistical methods to address safety, sustainability and resilience issues of the built civil infrastructure systems including bridges, buildings, and road/rail networks. His specific interests include Bayesian system identification, operational modal analysis, infrastructural network modeling and field test, for structure and infrastructure health management and resilience assessment.